The Reading Matrix: An International Online Journal
Volume 24, Number 2, September 2024

The Process of Coherence and Cohesion Judgment Tasks of Japanese University EFL Learners: Evidence from Eye-Tracking

Ken Fujita
Aichi Gakuin University

Mitsuo Ishida
Aichi Gakuin University

ABSTRACT

Readers should construct a coherent discourse during reading comprehension. The ability to build coherence has been examined using coherence and cohesion judgment tasks. Although eye-tracking studies have been conducted on building coherence or processing cohesion among native language users, few such studies have been conducted with second language learners. This study investigated the process of conducting a task using eye-tracking with 30 Japanese university students who were asked to judge the coherence between two sentences on a screen. The 24 stimulus items consisted of four conditions (with or without coherence and cohesion), and six items were assigned to each condition. The fixation duration in some areas and frequencies of the regressions were measured. The results showed that the difference in fixation duration between conditions was not statistically significant, but the frequencies of regressions were significantly different, with incoherent–cohesive pairs having more regressions than incoherent–incohesive pairs. Thus, the study revealed that readers might make more inferences because of the contradictory information of the condition, which partly supports previous research on native participants.

INTRODUCTION

When readers comprehend a discourse, they must decode what is written and grasp the word meanings from their mental lexicons, together with grammatical and syntactic knowledge. This is called the bottom–up process. However, the top–down process is also necessary, where they are required to establish coherence between sentences and paragraphs using cohesive devices like discourse markers, making inferences, and activating world knowledge. Good readers allow bottom–up and top–down processes to interact, constructing a situational model of the passage. Since these processes are invisible in their minds, they must be visualized to study the cognitive reading mechanism.

Readers comprehend a passage by constructing coherence as well as by using linguistic cohesive ties, such as connectives and lexical cue phrases. Fujita (2019, 2020, 2021a, 2021b) ensured that the efficiency of coherence and cohesion judgment tasks influences the outcome of

reading as a component skill, alongside vocabulary and grammatical–syntactic abilities. However, the processes of coherence and cohesion judgment must be carefully examined in future studies.

In this study, the processes of coherence and cohesion judgment tasks were visualized by following readers' eye movements. The authors analyzed the fixation durations before and after cohesive devices, which are regarded as the time necessary to process and understand each sentence component. Simultaneously, the eye movements of the regressions were observed. The results provide a picture of reading processes that have implications for reading instruction.

Coherence and Cohesion Judgment in Reading Comprehension

In this section, studies on coherence and cohesion judgment tasks are reviewed, and the research questions of this study are presented.

The ability to integrate the meanings of two sentences demands an understanding of the roles of coherence and cohesion. Li and D'Angelo (2016) emphasized that the difference between coherence and cohesion should be considered. Therefore, these two constructs must be measured separately, although they are often treated alike. Coherence implies that the written text is not contradictory as a unit; a coherent text allows readers to build a situation model in their mind by generating inferences and using world knowledge. Through cohesion, linguistic devices clarify coherence. Cohesive devices are classified into lexical cues (e.g., "that's why," "as a result"), anaphoric devices (e.g., demonstrative pronouns, personal pronouns), logical connectives (e.g., "because," "though"), and so on (Sánchez & García, 2009). In short, support for coherence through linguistic discourse markers can be defined as cohesion (Graesser, McNamara, & Louwerse, 2003; Nahatame, 2017).

Studies on cohesion's role in promoting readers' comprehension have been conducted in both L1 (Cain & Nash, 2011; Geva & Ryan, 1985; Sánchez & García, 2009; Sanders & Noordman, 2000) and L2 contexts (Crosson & Lesaux, 2013; Degand & Sanders, 2002; Demel, 1990; Geva, 1986, 1992; Jonz, 1987). The results have indicated that cohesive devices enabled readers to comprehend the text better, although there were some differences in participants' ages, types of cohesive ties, genres of texts, tasks for measurement, and so on. However, these studies paid little attention to distinguishing between coherence and cohesion and to treating them as different constructs.

Ferstl (2015) proposed a coherence judgment task that separates coherence from cohesion. The task was quite easy: the participants judged whether the two-sentence pairs had coherence. The accuracy and reaction time for each item were measured as task achievement indicators. The stimuli comprised a combination of coherence and cohesion. The items were then divided into four conditions: coherent–cohesive items (conditions C–C), coherent–incohesive ones (conditions C–I), incoherent–cohesive ones (conditions I–C), and incoherent–incohesive ones (conditions I–I). For example, the sentence pair "*Mary's exam was about to start. / Therefore, her palms were sweaty.*" belonged to Condition C–C, because the two sentences have coherence in terms of semantics, and, simultaneously, they are tied by the connective "therefore" and the pronoun "her." If another item is "*Mary's exam was about to start. / Some friends had remembered the birthday.*," the sentence pair has neither coherent nor cohesive devices. This is why the latter could be classified as Condition I–I. The other conditions were created in the same manner. The prediction of the task's results was that coherence would turn out to be obvious in the presence of the cohesive device, while the judgment was likely to be deterred by the contradictory information that the incoherence was connected with the cohesive tie. In summary, the items of Condition C–C could

be processed with relative ease compared with those of Condition C–I. In addition, dealing with Condition I–C was more difficult than dealing with Condition I–I. This is how the coherence judgment task was first created, especially for testing those with reading disabilities.

