Measures of Macroproposition Construction in EFL Reading:

Summary Writing Task vs. the Meaning Identification Technique

Yuji USHIRO
University of Tsukuba
Shuichi TAKAKI
Fukushima University
Mayuko KOBAYASHI
Yusuke HASEGAWA
Shingo NAHATAME
Akira HAMADA
Yukino KIMURA
Graduate School, University of Tsukuba
The Japan Society for the Promotion of Science

Abstract

The summary writing task has been widely used in order to examine how well readers comprehend texts (Alderson, 2000). As a scoring criterion of summary protocols, previous studies have considered whether or not a reader can effectively use macrorules, which reflect the process of readers’ construction of their mental representation (e.g., Johns & Mayes, 1990; Kim, 2001). In fact, the use of macrorules is assumed to be closely related to the process of how readers construct their mental representation of a passage (van Dijk & Kintsch, 1983). However, summary task has a methodological problem in that readers’ performance in the use of macrorules reflects not only the process of reading comprehension but also their writing skills (Cohen, 1993). Therefore, we have used another psycholinguistic measure for reading comprehension that suppressed the influence of writing skills: the Meaning Identification Technique (MIT). The present study examined English as a foreign language (EFL) learners’ macrorule use with the MIT as well as in a summary writing task. Focusing on three types of macrorules (i.e., deletion, generalization, and construction rules), Experimental Study 1 showed that EFL students used all three types of macrorules when they did summary writing. In Experimental Study 2 using the MIT, the generalization and construction rules were more difficult for learners to use than the deletion rule. These two rules required learners to generate inferences for constructing the implicit main ideas of texts, whereas the deletion rule simply required learners to select explicit main ideas from texts. Comparison of the two series of experimental studies indicated that summary writing encouraged the participants to use the generalization and construction rules by requiring them to integrate pieces of
information for making the summary. Therefore, macrorule use measured by the summary protocol should be interpreted carefully, taking the effects of the summary task itself into account in terms of promoted strategic macrorule use. Furthermore, it was implied that the MIT could be superior to a summary writing task as a reading comprehension test, as far as macrorule use under natural reading conditions reflected pure comprehension of a text.

1. Introduction

1.1 Macrorules in Text Comprehension

Reading comprehension is a complicated process that is achieved through lower and higher level text processing. As a first step (i.e., lower level processing), readers conduct several processes simultaneously, such as recognizing graphemic information, accessing the lexical information, and syntactic processing (Grabe, 2009). According to van Dijk and Kintsch (1978, 1983), readers process individual words and sentences in such lower level language processing, and construct micropropositions which represent the literal comprehension of a text; the readers’ mental representation of contents explicitly described in the text. Subsequently, readers are required to grasp the gist of a text (i.e., macroproposition). To reach a deeper understanding beyond surface comprehension, they have to integrate micropropositions into a macroproposition that represents the global meaning of a discourse (i.e., higher level processing).

![Figure 1. Micropropositions integrated into a macroproposition through macrorules.](image)

Figure 1 shows a macroproposition and micropropositions related hierarchically to each other from the lower to higher levels. The macroproposition is entailed by the sequence of micropropositions. That is, readers select important micropropositions or integrate certain ones in order to construct a macroproposition (van Dijk & Kintsch, 1983). Evidence suggests that the process of macroproposition construction from micropropositions is under the control of macrorules (e.g., Brown & Day, 1983; Kim, 2001; van Dijk, 1977; van Dijk & Kintsch, 1978, 1983). Although various types of macrorules have been suggested by previous researchers, van Dijk and Kintsch (1978, 1983) identified three main macrorule types: deletion, generalization and construction. The definition of each macrorule is briefly described as shown in Table 1.

---

The Japan Language Testing Association

NII-Electronic Library Service
Table 1  
Definitions of Each Macrorule (van Dijk & Kintsch, 1978)  

<table>
<thead>
<tr>
<th>Macrorule</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deletion</td>
<td>Each proposition that is neither a direct nor an indirect interpretation condition of a subsequent proposition may be deleted.</td>
<td>Mary played with a ball. The ball was blue. ⇒“Mary played with a ball”</td>
</tr>
<tr>
<td>Generalization</td>
<td>Each sequence of propositions may be substituted by more general proposition denoting an immediate superset.</td>
<td>Mary played with a doll. Mary played with blocks. ⇒“Mary played with toys.”</td>
</tr>
<tr>
<td>Construction</td>
<td>Each sequence of propositions may be substituted by a proposition denoting a global fact of which the facts denoted by the microstructure propositions are normal conditions, components, or consequences.</td>
<td>I went to the station, bought a ticket, and took a train. ⇒“I traveled by train.”</td>
</tr>
</tbody>
</table>

Note. The italicized sentences are a part of example text. The sentences in double quotations are example macropropositions constructed from parts of the example text. Deletion reflects literal comprehension, and Generalization and Construction represents inferential comprehension (Kintsch, 1990).

