



Text Coherence Necessary for the Promotion of Japanese EFL Learner's Predictive Inference Generation

Masahiro MAEDA^{1&2}

¹Kanazawa Seiryō University, Japan & ²University of Tsukuba, Japan

ARTICLE INFO

Article history:

Received Nov 28, 2023

Revised Dec 30, 2023

Accepted Jan 16, 2023

Keywords:

Reading,
Encoding of predictive
inference,
EFL learner's proficiency

Conflict of Interest:

None

Funding:

None

ABSTRACT

Reading is a very complex process consisting of lower- and higher-level processing. Inference generation is necessary to achieve higher-level processing. It is challenging for Japanese EFL readers because they must use many cognitive resources in lower-level processing, such as word recognition and syntax analysis. This study investigated the relationship between Japanese EFL learners' predictive inference generation as higher-level processing and text coherence from the encoding perspective rather than activation. We used a cued recall task whose rates would be evidence of the reader's encoding of predictive inferences. The recall rates in the high predictability target version, which allowed readers to predict an outcome easily, were more significant than other text versions, and results were affected by the reader's proficiency. These findings suggest that EFL learners' encoding of predictive inference can be influenced under the specific text conditions in higher-level proficiency groups.

Corresponding Author: Masahiro MAEDA, Associate Professor in the Faculty of Humanities at Kanazawa Seiryō University, Ishikawa, Japan. E-mail: maedam@seiryō-u.ac.jp.



© Masahiro MAEDA

This is an open access article under the CC BY-SA 4.0 international license

1. Introduction

1.1 The Importance of Predictive Inference Generation

Text choice is a significant challenge for language teachers, as they must consider numerous factors. For example, the readability or topic familiarity of each story in textbooks and learners' proficiency should be considered. However, textual constraints regarding how much the stories can contribute to students' inference generation should be remembered. It is because a few previous studies have reported that inference generation is a crucial process in successful reading comprehension, as it requires not only word and sentence structure analysis but also the connection of written facts with readers' world knowledge (e.g., Graesser et al., 1994).

Inferences include various types and methods of classification, according to researchers. Trabasso and van den Broek (1985) classified inferences into backward and forward inferences. Backward inferences are reinstatements, connecting inferences, and backward elaboration to produce more causal connections, while forward elaboration, orthogonal elaboration, and associative inferences are forward inferences. Forward inferences, often called predictive inferences, allow readers to generate inferences about the future consequences of a focal event. For example, a frequently used story begins: "Jeff was panicked when he saw the exam questions. He needed to pass this test to graduate and realized he was unprepared (Cranford & Moss, 2019)." The next scene is, "He looked around the classroom and saw the teacher was busy doing something else. He was able to get a clear view of the student's paper next to him without risk." The strong

coherence between his situation (without risk) and intention (to pass the test for his graduation) and the teacher's status (busy doing something) would lead readers to predict that he would cheat.

Table 1: *The Types of Knowledge-Based Inference*

Type of Inference	Brief Explanation
Backward Inference	Referential: A word or phrase is referentially tied to a previous element or constituent in the text (explicit or inferred).
	Causal Antecedent: A causal chain (bridge) between the current explicit action, event, or state and the previous passage's context.
Predictive Inference	Causal Consequence: A forecasted causal chain, including physical events and new plans of agents.

Table 1 shows the types of inferences and their characteristics based on Graesser et al. (1994). They insisted that L1 readers automatically produce backward inferences. However, predictive inferences are said not to be produced automatically, even by L1 readers. L2 readers allocate their cognitive resources in reading processes primarily to lower-level processes such as word coding or structure analysis. Therefore, L2 readers' inference generation is more complicated than L1 readers. This study investigates whether textual coherence helps L2 readers' predictive inference generation.

Predictive inference-making facilitates readers' context processing and construction of situational models (Allbritton, 2004; Linderholm, 2002). When readers understand a text, they form mental representations known as situation models (Johnson-Laird, 1983; van Dijk & Kintsch, 1983; Zwaan & Madden, 2004). Readers must construct and revise the necessary situation models to read effectively (Muramoto, 2000; Ushiro et al., 2012; van Dijk & Kintsch, 1983).