Ferstl and von Cramon (2001) conducted a coherence judgment task using fMRI before brain imaging. The participants were 24 German university students who were given a coherence judgment task in their native language German. The task consisted of 120 items, with 30 items for each condition, and the entire test was divided into four test lists. Each item's accuracy was measured by the correctness of each judgment. Reaction time was defined as the time spent between the end of reading the two-sentence pair and the judgment of coherence, indicated by typing on the keyboard. Consequently, the correctness rate for the entire test exceeded 90%, indicating a ceiling effect. The average reaction time for each condition was compared using a 2 x 2 analysis of variance (ANOVA) with the factors of coherence and cohesion. The interaction between coherence and cohesion was found to be statistically significant ($F(1,23) = 20.6$, $p < .001$). Cohesive ties facilitated the comprehension of coherent sentence pairs, while they made the detection of the inconsistent condition more difficult. Here, cohesive ties implied a false connection between incoherent sentence pairs. Thus, the results supported the predictions of the coherence and cohesion judgment tasks.

Hamilton (2011) attempted to explain reading comprehension using various component sub-skills: word decoding, vocabulary, working memory, and coherence judgment. Reading comprehension ability was examined on the basis of these components' interactions. The participants comprised 52 British university students. The coherence judgment task was created using the method proposed by Ferstl (2015), and four test lists of 96 items with 24 items for each condition were presented. Accuracy was measured by the correctness of the response. Reaction time was defined as the time between the presentation of the target sentence and the end of the judgment, which was transformed into a residual with the number of letters partialed out. A two-way ANOVA with the factors of coherence and cohesion was conducted. Regarding the correctness, the interaction between coherence and cohesion reached a level of significance ($F(1,51) = 12.73$, $p = .001$). Hence, cohesive devices increase judgment accuracy. Additionally, as for the reaction time, the interaction between coherence and cohesion was significant ($F(1,51) = 18.07$. $p < .001$). These results lead to the conclusion that cohesive ties promote the processing of judgment in the coherent context but delay it in the incoherent condition. This accords with the results of Ferstl and van Cramon (2001).

Fujita (2019) conducted the first preliminary study of Japanese university students on coherence and cohesion judgment tasks based on Hamilton (2011). The participants, consisting of 15 Japanese university English as a Foreign Language (EFL) learners, were given a task comprising 96 items, with 24 items for each condition. Correctness and reaction time were measured in the same way as in Hamilton (2011). The results showed that neither correctness nor reaction time was statistically different. However, when effect sizes were compared, the tendency of cohesion to assist judgment in coherent sentence pairs was apparent. The influence of cohesive ties on incoherent pairs, however, was not discovered, which contradicts the results of previous studies. The participants might have prioritized the judgment without making inferences brought by cohesive ties in an incoherent context. This is why the difficulty did not increase in Condition I–C for Japanese students, although this phenomenon was observed in the native students' experiments.

Furthermore, this topic was examined in Fujita's study (2021b), which had a higher number of Japanese university EFL learners (100). The same task as in Fujita (2019) was given to them in

both English (L2) and Japanese translations (L1). A two-way ANOVA with the factors of language and condition was performed, with the dependent variables being correctness and reaction time. The results showed that the interaction between language and condition was statistically significant for both correctness and reaction time (for correctness, F (2.52, 229.69) = 15.82, p <.001, η^2 = .017; for reaction time, F (2.65, 241.02) = 57.15, p <.001, η^2 = .043). Moreover, the main effects of language and condition were significant. The Japanese task was judged more correctly and promptly than the English one, and cohesive ties strengthen the coherence between the two sentences both in English and Japanese. However, in an incoherent context, cohesion did not deter the judgment process for the two-sentence pairs. In contrast, the judgment of Condition I–C was faster than that of Condition I–I, which again differed from the results of native participants in previous literature. Further investigation is required to reach a conclusion focusing on the processing procedure.

Eye-tracking Studies in Reading Comprehension

Eye-tracking has often been used in research on discourse processes involving more than two sentences. According to a review by Nahatame (2022), eye-tracking naturally detects the course of reading without interrupting it, while the methods of think-aloud or self-paced reading through a moving window do not. In addition, eye-tracking yields various numerical data at certain stages of the reading process. Staub and Ryner (2007) pointed out that few eye-tracking studies have been conducted on readers' discourse processing and that these lines of experiments will prevail widely in the future. This is because the discourse process is a complicated procedure in which readers build a representation of characters and events, or capture the relationship between them, by making inferences based on lower processes, such as word recognition and syntactic parsing. In addition, some eye-tracking studies on the process of anaphor resolution and elaborative inference generation have been reviewed (Staub & Rayner, 2007).

There are several studies on eye-tracking studies on building coherence or processing cohesion among native language users. Some use the paradigm of inconsistency detection between the target word and the information around it. Others introduce various kinds of texts, such as those with or without coherence and those with or without cohesion.