So far, the use of macrorules in text comprehension has been examined mainly in first language (L1) reading research and this has already provided some valuable results (Brown & Day, 1983; Kintsch, 1990; Kintsch & van Dijk, 1978; Winograd, 1984). For example, deletion is easier to apply because it only requires readers to exclude trivial and redundant propositions from their mental representation. Although Kintsch (1990) did not describe the deletion rule explicitly, this rule could be included in the lower level processing for shallow comprehension unlike the other two macrorules which require higher level processing. Meanwhile, readers sometimes have difficulty in using generalization or construction rules because they need to integrate a sequence of detailed information into an inclusive main idea. For example, when a reader comprehends the sentence I went to the station, bought a ticket, and took a train, she or he would never achieve an understanding that the person “traveled by train” without referring to the background knowledge that all actions the person was carrying out were for travelling. Some researchers consider the generalization and construction rules as an inference generation process (Kintsch, 1990; Ritchey, 2011). The inference generation and those macrorules are similar in that they derive implicit propositions from text-based information. One of the definitions of inference was to activate text-based arguments or propositions that were not explicitly mentioned in a text (Graesser & Kreuz, 1993). What should be carefully considered is that the difficulty of
macrorule use is quite different among the three types; in particular, the generalization and construction rules are more difficult to apply than the deletion rule.

1.2 Using a Summary Writing Task to Measure Macrorule Use

Many previous L1 studies have used a summary task to examine how readers use macrorules in reading comprehension. This serves as a useful tool to test whether learners can understand the main ideas, separate relevant from irrelevant information, and organize their thoughts on a text (Alderson, 2000). For example, Brown and Day (1983) used a summary task to investigate the development of L1 readers’ ability to use macrorules. They examined the summary protocols written by L1 readers of a variety of grades in terms of macrorule use. Although the labels for the two macrorules examined in their study (i.e., superordination and invention) were a little different from those of van Dijk and Kintsch (1978), the superordination and invention rules were almost equivalent to generalization and construction rules, respectively. Their result showed that the deletion rule was easiest to use; even fifth grade children could use it. Meanwhile, the generalization (i.e., superordination) rule was more difficult to use than deletion. The construction (i.e., invention) rule was most difficult; as many as half of the L1 university students could not use it. In sum, the researchers investigated the developmental stages for using each macrorule was deletion < generalization < construction, in ascending order of difficulty.

Some second language (L2) studies also used a summary task (e.g., Cohen, 1993; Johns & Mayes, 1990; Kim, 2001; Ushiro, Nakagawa, Kai, Watanabe, & Shimizu, 2008; Ushiro et al., 2009). They assumed the use of macrorules in summary protocols as a criterion to measure students’ macroproposition construction (i.e., main idea comprehension). They showed that readers who used difficult macrorules (i.e., generalization or construction, rather than deletion) could construct elaborate macropropositions; readers who used only an easy macrorule (i.e., deletion) could construct simple macropropositions. For instance, Kim (2001) examined the three types of macrorules used by Korean learners of English as a foreign language (EFL): deletion, selection (i.e., selection of important idea), and transformation (i.e., invention of macroproposition). Although Kim used different terminology from van Dijk and Kintsch (1978), the transformation rule is about the same process as generalization and construction rules. The results showed that EFL learners could distinguish important ideas from trivial ones in L2 texts through deletion and selection rules. However, they had difficulty in using transformation (i.e., generalization plus construction) because the integration of information across sentences or paragraphs requires higher level language processing.

On the other hand, Ushiro et al. (2009) found that EFL university students could use complex generalization and construction rules when they wrote summaries in their L1; to be more precise, about half of the participants used both rules. This indicates that
the difficulty of the generalization and construction rules, unlike the deletion rule, depends on what kind of task instructions are given to the participants. For example, Ushiro et al. gave their participants an instruction to read a text carefully, paying attention to the global ideas of the text. They explained that this kind of such an instruction was given to investigate the influence of reading goals. The conclusion is that whereas the deletion rule can always be used by L2 learners, the difficulty of the generalization and construction rules is significantly affected by task instructions.

However, some L1 and L2 studies have indicated that examining macrorule use in a summary protocol is not appropriate for measuring comprehension of a text (Cohen, 1993; Winograd, 1984). For example, Winograd (1984) showed that macrorule use was a characteristic strategy accompanying summary writing; readers constructed macropropositions using macrorules only when they were required to do so by tasks. As in a typical summary writing task, the participants in his study were asked to read a passage and write a summary of it. In addition, they were assigned several tasks such as multiple choice comprehension questions on the passage. The result of a hierarchical multiple regression analysis showed that macrorule use in a summary protocol did not predict comprehension of the passage; rather, it was a significant predictor of the quality of written summaries. Following the results, he concluded that macrorule use was a kind of strategy for a summary task rather than an unconscious mental process of reading comprehension; it was strategically triggered by the task. Cohen (1993) also pointed out that writing proficiency as well as reading proficiency influence the performance of a summary task. That is, when a subject’s writing ability was not high enough to produce macropropositions, a summary task could not measure all macropropositions constructed while reading. Therefore, it cannot be said that the macrorules used in a summary protocol perfectly reflect readers’ pure comprehension of a text.