1.2 Predictive Inference Generation and Coherence-building

Predictive inferences are difficult to measure and are not often found, except when a coherence break is resolved (Murray et al., 1993), when the prediction is causally constrained within the text (Murray et al., 1993), or when the prediction is foregrounded explicitly in the text (Whitney et al., 1991). These studies reported that text coherence affected readers' predictions. Cranford and Moss (2019) examined text-related factors under different test conditions that could affect predictive inference generation for L1 readers. They compared the readers' response times to a target probe word, which summarized the likely outcome in one word, such as "CHEAT" for Jeff's story mentioned previously, thereby changing the text coherence. There are four text versions in their study with different coherences in the stories: High-Predictability Target (HPT), High-Predictability Alternative (HPA), Mutual Predictability (MP), and Control Version (CO). The stories from the HPT version have strong coherence with resultative story development, similar to Jeff's story in section 1.1. The stories from the HPA version likely had consequences with paradox story development that could interfere with readers' generation of target inferences.

For example, in the story about Jeff, the sentence "He needed to pass this test to graduate, and he realized he wasn't prepared" would have readers recognize that Jeff had to pass the test for his graduation. However, the next scene would be: "He looked around the classroom and saw the teacher was looking right at him. It would be too risky to try to get a clear view of the student's paper next to him," which would lead readers to infer that he would be unable to cheat. His situation (he must pass the test) and the teacher's situation (looking at Jeff) expresses the antinomy development, and readers will predict Jeff would not cheat. This paradox is thought to interfere with the recall process in the present study. The stories from the MP version had mutual predictability regarding target consequences. For example, the first scene is the same as in the HPT and HPA version, "He needed to pass this test to graduate, and he realized he wasn't prepared" would have readers recognize that Jeff had to pass the test for his graduation. However, in the next scene, "He saw teacher walking on the other side of the classroom. He considered trying to get a clear view of the student's paper next to him, but it would be risky." This description expresses that his safe situation (he saw the teacher walking on the other side of the classroom) and the risk (it would be risky) are mixed. This description will allow readers to predict the two mutual possibilities that Jeff will cheat or give up cheating, influencing readers' recall in this study. Finally, the stories from the CO version could function as filler stories, which do not strongly lead to predictable consequences. For example, "Jeff was happy when he saw the exam questions. He needed to pass the test to graduate, and he realized that he was very well prepared." this will have the reader believe that Jeff does not need to cheat. Cranford and Moss (2019) concluded that generating predictive inferences can be difficult when story development has multiple possible outcomes without strong coherence, as in the HPA situation.

1.3 Study Purpose

The present study attempts to verify the influence of differences in text coherence, which have different strengths of coherence with different storylines, using the experimental materials HPT, HPA, MP, and CO versions from Cranford and Moss (2019) on predictive inference generation, considering its benefit. As mentioned above, L1 research has previously been conducted to examine text-related factors on predictive inference generation, but only some studies exist for L2 research. One of the characteristics of L2 readers is their significant individual differences. Therefore, this study considers readers' proficiency in addition to textual differences.

The measurement of inference generation, which refers to how readers' inference generation can be assessed, is controversial. Both online and offline measurements were available. Regarding online measurement, the response time to target probe words is widely used (Cranford & Moss, 2019). However, regarding offline measurements, free recall is a popular option (Ushiro et al., 2012; Nahatame, 2014). Offline measurements, such as recall tasks, will be adopted rather than online measurements while reading because previous researchers have revealed that EFL readers' predictive inference generation will be delayed (e.g., Ushiro et al., 2012). The present study will examine whether predictive inferences could be generated and maintained (encoded) in readers' mental representation from the recall rate results, as investigated by Ushiro et al. (2012). This study contributes to story selection by EFL readers in terms of predictive inference generation. The following research questions (RQs) were established:

RQ1: Do coherent differences affect EFL learners' encoding of predictive inferences in their mental representations?

RQ2: Do differences in readers' proficiency affect EFL learners' encoding of predictive inferences in their mental representations?

2. Method

2.1 Participants

Participants were 64 Japanese freshman undergraduate students (35 men and 29 women, aged 19-20 years, $M = 19.8$, $SD = 1.52$) from a Japanese university. Four participants did not complete the tasks, and their scores were excluded from the analysis.