However, few eye-tracking studies have been conducted with second language learners. Moreover, the aims and experimental designs of these studies differ. One study focused on the cohesive devices of connectives and coherence-building (van den Bosch, Segers, & Verhoeven, 2018), whereas others examined the reading process of coherence construction with inconsistency detection (Hessel, Nation, & Murphy, 2021; Hessel & Schroeder, 2020; Ushiro et al., 2016). The characteristic of second language research is that participants' language proficiency varies greatly; therefore, individual differences in decoding, vocabulary knowledge, and grammatical–syntactic parsing often mediate the results. For example, van den Bosch et al. (2018) investigated the eye movements of 46 elementary school children (L1 Dutch: 27; L2 Dutch: 19) in Holland. The fixation durations per word of L2 readers were reduced significantly when the two sentences were connected with a cohesive tie, "because." This effect was more apparent for participants with lower syntactic knowledge.

These studies tracked the eye movements of readers for the target words or areas around them. Eye-tracking data included fixations on the eyes or regressions from them, with time duration or frequencies measured. These data were then compared between texts or conditions on the assumption that the reading process was reflected in eye movements. However, the differences

in such movements depended on the researchers' interpretations. Hence, data triangulation is necessary when other methods, such as reaction time or introspective reflection, are used simultaneously.

Current Study

As aforementioned, Fujita (2021b) investigated the correctness and reaction time of coherence and cohesion judgment tasks in L1 and L2 with 100 Japanese university EFL learners. The results showed that cohesion promoted the comprehension of coherence in a coherent context since the correctness and reaction time was better in Condition C–C than that in Condition I–C, which supported the results of native language users. The cohesive ties in the incoherent condition did not disturb judgment, on which effect was reported in native speakers' research (Ferstl & von Cramon, 2001). In contrast, Fujita (2021b) revealed that cohesion did not change the accuracy rate in Condition I–C; rather, it accelerated the speed of judgment in an incoherent context in both L1 and L2. These results indicate the need for further investigation of the processing procedure using eye-tracking methodology.

Based on a series of previous studies, the current study focused on the following two research questions (RQs) to shed light on the processes of coherence and cohesion judgment using eye-tracking methodology.

1. What is the process through which Japanese EFL learners judge coherence under coherent conditions? How do cohesive ties affect judgments?
2. What is the process through which Japanese EFL learners judge coherence under incoherent conditions? How do cohesive ties affect judgments?

To address these questions, participants' eye movements were tracked during tasks. The results clarify whether the Japanese EFL learning process differs from that of native language users. Additionally, the results are expected to have practical implications for classroom English teaching.

METHODOLOGY

Participants

Thirty Japanese university EFL learners with sufficient eyesight to see letters on the display monitor were selected for this study. At the time of the experiment, all of them had studied English for about 10 years: from 5th grade in elementary school to 1st, 2nd, or 3rd year at university. Before the experiment, the participants were informed of the research ethics, including the aim, procedure, and data publication. All the participants provided consent for participation. The accuracy rate of two students did not exceed 66.7% under any of the four conditions. Therefore, they were excluded, and statistical analysis was conducted for 28 participants ($N = 28$).

Materials

A total of 24 items, with six categories in each condition, were used as the coherence and cohesion judgment tasks. They were selected from the 96 items in Fujita (2021b), with an accuracy

of more than 80% in previous research. In addition, the second sentence for each item differed. To control the width of the area measured by the eye-tracking device, the cohesive ties connecting two sentences were made up of three words. Regarding the second sentences, the areas of "subject + verb" (in the case of copula be, "subject + be + adjective") had four to five words, and the final parts of the sentence were limited to three to five words. Participants were randomly assigned to the 24 test items. A list of the items is provided in the appendix. Since the average accuracy score of the test had almost reached the maximum (over five out of six points), reliability was estimated by the correlation of the reading times of all 24 items divided into two groups of odd- and even-numbered items, which turned out to be 0.67.

Apparatus, Presentation, and Procedure

The eye movements of the participants while reading the items were tracked using the Talk Eye Free apparatus (Takei Scientific Instruments Co., LTD.). The time series were recorded at a sampling rate of 30 Hz, which was sufficient to detect real-time eye gaze in the areas of three to five words, since the average fixation per word for native readers is 225–325 ms (Conklin, Pellicer-Sánchez, & Carrol, 2018). A chinrest was used to fix the participants' head positions. At the beginning of the experiment, calibrations were performed between the nine dots on the screen and the positions of eye fixations. Synchronization between the recording of eye movements and the onset of stimuli was carried out using the psycholinguistic software PsychoPy (Peirce & MacAskill, 2018) and serial communications with a baud rate of 19200 bps.

The items of the coherence and cohesion judgment tasks were presented with white letters (R:G:B = 255:255:255) on a gray background (R:G:B = 127:127:127). The monitor was 32 inches wide (1920 x 1080 pixels). The participants sat 80 cm away from the screen, and the sentences were presented in monospace font (terminal, 32 pixels) with the angle of vision per letter at 0.83° (height) x 0.42° (width). The first sentence for each item was presented on the left side of the screen. The subject of the second sentence was presented in the middle of the screen. In the cohesive tie condition, the subject following the cohesive tie was placed at the center of the screen. The two sentences were written on a single line. This presentation was expected to minimize measurement errors owing to the sentence positions.