1.3 Using a Meaning Identification Technique to Measure Macrorule Use

Only a few studies have examined readers’ macrorule use without a summary task (Guindon & Kintsch, 1984; Ritchey, 2011). For example, Guindon and Kintsch (1984) investigated the construction of macroproposition while reading by using a priming task, following the methodology proposed in Ratcliff and McKoon (1978). In their study, participants read a short passage before they were presented with word pairs. There were three target word pairs in terms of the relationships between them: (a) macro pair, (b) micro pair and (c) different pair. Among the three, (a) and (b) were word pairs in the same sentence, organizing macropropositions or micropropositions, respectively, and (c) was a word pair in two separate sentences. The comparison of (a) and (b) was the closest to the current study’s interest; if readers used “macrorules” during text comprehension, (a) the “macro” pair would be recognized faster than (b). In the priming task, participants were required to judge whether the word pairs were

— 189 —
included in the same sentence or not. The result showed that L1 readers used macrorules and constructed macropropositions while reading even when they were not assigned a summary task. Although Guindon and Kintsch did not take types of macrorules into account and we cannot predict results in L2 reading, the logic in their study can be applied to the present study, i.e., if readers' judgment responses to a macrorule-based prime are faster and more accurate than another type of prime, the readers can be assumed to have constructed that kind of macroproposition.

Ritchey (2011) also investigated L1 readers' use of the generalization rule. Her participants read a passage that included a target sentence constructed through the generalization rule at the end of the passage, and the reading time was recorded. The target sentence was either consistent or inconsistent with a macroproposition sentence of the passage. The method was based on the following logic: If the reading time of the consistent target sentences was faster than that of inconsistent ones, participants constructed macropropositions using the generalization rule. The result showed that the reading time of the former was faster than that of the latter, as she predicted. In short, even when the summary task was not assigned, L1 readers used the generalization rule while reading. The limitation of Ritchey's study was that she did not examine the other types of macrorules (i.e., deletion and construction).

A set of previous findings suggest that L1 readers constructed macropropositions using macrorules while reading. However, there were two limitations: L1 studies did not focus on all three types of macrorule use, and they did not examine macrorule use in L2 reading. Therefore, as a new test method to overcome these limitations, we have chosen the Meaning Identification Technique (MIT) test to measure the ability to construct macropropositions using macrorules by L2 learners. It was developed as a reading test with high reliability and validity by Marchant, Royer, and Greene (1988). It consists of a set of passages and two types of test items: (a) consistent paraphrases and (b) inconsistent paraphrases. Item (a) is constructed by changing some words from the original text without altering the meaning of the passage, and (b) is constructed by changing some words in the original text so that the meaning of the passage is altered. In the test session, readers need to judge accurately if (a) is consistent and (b) is inconsistent with passage content. This can give evidence that they have been able to understand the text meaning beyond surface comprehension.

There are two reasons to adopt the MIT as a valid method for measuring the extent to which readers used macrorules. The first reason is that the MIT is suitable for assessing the propositional information in readers' mental representation based on a theoretical study (van Dijk & Kintsch, 1978); the examiners using the MIT do not measure examinees' comprehension of the exact words but their mental representation, preserving the meanings of what has been read (i.e., macroproposition; Royer, 1990). The MIT is based on the theoretical assumption that reading comprehension (i.e., construction of mental representation) is accomplished through a constructive process.
involving an interaction between literal understanding of text and readers’ prior knowledge (e.g., van Dijk & Kintsch, 1983). The second reason is that MIT is easier to use in a real classroom situation than other methods used in past research, such as measurement of a priming task (e.g., Guindon & Kintsch, 1984) or reading times (e.g., Ritchey, 2011). Since the MIT test has a similar form to the True-False questions that teachers often use to examine students’ understanding, it can be a familiar method for both teachers and learners. For these reasons, the present study used MIT in order to measure the three types of macroproposition construction in L2 reading under no summary condition.

In the present study, the MIT is composed of two items: (a) macroproposition sentences and (b) filler sentences. Item (a) are subsequently three types of paraphrase sentences representing a consistent macroproposition created using each macrorule, and (b) are paraphrase sentences indicating an inconsistent macroproposition. In this study, participants were presented with (a) and (b) after reading a passage and were required to judge whether the sentences were consistent or inconsistent with the passage. The logic was that if the participants judged the sentence (a) accurately, they constructed the macropropositions using each macrorule in their mental representation.

1.4 The Purpose of the Present Study, and the Research Question

As several L1 and L2 studies have indicated, the traditional methodology relying on a summary task has two serious problems: Readers’ macrorule use in the protocol is different from that in the natural reading condition, and few studies have examined the three types of macrorule use in L2 reading. The present study investigated the three types of L2 readers’ macrorule use measured by the MIT as well as conducting a summary writing task. These were then compared with each other. As a methodology for measuring readers’ pure comprehension of a text under natural reading conditions, the MIT is often seen to be superior to the summary task. Whereas macrorule use measured by the MIT reflected macropropositions constructed under the natural reading condition, one measured by the summary writing task did so under strategic reading conditions for writing the summary. In addition, the MIT seems to be an appropriate method for measuring L2 learners’ macrorule use because it is based on readers’ cognitive processes (van Dijk & Kintsch, 1983), and it can easily be applied to the L2 classroom situation.

In line with a review of previous studies, we predict that there will be a noticeable difference in EFL learners’ macrorule use between the summary writing task and the MIT, especially with regard to generalization and construction rules. Thus, the research question (RQ) was set as follows:

RQ: Does EFL learners’ performance in the three types of macrorules differ between the measures using a summary writing task and the MIT?
Two experimental studies were conducted in the present study. In Experiment Study 1, EFL learners read a text and worked on a summary task in their L1. The study was conducted to examine what kinds of macrorules were used in a summary writing task. The results of previous L2 studies using a summary writing task were not consistent in that EFL learners used generalization and construction rules, and it is highly likely that the use of the generalization and construction rules is significantly affected by task instructions. Hence, we expected to gain a clearer understanding of whether EFL university students in this study used these complex macrorules or not. In Experimental Study 2, the participants, who belonged to the same university as those in Experiment Study 1, were assigned to read the same text as the Experimental Study 1, and the use of macrorules was examined by the MIT. Results of these two studies were compared to address the RQ.

2. Experimental Study 1

2.1 Pilot Studies

Prior to Experiment 1, two pilot studies were performed to determine the four implementation conditions of the summary task: (a) the scoring criterion for a summary protocol, (b) the time required to read a passage, (c) the time required to accomplish a summary task, and (d) the word limit for writing a summary protocol. The first study was designed to address (a); the scoring criterion for the summary protocol was made through a cluster analysis with important rating of each idea units (IUs) as a dependent variable following Taniguchi (1999). A cluster analysis can identify the united group of IUs (i.e., macropropositions) which have the same degree of importance. It ensured more objective scoring compared to using the scoring criteria only based on the researchers’ judgment used in past studies (e.g., Kim, 2001). The second study was for (b), (c), and (d); it was necessary to control these factors strictly because they influence L2 and EFL learners’ reading processes (e.g., Kim, 2001; Ushiro et al., 2008, 2009).

Table 2

<table>
<thead>
<tr>
<th>Text</th>
<th>FKGL</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>FKGL</td>
<td>8.1</td>
<td>8.8</td>
<td>8.1</td>
<td>8.7</td>
<td>8.3</td>
<td>10.9</td>
<td></td>
</tr>
<tr>
<td>Token</td>
<td>357</td>
<td>360</td>
<td>345</td>
<td>368</td>
<td>385</td>
<td>309</td>
<td></td>
</tr>
</tbody>
</table>

*Note. FKGGL = Flesch-Kincaid Grade Level; FKGL was calculated by Microsoft Word 2010; Texts A, C, and E were also used in Experiment Study 2.*

Experimental passages were adopted from six passages used in reading sections
of the Society for Testing English Proficiency (STEP) 2nd Grade test (2003, 2009, 2011). They were selected by the researchers so that full text comprehension required readers to use the three types of macrorules. These passages were divided into IUs following Ikeno (1996). Since one of the passages (i.e., Plastic money) had already been analyzed in Ushiro et al. (2009), the others were analyzed by two judges independently. The agreement rates were from 95.10% to 95.51% and the mean rate was 95.69%, which showed a sufficiently high reliability. Table 2 shows the mean token and readability of these six passages.

In the first pilot study, 15 graduates and fourth-grade undergraduates majoring in English education rated the IU importance in all six passages using a 5-point Likert scale. The cluster analysis was conducted using a between-group linkage method based on the squared Euclidean distance. The cutting point was set at 3.75 and it was found that four or five clusters were composed of IUs in all the passages. Based on the results, macropropositions were identified by four of the researchers. The macropropositions were constructed by combining IUs belonging in the same cluster within each paragraph. Table 3 shows examples.

Table 3

Examples of Macropropositions Made Using Each Macrorule in Text A

<table>
<thead>
<tr>
<th>Macrorules</th>
<th>Macropropositions and IUs</th>
</tr>
</thead>
</table>
| Deletion   | "A whale fall describes the dead body of a whale and becomes a source of food for other sea creatures."
| (a) A "whale fall" is a term used by scientists (IU1)  |
| (b) to describe the dead body of a whale (IU2)  |
| (c) and become a source of food for other sea creatures (IU5) |
| Generalization | "There are three different stages for sea creatures to eat a whale fall."
| (a) that there are three different stages (IU15)  |
| (b) creatures such as crabs and fish eat the flesh of the whale (IU18)  |
| (c) and leave only the skeleton (IU19)  |
| (d) worms and other animals eat the surface of the bones. (IU21)  |
| (e) eat the oil (IU24)  |
| Construction | "A whale fall allows us to learn about the ecosystem of the deep ocean."
| (a) from studying whale falls (IU30)  |
| (b) about what happens to waste (IU42)  |
| (c) for dealing with organic waste in the future (IU46)  |