Participants were asked to determine whether the two sentences were coherent and to press either button as soon as possible. If they judged the sentences as coherent, they pressed the left button; otherwise, they pressed the right button. The four-item practice session was conducted before the start of the 24-item experiment.

Data Collection and Analysis

The time-series data of eye-tracking were valid only when the participants could judge correctly, and the items with inaccurate answers were excluded from the analysis. Three regions were included in the second sentence: a cohesive phrase made up Region 0 (R0); the following "Subject + V (in the case of copula be, Subject + be + adjective)" was placed in Region 1 (R1); and the final part of the second sentence became Region 2 (R2). The fixation duration in each region and the frequency of regression were measured by tracking the eye movements.

Since each condition had only six items, data corrections were implemented. First, two participants who answered fewer than three of the six items, even in one condition, were excluded from the analysis. When the average fixation duration of all the valid data per condition was

calculated, the average score was winsorized; that is, the scores surpassing $M \pm 2SD$ were changed to $M \pm 2SD$ (Sugiura & Yamashita, 2011). Then, the two scores of the different conditions in R1 and R2 were compared according to the RQs. The frequency of regression was calculated for each participant for each item. The average frequencies of all the participants in each condition were winsorized, and the scores were statistically compared according to the RQs. Statistical analysis was conducted using the web-based statistical tool LangTest (Mizumoto & Plonsky, 2016).

Prediction of the Results

The results of previous studies (Fujita, 2019, 2020, 2021b) showed that cohesive ties strengthen the coherence of the two sentences; thus, the accuracy rate was higher, and the reaction time became shorter in the coherent condition with cohesive devices; that is, in Condition C–C than in Condition C–I. Therefore, the eye-tracking data revealed that the fixation duration in the second sentence was shorter, and the frequency of regression was lower in Condition C–C than in Condition C–I.

In contrast, Fujita (2019, 2020, 2021b) clarified that in an incoherent context, the accuracy rate was the same for Condition I–C and Condition I–I. Regarding the reaction time, contrary to the results for native speakers, it became shorter, rather than longer, under Condition I–C than under Condition I–I. Studies with native-speaker participants indicated that cohesive ties in an incoherent context deterred the judgment process because of the contradictory information of cohesion in an incoherent setting. These inconsistent results are thought to bring about ambivalent conclusions in the current study. If the results supported the results of the native speakers, the fixation durations would be longer and, simultaneously, the frequencies of regressions would be higher in Condition I–C than in Condition I–I. The reverse results would occur if they were in accordance with Fujita (2021b).

RESULTS

Descriptive Statistics

Table 1 displays the fixation durations in R0, R1, and R2, together with the combined times in R1 and R2 (R1 + 2). Since R0 represents the area of cohesive phrases, it appears only under Condition C–C and Condition I–C. The average fixation duration in R0 was the same for the two conditions (M = 0.85 s). That in R1 was from 1.18 s to 1.43 s, and that in R2 was from 3.34 s to 3.59 s. For total R1+2, the average fixation time varied little from 4.63 s to 4.93 s.

Table 1. Means and Standard Deviations for Eye Fixation Durations(s)
(Regions by Conditions)

Condition	Region 0		Region 1		Region 2		Region 1+2	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
C-C	0.85	0.37	1.29	0.43	3.38	1.50	4.68	1.78
C-I			1.18	0.36	3.46	1.37	4.63	1.49
I-C	0.85	0.35	1.31	0.50	3.59	1.56	4.93	1.90
I-I			1.43	0.46	3.34	1.16	4.78	1.36

Note: The four conditions are C–C: coherent cohesive, C–I: coherent incohesive, I–C: incoherent cohesive, and I–I: incoherent incohesive. N = 28.

Table 2 shows the average frequencies of the regression and the average number of correct answers out of six for each condition. In coherent conditions, the average frequencies were 0.82 (Condition C–C) and 0.60 (Condition C–I). However, in the incoherent setting, they were 1.05 (Condition I–C) and 0.70 (Condition I–I). The number of correct answers exceeded 5.5 points out of 6.

Table 2. Means and Standard Deviations for Frequencies of Regressions and Correct Answers (by Conditions)

Condition	Regression		Correct Answers	
	M	*SD*	*M*	*SD*
C-C	0.82	0.68	5.50	0.64
C-I	0.60	0.42	5.57	0.63
I-C	1.05	0.71	5.50	0.64
I-I	0.70	0.49	5.61	0.63

Note: The four conditions are C–C: coherent cohesive, C–I: coherent incohesive, I–C: incoherent cohesive, and I–I: incoherent incohesive. *N* = 28.

Analysis of Results in Coherent Condition (RQ1)

The reading processes of the second sentence under coherent conditions (Condition C–C and Condition C–I) were first investigated. The difference in the average fixation duration between Condition C–C and Condition C–I was examined statistically using the Wilcoxon signed-rank test because the data were not normally distributed. The results of the test showed that the difference in the average fixation duration in R1 and R2 was not statistically significant, with small effect sizes (R1: $z = 1.02$, $p = .307$, $r = .19$; R2: $z = 0.91$, $p = .361$, $r = .17$).