In the second pilot study, the participants were eight graduates and fourth-grade
undergraduates majoring in English education. They were assigned all six passages and asked to read them at their own pace. The reading time was recorded to determine the time required to read a passage in the experiment. After reading a passage, they wrote the two summaries in Japanese with each being limited in the number of characters they could write. One was limited to 120 characters per passage, and the other was limited to 200 characters based on Ushiro et al. (2009). There was a possibility that the 120-character limit was too short for participants to write down a summary protocol composed of information on a whole passage. In addition, the time for the summary task was recorded. Following the results of the second pilot study, the time for reading each text was set at five minutes. Also, the time required to write the summary was set at 15 minutes and the length for writing a summary protocol was set at 200 characters.

2.2 Experiment
2.2.1 Participants
A total of 26 Japanese national university students took part in the experiment. Their majors were humanities, education, or international studies. They were second- or third-year undergraduates and had studied English for more than seven years.

2.2.2 Materials
The same six experiment passages as the pilot studies were used in this study. They were attached a summary writing task with 200 character limitation.

2.2.3 Procedure
The experiment was conducted with all participants in a single session. The participants were given a booklet composed of three of the six passages. We used several types of booklets by shuffling the order and choice of the passages to be presented because it was necessary to counterbalance any effects related to the pattern of item presentation. Each worksheet was printed together with an answer column with 200 squares for the summary task. Before the experiment, the participants were asked to read a passage and then write a summary about it. The time limits were five minutes for reading and 15 minutes for summary writing. To avoid the effects of English (i.e., L2) writing ability on performance of the summary task (e.g., Kim, 2001), we asked the participants to write the summary in Japanese (i.e., L1).

2.2.4 Scoring and Analysis
The scoring criterion of the summary protocol was based on whether a participant produced a macroproposition which is identified through the first pilot study. For example, when a macroproposition in Table 3, A whale fall describes the dead body of a whale which becomes a source of food for other sea creatures was produced, it was judged that the deletion rule had been used, which deleted
unimportant IUs 3 and 4, and one point was given. Two judges rated independently whether a participant had produced each macroproposition using the three types of macrorules. The inter-rater agreement of 30% of participants’ data was 81.38%, which was sufficiently high reliability. The points of disagreement were resolved through discussion, and the scoring criteria were shared based on the discussion. Then, the same two judges rated the rest of the data. The total agreement rate was 86.33%, and the points of disagreement were resolved by discussion again.

As a whole, the average length of the written summary protocols was 93.77% (i.e., 187 characters). Hence, eight out of 78 protocols which had less than 80% of the word limit (i.e., 160 characters) and over 200 characters were excluded from the analysis because the data were considered to show that the accomplishment of the task was incomplete. The deleted data were estimated to represent 10.26% of all data. The main focus or Experimental Study 1, the mean production rates by macrorules were calculated and normalized by angular transformation because the numbers of target macropropositions were different by macrorules; the variances between variables were not equal. A one-way analysis of variance (ANOVA) was conducted with the mean production rate of macropropositions as a dependent variable. The independent variable was the macrorule type. The criteria for results interpretation was as follows: if the participants used the generalization or construction rule, the mean production rates of these rules were not different from the production rate of the deletion rule. In this case, the deletion was treated as a control condition because past studies were consistent in that L2 learners used the deletion rule.

2.3 Results and Discussion

Table 4 shows the descriptive statistics of the raw mean rates of each macrorule. It was revealed that production rates of macropropositions for all macrorules were over 50%, even for the construction rule. These high rates indicated that participants constructed macropropositions using all types of macrorules. This was fully consistent with Ushiro et al. (2009) which showed that Japanese EFL learners (about half of the participants in their study) used generalization and construction rules.

The result of the ANOVA supported this interpretation, which showed that the main effect of the macrorule types was not significant, $F(1.51, 37.75) = 0.60, p = .511, \eta^2_p = .023$. That is, the production rates of macropropositions constructed through the generalization and construction rules were not significantly different from that of macropropositions constructed using the deletion rule. Although some studies, such as Kim (2001), indicated that the generalization and construction rules were difficult for EFL learners to use, it was found that the participants in this study used the generalization and construction rules as well as the deletion rule.
Table 4
Descriptive Statistics for Raw Production Rates of Macropropositions (N = 26)

<table>
<thead>
<tr>
<th>Macrorules</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deletion</td>
<td>.58</td>
<td>.12</td>
</tr>
<tr>
<td>Generalization</td>
<td>.51</td>
<td>.27</td>
</tr>
<tr>
<td>Construction</td>
<td>.56</td>
<td>.18</td>
</tr>
</tbody>
</table>