Next, the difference in the frequencies of regression between Condition C–C and Condition C–I was tested with the Wilcoxon signed-rank test because the data were not normally distributed. The difference was not statistically significant, with a small effect size ($z = 1.10$, $p = .271$, $r = .21$).

In summary, the coherence and cohesion judgments in Condition C–C and Condition C–I were not different from each other when tested through eye-tracking.

Analysis of Results in Incoherent Condition (RQ2)

To address RQ2, the processes of coherence and cohesion judgment under incoherent conditions (Condition I–C and Condition I–I) were examined. The difference in the average fixation duration between Condition I–C and Condition I–I was checked statistically using the Wilcoxon signed-rank test, without the data distribution of normality. The results of the test indicated that the average fixation durations in R1 and R2 did not reach a significant level, with small effect sizes (R1: $z = 1.52$, $p = .127$, $r = .28$; R2: $z = 0.46$, $p = .646$, $r = .09$).

Finally, the difference in the average frequencies of regression between Conditions I–C and Condition I–I was tested using the Wilcoxon signed-rank test because of the non-normal data

distribution. The results revealed that the difference in the average frequencies of regression between the conditions was statistically significant after Bonferroni adjustment ($p < .025$), with a medium effect size ($z = 2.43$, $p = .014$, $r = .46$).

Hence, participants' eyes moved backward more often in Condition I–C than in Condition I–I, although the fixation durations were almost the same between the two conditions under incoherent contexts.

DISCUSSION

The results of this study are discussed in the context of the two RQs: first, in the coherent condition (RQ1), and then in the incoherent condition (RQ2). This section reviews the results and compares them with previous studies. Finally, the implications for teaching practice are presented.

Coherence and Cohesion Judgment in Coherent Condition (RQ1)

The process of coherence and cohesion judgment was examined using an eye-tracking methodology. The results in the coherent condition showed that the difference between Condition C–C and Condition C–I was not statistically significant either in terms of the eye fixation duration in R1 or R2 or the frequency of regression per item.

Previous studies have indicated that the process of judgment in coherent conditions was promoted by the existence of cohesive ties, with the accuracy rate increasing and the reaction time decreasing (Hamilton, 2011; Fujita, 2021b). Especially for second language learners, coherence can be built more easily if two sentences are connected by a cohesive tie, as pointed out using eye-tracking (van den Bosch et al., 2018). The results of the present study do not support those of previous studies. No phenomenon resulted in which coherence was strengthened by cohesive devices.

Hamilton (2011) gathered 52 participants and used 96 stimulus items, with 24 items for each condition. In Fujita (2021b), 100 Japanese university students participated and judged the same number of items (a total of 96) as in Hamilton (2011). Regarding van den Bosch et al. (2018), although the number of participants was 46 (L1: $n = 27$; L2: $n = 19$), the apparatus had a sampling rate of 300 Hz, which may have led to more detailed results.

The number of participants in the current study was limited to 30 Japanese university students. Moreover, a total of 24 items, with only six per condition, were presented because of methodological restrictions. If the number of participants and stimulus items are increased, more results may be obtained. The sampling rate of the apparatus was 30 Hz. To address this issue, an apparatus with higher performance is necessary. Therefore, further studies are required.

Coherence and Cohesion Judgment in Incoherent Condition (RQ2)

RQ2 addressed the difference in the judgment process between Condition I–C and Condition I–I. In Condition I–C, the two sentences were connected with cohesive ties (e.g., that's why), although they had no coherence. Therefore, this contradictory information was predicted to deter judgments under this condition.

The results indicate that the difference in eye fixation duration between R1 and R2 did not reach statistical significance. The eye fixation durations in R1 and R2 were almost the same under Condition I–C and Condition I–I. The regression frequencies differed significantly. This increase in the number of regressions in Condition I–C could be because the contradictory information from

the cohesive ties in the incoherent context made it more difficult for the participants to build the situation model while processing the discourse.

These results partly support studies with native-speaker participants, such as Ferstl and von Cramon (2001) and Hamilton (2011). However, they were not in accordance with Fujita (2021b), in which Japanese university students did not have difficulty judging items under Condition I–C. The results of the current study showed that the participants may have tried to connect incoherent sentences through cues from cohesive ties, which caused greater regression in the task. They were supposed to look back at the former parts of the item, check the discourse flow, and make use of inferences and world knowledge more often under Condition I–C than under Condition I–I.

The eye fixation durations in R1 and R2 between Condition C–I and Condition I–I did not differ significantly, which meant that the processing speed of judgment was almost the same. The total processing time of the second sentence (R1 +2) was 4.98 s in Condition I–C and 4.78 s in Condition I–I. If the difference between more participants and stimulus items is investigated, the difference might be statistically significant.

Another possibility is that the processes of Condition I–C and Condition I–I may be more complicated than expected. The fixation duration in R1 was 1.31 s in Condition I–C and 1.43 s in Condition I–I, while that in R2 it was 1.56 s in Condition I–C and 1.16 s in Condition I–I. Thus, the processing time is reversed. As Millis and Just (1994) pointed out, when two sentences were connected with a cohesive tie, content integration began immediately, which made processing in R1 faster than in the case without the tie. In Condition I–C, the participants were thought to begin integrating the two sentences immediately after the cohesive ties, which made the fixation duration in R1 shorter than that in Condition I–I. However, the wrap-up of the integration in R2 might have been more difficult in Condition I–C because of the contradictory information compared with Condition I–I. This explains why the eye fixation duration in R2 was longer in Condition I–C than in Condition I–I.