Macropropositions constructed through the generalization and construction rules in the present study are reviewed taking the case of Text A as an example in Table 3 (see Section 2.1 for details). There are three different stages for sea creatures to eat a whale fall was made using the generalization rule; it was produced by 80.00% of the participants. To construct this macroproposition, they had to replace some detailed text information (i.e., IUs 15, 18, 19, 21, and 24) with the implicit whole information. In the summary protocols, the detailed expression in the text, crabs, fish, worms, and other animals were often replaced with inclusive expressions such as sea creatures or deep-sea animals. This indicated that the detailed information was integrated into the main information (i.e., sea creatures) which was maintained in participants’ mental representation. Another macroproposition, such as A whale fall allows us to learn about the ecosystem of the deep ocean was constructed through the construction rule. A total of 73.33% of the participants produced this macroproposition in their summary protocols. It required the participants to integrate pieces of information located separately to each other in the text such as IUs 30, 42, and 46. They omitted the detailed information IUs from 30 to 41 (i.e., A certain company made soap using one of the worms that live on whale bone) and constructed the implicit main idea: A useful chemical for the human was found from the whale fall.

The purpose of Experimental Study 1 was to examine which macrorules the EFL university students used for making a summary of the experimental passage. The results showed that the participants in this study succeeded in using both generalization and construction rules almost to the same extent as deletion rules. Based on this, Experimental Study 2 was conducted with the MIT in order to examine EFL readers’ macrorule use while engaged in reading comprehension.

3. Experimental Study 2

3.1 Participants

A total of 66 Japanese first-grade undergraduates took part in Experimental
Study 2. None of them took part in the pilot study, but they belonged to the same university as the participants in the pilot study. Their majors were engineering, social sciences, or international studies. Their reading proficiency was assessed by the reading subsection of paper-based TOEFL practice test (Educational Testing Service, 2002), which composed of 50 items with five passages \((M = 26.73, SD = 6.49, \text{Min} = 13, \text{Max} = 40)\). Their average estimated TOEFL scores were approximately 500 (ranging from 310-580).

3.2 Material

In Experimental Study 2, the MIT was adopted to examine EFL learners’ macrorule use. The task requires participants to judge whether target sentences are consistent or inconsistent with the passage content. The correct response to the presented sentence generally indicates that a participant maintains text information about the sentence in their mental representation; a correct response to a macroproposition sentence means that a participant have succeeded in constructing it in their mental representation.

<table>
<thead>
<tr>
<th>Macrorules</th>
<th>Macroproposition Sentences</th>
<th>Production Rates</th>
<th>Importance Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deletion</td>
<td>A whale fall describes the dead body of a whale and becomes a source of food for other sea creatures.</td>
<td>80.00%</td>
<td>4.64</td>
</tr>
<tr>
<td>Generalization</td>
<td>There are three different stages for sea creatures to eat a whale fall.</td>
<td>80.00%</td>
<td>4.24</td>
</tr>
<tr>
<td>Construction</td>
<td>A whale fall allows us to learn about the ecosystem of the deep ocean.</td>
<td>73.33%</td>
<td>4.10</td>
</tr>
<tr>
<td>Filler</td>
<td>The deep parts of the ocean are very cold but there are many foods.</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The MIT in Experimental Study 2 was composed of target sentences and filler sentences. The target sentences were correct macropropositions, and filler sentences were incorrect macropropositions; therefore, correct responses to the target sentences were “Yes,” while correct responses to the filler sentences were “No.” Filler sentences were made by changing macroproposition sentences to negative sentences so that the ratio of “Yes” and “No” responses in correct responses were equal. For this reason, macroproposition sentences in Experimental Study 1 were divided into target and filler sentences. Therefore, we selected three of six passages (Texts A, C, E). These three texts were carefully selected to maintain a good balance in terms of the importance
rates of IUs between macrorules. This text selection was necessary to control item difficulty of between macrorule types after macroproposition sentences were divided into target and filler sentences. It was expected that the participants' performance on the MIT would be influenced by the importance of macropropositions in terms of the task difficulty; more important macroproposition sentences can be easier to answer, regardless of macrorule type. The mean importance rates on the 5-point Likert scale were 3.85 for deletion, 3.82 for generalization, and 3.78 for construction. Table 5 shows the example macroproposition sentences in Text A. We used four target sentences allocated to one of the macrorules (i.e., deletion, generalization, construction) and four filler sentences. Thus, there were eight sentences for each passage; a total of 24 sentences in three passages.

3.3 Procedure

The experiment was administered to each participant individually. Three experimental passages were assigned to each participant in a random order. The participants read the texts sentence by sentence on a computer screen at their own pace. After reading the last sentence in each passage, one of four target sentences or four fillers for MIT appeared randomly one by one. The participants were instructed to push the red button (i.e., the "Yes" button) if they thought the sentence was consistent with the content of the passage, and to push the blue button (i.e., the "No" button) if they thought the sentence was inconsistent with the content of the passage. After finishing the MIT, they continued to read the other passage and answer the MIT, which was repeated three times. All the participants accomplished all assignments within 30 minutes. As for the criteria for interpreting results and discussing the RQ, if the summary task had little effect on macrorule use in EFL reading, it was predicted that the correct response rates of each macrorule would show no differences, similar to the result of Experimental Study 1. On the other hand, if the summary task had an effect on macrorule use, it was expected that there would be a significant difference between the correct response rates of each macrorule; results different from Experimental Study 1 would be obtained. All sessions were conducted with computers in order to avoid re-reading the passages too much and taking too much time to answer the one item; the participants were not allowed to use extra time strategically for accomplishing the task. The results of this experimental study using the MIT were compared to those of Experimental Study 1 using a summary writing task.

3.4 Scoring and Analysis

The responses to the MIT were scored according to macrorule types. The correct response rates to sentences for each macrorule were calculated. Some missing data (1.00%) caused by technical errors were removed from the analysis. A one-way ANOVA whose dependent variable was correct response rates was conducted. The
independent variable was the macrorule type (i.e., deletion, generalization, and construction).

3.5 Results and Discussion

Table 6 shows the descriptive statistics of mean correct rates for each macrorule. The results showed that correct rates for all macrorules were over 80%, even for the construction rule. These high rates indicated that the participants constructed macropropositions using all types of macrorules. The ANOVA showed that the main effect of the macrorule type was significant, $F(2, 132) = 11.96, p < .001, \eta^2_p = .153$. Since the main effect of macrorule type was significant, a post-hoc test was conducted using Bonferroni correction. As a result, the correct rate of deletion was higher than both generalization and construction ($p < .001$ each). The correct rates of generalization and construction were not significantly different ($p = .279$).

<table>
<thead>
<tr>
<th>Macrorule Types</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deletion</td>
<td>.94</td>
<td>.12</td>
</tr>
<tr>
<td>Generalization</td>
<td>.86</td>
<td>.17</td>
</tr>
<tr>
<td>Construction</td>
<td>.81</td>
<td>.20</td>
</tr>
</tbody>
</table>

Table 6

Descriptive Statistics for Correct Response Rates to Macroproposition Sentences ($N = 66$)

Figure 3. Mean correct response rates to macroproposition sentences.

We conducted the MIT to measure the readers' macrorule use process while reading without production in the summary task. The correct response data to the MIT showed how learners used macrorules to construct macropropositions. The correct response rates for MIT indicated that all types of macrorules were used without the summary task at high rates. However, the mean correct rates for generalization and construction were lower than that for the deletion rule. These results demonstrated that generalization and construction rules were more difficult compared to the deletion rule when the summary task was not assigned, which was different from the results of Experimental Study 1 using the summary task.

A valid reason for this may be the different processes included in the use of each macrorule. The deletion rule was applied for text-based level comprehension, whereas generalization and construction rules were used for more global level comprehension (van Dijk, 1977). The process of using the deletion rule consists of selecting necessary propositions and removing unnecessary ones to construct macropropositions. For instance, the target sentence showing macropropositions made by the deletion rule ($A$
whale fall describes the dead body of a whale and becomes a source of food for other sea creatures) which is composed of IUs 1 (A "whale fall" is a term used by scientists), 2 (to describe the dead body of a whale), and 5 (and become a source of food for other sea creatures). In order to respond to this macroproposition sentence correctly, the participants had only to delete the trivial information, IUs 3 (that has fallen) and 4 (to the ocean floor); therefore, the presented macroproposition almost reflected the explicit text information as it was. On the other hand, the process of using the generalization and construction rules involved integrating a sequence of propositions or replacing individual propositions with more general statements. In other words, the participants were required to produce implicit macropropositions by integrating or replacing as well as comprehending propositions in a text. They were required to collect pieces of explicit information and integrate them to construct implicit information. In sum, Experimental Study 2 using the MIT showed results consistent with theoretical studies. The reasons why generalization and construction rules were used to the same extent as the deletion rule in the summary protocol are discussed in the next section.

3.6 Comparison Between Experiment Studies 1 and 2

Both of Experimental Studies 1 and 2 revealed that EFL learners constructed macropropositions using the deletion, generalization, and construction rules. However, the results of the studies were partially different in terms of difficulty in using generalization and construction rules compared to the deletion rule. Specifically, the participants in Experimental Study 1 used the generalization and construction rules to the same extent as the deletion rule, but those in Experimental Study 2 used these two rules to a lesser extent than deletion. There are two possible reasons for the differences between the experiments: (a) EFL readers' macrorule use while reading was limited, unlike in the summary task after reading, and (b) EFL readers' macrorule use was facilitated by the summary task.

The first reason (a), can be attributed to the lack of cognitive resources for using macrorules while reading. Kintsch (1990) showed that unskilled readers used their cognitive resources on the lower level processing (e.g., literal comprehension) and did not have spare resources for generating inferences (i.e., the generalization and construction rules) while reading. Whereas the MIT reflected macropropositions that were constructed while reading, the summary task reflected ones which were made during summary writing as well. Thus, it was possible that EFL learners who did not use generalization and construction rules efficiently in the MIT succeeded in using these rules in the summary task when cognitive resources remained. However, this possibility is not valid because summary writing is an activity which demands a lot of cognitive resources; the participants given the summary task had to allocate their resources to completing the summary task. The task requires readers to engage in
several processes, such as separating relevant and irrelevant information, as well as writing the summary (Alderson, 2000). Therefore, the first possibility (a) is ruled out for interpreting the different results between the two tasks.