Future studies should be conducted on coherence and cohesion judgment in incoherent conditions involving more participants, stimulus items, and delicate eye-tracking equipment. Moreover, as pointed out before, the meaning of eye movements depends on the researcher's interpretation; different data, generated by interviews or think-aloud, may be necessary. This provides a more precise picture of the aforementioned issues.

Implications for Teaching Practice

The results indicate that participants frequently revisited earlier parts of the text when sentences were connected with cohesive phrases in incoherent contexts. This suggests that they struggled to make inferences or apply world knowledge to construct a coherent mental model of the text. By contrast, when the sentences were coherent, the presence of cohesive ties did not significantly impact the reading process. These results imply that readers should be mindful of cohesive devices, especially when reading difficult texts that lack clarity.

In general, writers attempt to make their texts easier to understand. The use of cohesive devices is one such writing technique commonly employed by skilled writers. Therefore, when a passage is difficult to comprehend, readers should pay close attention to its cohesive ties. For example, in the passage "The shower curtain did not quite shut. *As a result*, the bathroom floor was wet," recognizing the cohesive tie "As a result" helps readers understand the cause-and-effect relationship between the sentences. Simultaneously, readers infer that the person had forgotten to

shut the curtain. This awareness, combined with general knowledge about showers, aids in improving reading comprehension.

Based on these findings, the study proposes several techniques for teaching English. Various types of discourse markers can aid comprehension, including lexical cues (e.g., that's why, as a result), anaphoric devices (e.g., demonstrative and personal pronouns), and logical connectives (e.g., because, though) (Sánchez & García, 2009). Teachers should first explicitly discuss the roles of these markers. Next, students can be provided with texts that contain cohesive ties. They should be encouraged to identify these ties and discuss the logical flow of the discourse. This activity helps students recognize cohesion when reading longer passages.

Another technique involves using passages with blanks for cohesive ties, creating a kind of Cloze test. Students fill in the blanks with appropriate cohesive ties based on the context. This task not only reinforces reading skills but also helps students transition to writing by applying their understanding of cohesion.

The think-aloud method is another effective technique in English classes. Teachers read aloud and explain their thought process, demonstrating how they manage cohesive devices while reading. This helps students learn to build coherence and form a mental representation of the text, thereby improving their reading and comprehension skills.

Additionally, teachers can engage students in a fun group activity where they create a story using cohesive ties. In groups of four, each student takes on a specific role: one sets the situation, another introduces the characters, the third describes the events, and the fourth provides the conclusion. Each student must use at least one cohesive tie in their narration. This group storytelling exercise is not only enjoyable but also helps students focus on the effective use of cohesive ties.

CONCLUSION

The current study investigated the process of coherence and cohesion judgment tasks using eye-tracking evidence. First, the cognitive mechanism of reading was explained, with a focus on the construction of a situational (or mental) model. As one of the component skills of reading, the ability to deal with coherence and cohesion is important. Previous studies on coherence judgment tasks were reviewed, including those with Japanese university EFL learners.

The advantage of using eye-tracking in reading research is that it can detect the process without disturbing natural reading. To date, only a few such studies on reading in a second language exist. Therefore, in this study, evidence of coherence and cohesion was collected through eye-tracking. Thirty Japanese university EFL learners were required to judge whether two-sentence pairs were coherent. The stimulus items consisted of a combination of coherence–incoherence and cohesion–incohesion, for a total of 24 items with six items per condition. Fixation durations on the two regions in the second sentence were measured, and the frequency of regressions per item was counted. The results showed that only the frequencies of regressions increased in Condition I–C, in which the incoherent pairs were connected with cohesive phrases. These results support studies with native-speaker participants (e.g., Ferstl & von Cramon, 2001), but not those with Japanese university participants (e.g., Fujita, 2021b).

The limitations of the study lie in the number of participants, the number of stimulus items, and the eye-tracking apparatus's performance. The participants were 30 Japanese university students. If the number increases, the difference in fixation duration between conditions could be

investigated more precisely, which could yield statistical significance. Second, the total number of stimuli was 24 with each four condition including six items. When the total number of items is increased, the results might be obtained with greater confidence. Finally, with the same number of participants and items, the reading process could be examined more minutely and certainly with a better eye-tracking apparatus. The sampling rate in this study was 30 Hz, whereas some recent studies have used an apparatus at 1000 Hz. If a high-performance apparatus is used, eye movements on each word could be detected, which might bring about new findings. In addition, eye-tracking data should be backed by other methodologies, such as interviews, questionnaires, or think-aloud.

Future studies should proceed in two ways. First, the results of this study should be elaborated using larger numbers of participants and stimuli, and a better apparatus. Eye-tracking data should be reinforced by other methodologies, such as think-aloud or interviewing. This would produce a clearer picture of the processes of coherent and cohesive judgment tasks, which would lead to a more precise mechanism of reading. Another line of studies should focus on using more natural and longer reading materials. These materials have drawn researchers' attention in the direction of everyday reading environments. In addition, coherence and cohesion can be manipulated more naturally when using such materials. This might make coherence and cohesion judgments a method for comprehending the monitoring strategy.