As for the second possibility (b), involvement in a summary task helped the learners to reorganize the understood information. That is, the summary task worked well to provide instructions that facilitated the use of complex macrorules. This possibility is supported by Winograd (1984) claiming that the macrorule use was a typical strategy for a summary task. By writing a summary, EFL readers were required to reconstruct the comprehended information in their mind, which encouraged them to use generalization and construction rules to the same degree as the deletion rules in Experimental Study 1. On the other hand, since the participants were not given such a reconstruction process, the generalization and construction rules were difficult in Experiment Study 2 with the MIT. Kintsch (1990) categorized the generalization and elaboration (i.e., construction) rules as inference processes. Generation of certain types of inference was strategically promoted by reading goals or reading instructions (Magliano, Trabasson, & Graesser, 1999). Thus, even when enough cognitive resources were available while reading, it was logical that generalization and construction inferences that were not easily generated automatically (i.e., without writing a summary) were in fact due to the summary task.

In conclusion, the answer to the RQ followed our predictions; EFL learners’ macrorule use measured by the MIT was different from that measured by the summary task. That is, the generalization and construction rules were more difficult than the deletion rule when the learners were not given a summary writing task, although they constructed macropropositions using macrorules. This was because the summary task encouraged readers to use complex macrorules. This finding is remarkable because it showed that the summary task used in many previous studies reflected macrorule use, not only in reading comprehension but also in the summary task.

4. Conclusion

4.1 Major Findings

Previous L1 and L2 studies adopted a summary task to measure text comprehension by assessing readers’ use of macrorules in their summary protocol. However, the summary task had a fatal inadequacy; macrorule use in a summary protocol did not reflect pure text comprehension. Therefore, the present study measured EFL readers’ complex macrorule use with the summary task and MIT, and the results were compared with each other. The results of Experimental Study 1 showed that EFL university students used generalization and construction rules to the same extent as the deletion rule when they were given a summary task. That is, according to the traditional measure of macrorule use, EFL readers constructed all
three types of macropropositions. Experimental Study 1 using the summary task showed no difference among three types of macrorules. In contrast, by utilizing the MIT as another measurement of macrorules without involving a writing process, Experimental Study 2 showed different results from those of Experimental Study 1. The results indicated that the frequency of generalization and construction rule use was lower than that of the deletion rule. That is, construction of implicit macropropositions using rather complex macrorules was difficult for EFL university students when they were not given a summary task.

Putting the results of the two experimental studies together, an important finding was obtained: EFL learners can use the deletion rule routinely and it is relatively easy for them to distinguish important ideas from trivial ones in the text even while they were reading. Meanwhile, the construction of macropropositions using the generalization and construction rules was difficult while reading, but was promoted by the summary writing process. In terms of the lower and higher level processing, the generalization and construction rules were recognized as a certain type of inference generation process which is one of higher level processing (Grabe, 2009). The present results were no exception to this general assumption as far as the MIT results are concerned. However, EFL learners succeeded in using complex macrorules in summary writing. The reason for this is probably that the summary task encouraged learners to construct implicit macropropositions using complex macrorules by requiring them to combine comprehended and explicit text information.

4.2 Suggestions for Further Research

Two suggestions can be made for future research based on the present study. First, Experimental Study 1 did not examine exactly how a summary task helped EFL learners to use complex macrorules. To reveal the task effect on the macrorule use, the mental processes in writing a summary could be explored by interviews or the think-aloud method. Think-aloud protocols during summary writing may reveal how the task instruction facilitates the construction of macropropositions using complex generalization and construction rules. Second, a series of experiments in our study found a difference between the MIT and the traditional summary-writing-based measures of reading comprehension. In particular, the MIT measure was more sensitive to readers’ use of generalization and construction rules. Therefore, it is advisable to develop a new criterion to assess reading comprehension of main ideas more precisely. The present study suggests the practicality of the MIT as a new tool for measuring reading comprehension, as long as apparatus for a computer-based test is available. Lastly, the pedagogical implications of the present study should be noted. Teachers should know that the main ideas written in students’ summaries are not always an indication of accurate reading comprehension; they might instead be the result of good writing skills rather than reading skills.
Acknowledgement

This research was conducted by Tsukuba Reading Research Group and was partially supported by a Grant-in-Aid for Scientific Research (B) (No. 25284099), from 2013 to 2016. We are grateful to all of the graduate and undergraduate members of our research group for their cooperation on this study. We would like to thank two anonymous reviewers for their constructive comments.

References


Taniguchi, A. (1999). Concrete information that enhances comprehension and memorization of a text [Bunsho no rikai to kioku wo sokushin suru gutaiha joho]. Tokyo, Japan: Kazamashobo.