This study is a starting point. The process of coherence and cohesion judgment must be investigated further in the future.

ACKNOWLEDGEMENT

Another version of this study was presented at the 63rd JACET International Convention at Aichi University in 2024. This work was supported by JSPS KAKENHI Grant Number JP19K00839.

REFERENCES

Cain, K., & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, *103*(2), 429–441. doi/10.1037/a0022824

Conklin, K., Pellicer-Sánchez, A, & Carrol, A. (2018). *Eye-tracking: A guide for applied linguistic research.* Cambridge: Cambridge University Press.

Crosson, A. C., & Lesaux, N. K. (2013). Does knowledge of connectives play a unique role in the reading comprehension of English learners and English-only students? *Journal of Research in Reading*, *36*(3), 241–260. doi/10.1111/j.1467-9817.2011.01501.x

Degand, L., & Sanders, T. (2002). The impact of relational markers on expository text comprehension in L1 and L2. *Reading and Writing*, *15*(7–8), 739–757. doi.org/10.1023/A:1020932715838

Demel, M. C. (1990). The relationship between overall reading comprehension and comprehension of coreferential ties for second language readers of English. *TESOL Quarterly*, *24*(2), 267–292. doi.org/10.2307/3586902

Ferstl, E. C. (2015). Inferences during text comprehension: What neuroscience can (or cannot) contribute. In E. J. O'Brien, A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading* (pp. 230–259). Cambridge: Cambridge University Press.

Ferstl, E. C., & von Cramon, D. Y. (2001). The role of coherence and cohesion in text comprehension: An event-related fMRI study. *Cognitive Brain Research*, *11*(3), 325–340. doi.org/10.1016/S0926-6410(01)00007-6

Fujita, K. (2019). Measurement of coherence and cohesion of Japanese EFL learners: A preliminary research. *Bulletin of the Faculty of Letters of Aichi Gakuin University, 48,* 107–118.

Fujita, K. (2020). Coherence and cohesion judgement in English and Japanese, and its relation to English comprehension ability. *Bulletin of the Faculty of Letters of Aichi Gakuin University, 49,* 81–92.

Fujita, K (2021a). The analysis of test items and content validity of a coherence and cohesion judgement task for Japanese EFL learners. *Bulletin of the Faculty of Letters of Aichi Gakuin University, 50,* 131–142.

Fujita, K (2021b). The relationship between English reading comprehension and cross-linguistic coherence-cohesion judgement of Japanese EFL learners. *Transactions of the Institute for Cultural Studies Aichi Gakuin University, 36,* 59–89.

Geva, E. (1986). Reading comprehension in a second language: The role of conjunctions. *TESL Canada Journal*, *3*(0), 85–96. doi.org/10.18806/tesl.v3i0.996

Geva, E. (1992). The role of conjunctions in L2 text comprehension. *TESOL Quarterly*, *26*(4), 731–747. doi.org/10.2307/3586871

Geva, E., & Ryan, E. B. (1985). Use of conjunctions in expository texts by skilled and less skilled readers. *Journal of Reading Behavior*, *17*(4), 331–346. doi.org/10.1080/10862968509547549

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do learners need to learn in order to process coherence relations in narrative and expository text? In A. P. Sweets, & C. E. Snow. (Eds.), *Rethinking reading comprehension* (pp. 82–98). New York, NY: Guilford.

Hamilton, S. T. (2011). *Reading comprehension in adults: Component skills; false memories; and judgements of coherence* [Doctoral dissertation]. UK: University of Sussex. Retrieved from https://hdl.handle.net/10779/uos.23315099.v1

Hessel, A. K., Nation, K., & Murphy, V. A. (2021). Comprehension monitoring during reading: An eye-tracking study with children learning English as an additional language. *Scientific Studies of Reading*, *25*(2), 159–178. doi.org/10.1080/10888438.2020.1740227

Hessel, A. K., & Schroeder, S. (2020). Interactions between lower- and higher-level processing when reading in a second language: An eye-tracking study. *Discourse Processes*, *57*(10), 940–964. doi.org/10.1080/0163853X.2020.1833673

Jonz, J. (1987). Textual cohesion and second-language comprehension. *Language Learning*, *37*(3), 409–438. doi.org/10.1111/j.1467-1770.1987.tb00578.x

Li, M., & D'Angelo, N. (2016). Higher-level processes in second language reading comprehension. In X. Chen, V. Dronjic, & R. Helms-Park (Eds.), *Reading in a second language: Cognitive and psycholinguistic issues* (pp. 159–194). London: Routledge.

Millis, K. K., & Just, M. A. (1994). The influence of connectives on sentence comprehension. *Journal of Memory and Language*, *33*(1), 128–147. doi.org/10.1006/jmla.1994.1007

Mizumoto, A., & Plonsky, L. (2016). R as a lingua franca: Advantages of using R for quantitative research in applied linguistics. *Applied Linguistics, 37*(2), 284–291. doi.org/10.1093/applin/amv025

Nahatame, S. (2017). Standards of coherence in second language reading: Sentence connectivity and reading proficiency. *Reading in a Foreign Language*, *29*(1), 86–112. http://hdl.handle.net/10125/66729

Nahatame, S. (2022). Causal and semantic relations in L2 text processing: An eye-tracking study. *Reading in a Foreign Language*, *34*(1), 91–115. http://hdl.handle.net/10125/67414

Peirce, J., & MacAskill, M. (2018). *Building experiments in PsychoPy*. London: SAGE.

Sánchez, E., & García, J. R. (2009). The relation of knowledge of textual integration devices to expository text comprehension under different assessment conditions. *Reading and Writing*, *22*(9), 1081–1108. doi.org/10.1007/s11145-008-9145-7

Sanders, T. J. M., & Noordman, L. G. M. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes*, *29*(1), 37–60. doi.org/10.1207/S15326950dp2901_3

Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 327–342). Oxford: Oxford University Press.

Sugiura, M., & Yamashita, J. (2011). Aitorakkingu wo tsukatta gengoshorikatei no kenkyu [A study on language processing by eye-tracking]. In I. Fujimura, & N. Takizawa (Eds.) *Gengokenkyu no giho data no shushu to bunseki [Methodologies of language research, data collection and analysis]* (pp. 159–177). Tokyo: Hitsuji Shobo.

Ushiro, Y., Nahatame, S., Hasegawa, Y., Mimura, Y., Hamada, A., Tanaka, N., Hosoda, M., & Mori, Y. (2016). Maintaining the coherence of situation models in EFL reading: Evidence from eye movements. *JACET Journal*, *60*, 37–55.

van den Bosch, L. J., Segers, E., & Verhoeven, L. (2018). Online processing of causal relations in beginning first and second language readers. *Learning and Individual Differences*, *61*, 59–67. doi/10.1016/j.lindif.2017.11.007

van der Schoot, M., Reijntjes, A., & van Lieshout, E. C. D. M. (2012). How do children deal with inconsistencies in text? An eye fixation and self-paced reading study in good and poor reading comprehenders. *Reading and Writing*, *25*(7), 1665–1690. doi.org/10.1007/s11145-011-9337-4

APPENDIX

No.	First Sentence	R0	R1	R2	Judge
111	Global warming will make ices melt.	As a result,	the sea level will rise	around the world.	○
112	Last night there was a terrible storm.	For that reason,	the roads are covered	with water today.	○
113	The center of the earthquake was near Tokyo.	In that area,	some houses broke down	all over the town.	○
114	The university holidays are in August.	At that time,	there are few students	in the campus.	○
115	The pen stopped working just now.	It is because	its ink is empty	for the red color.	○
116	The shower curtain did not quite shut.	As a result,	the bathroom floor is wet	with much water.	○
101	My daughter buys some flowers at the shop.		It is the birthday	of her friend.	○
102	Cindy had a car accident on the street.		The legs were broken	by the crash.	○
103	The speech contest starts on the stage.		Rumi has turned off	the smartphone switch.	○
104	Monica has lost weight by 20 kilograms.		The old skirts fit	her waist again.	○
105	In the autumn the days are getting shorter.		Mary has to go back	home before dark.	○
106	Yoko went to sleep on the beach.		The neck is red	with sunburn now.	○
011	My daughter buys some flowers at the shop.	As a result,	my wife has to pay	a high telephone bill.	×
012	The university holidays are in August.	At that time,	the gas company will cut	the gas line.	×
013	The pen stopped working just now.	As a result,	there are few people	on the streets.	×
014	Kate is going to have some visitors.	In her room,	a tissue box is put	on the desk.	×

015	Kate has a bad cold in bed today.	For that reason,	she wants to eat out	for dinner today.	×
016	The center of the earthquake was near Tokyo.	It is because	money is very important	to the family.	×
001	The shower curtain did not quite shut.		A coffee is needed	for my work.	×
002	An hour has passed since the cooking.		Pink clothes are bought	at the department.	×
003	The news announced much snow on the road.		The monitor has showed	a blue screen.	×
004	The computer has just crashed suddenly.		Fred has left home	an hour earlier.	×
005	Mrs. Green hopes to have a baby girl.		The cake will be made	for today's lunch.	×
006	The speech contest starts on the stage.		The time has come	for a job interview.	×

Note: The item number is created by the condition + item number; #111 indicates Condition C–C, item No. 1, #101 indicates Condition C–I, item No. 1, #011 indicates Condition I–C. item No. 1, and #001 indicates Condition I–I, item No. 1.

Ken Fujita *is a Professor at Global English Department of Aichi Gakuin University, Aichi, Japan. Since 2017, he has been teaching English language, and language teaching methodology. His research interests include second language reading, English language teaching, language teacher training in Japan.*

Email: kfujita@dpc.agu.ac.jp

Mitsuo Ishida *is a Professor at the Faculty of Psychology at Aichi Gakuin University, Aichi, Japan. Since earning his Ph.D. from Aichi Gakuin University in 2005, he has been teaching experimental psychology and psychophysiology. His interests include visuomotor coordination and spatial attention in executing cognitive tasks.*

Email: mitsu-da@dpc.agu.ac.jp