2.2 Results of the Reading and Listening Test

On the TOEIC® IP test, which includes a reading section and a writing section with 100 questions each, the average score for all participants was relatively low: $M = 323.33$ (out of 990), $SD = 100.69$, $Max = 610$, $Min = 235$. The average score for the reading section was $M = 159.21$ (out of 495), $SD = 35.51$.

2.3 Results of Vocabulary Test

In addition to reading proficiency, participants' vocabulary level was tested using "The R vocabulary test" (Oba, 2016), which was created based on the "Japan Association of College English Teachers (JACET) 8000" (University English Education Society, 2016). The results showed that their vocabulary levels were $M = 33.75$, $SD = 3.69$. This result indicates that their vocabulary level was estimated to be between 3300 and 3400 in the JACET800 word list.

2.4 Materials

Eight English stories (four-story versions $\times$ 2) cited from Cranford and Moss (2019) were used. As mentioned above, the story version consists of a High-Predictability Target (HPT), High-Predictability Alternative (HPA), Mutual Predictability (MP), and Control Version (CO). As mentioned in section 1.2, the stories from the HPT version have strong coherence with resultative story development. The stories from the HPA version will provide readers with alternative predictions rather than target predictions, whose contradictions may confuse readers—the stories from the MP version are the two mutual possibilities to occur. Finally, the CO version's stories work as filler stories, which will not let readers generate the target predictive inference.

2.5 Procedure

Each participant received a booklet containing each condition (four-story versions $\times$ 2). The order of stories was counterbalanced, changing the presentation order of stories one by one. Participants were informed that they would have to perform a task after reading. However, they were unaware of having to perform a cued recall task. Readers should be able to predict the final scene of each story, which is not written in the story. After reading all the passages, they received another booklet with a space where they had to recall and write down each passage. A recall task, which was classified based on offline measurement, was adopted to assess learners' encoding of inferences (Rapp & Mensink, 2011). Here, an offline measure assessed the encoding of

predictive inferences in their mental representation. Participants recalled and wrote down as much as they could about the passages in Japanese.

2.6 Scoring and Data Analysis

Recall rates were analyzed. First, all experimental texts were divided into idea units (IUs) based on Ikeno (1996). Two raters conducted this process ($\alpha = .72$ (substantial)). Disagreements were resolved through discussion. Two raters randomly selected and scored 30% of the cued recall data independently ($\alpha = .77$ (substantial)). Disagreements were resolved through discussion, and one rater scored the remaining data individually. Taking the participants' differences in English proficiency based on reading scores from TOEIC® IP and vocabulary scores from Oba (2016), the cued recall data were analyzed using a generalized linear mixed model (GLMM) using SPSS advanced version 28.0.1.

3. Results

3.1 Quantitative Analysis

Table 2 presents the recall production rates for each story version. The High-Predictability Target obtained the highest portion of recall ($M = .42$, $SD = .11$), while the control version obtained the lowest ($M = .30$, $SD = .08$). This result suggests that resultative story development (HPT) promotes readers' predictive inference generation.

Table 2: Recall Production Rates in Each Story Version

HPT		HPA		MP		CO	
<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
.42	.11	.31	.08	.31	.08	.30	.08

Table 3 shows the estimates of the recall rates among the story versions. The results show that the estimate of the High-Predictability Target for the recall rate had a more significant effect (estimate = .22, $SE = .02$). This finding strengthens the suggestions presented in Table 2.

Table 3: Estimate for Response Times among Story Versions

Story Version	Estimate	SE	95% CI	
			Lower	Upper
High-predictability target	.22	.02	.18	.27
High-predictability alternative	.20	.02	.15	.24
Mutual predictability	.20	.02	.15	.24
Control version	.19	.02	.14	.22

Table 4 compares the estimates for recall rates between each version and the whole mean score. The results indicate that the recall rate for the High-Predictability Target had a significant effect (estimate = .02, $SE = .01$, $t = 1.98$, $p = .048$ ($p < .05$)). In addition to the results in Tables 2 and 3, this result suggests that only the HPT version would help readers promote predictive inference generation.

Table 4: Comparison of Estimate for Recall Rates between Each Version and Mean

Comparison	Estimate	SE	<i>t</i>	<i>p</i>	95% CI	
					Lower	Upper
HPT – Mean	.02	.01	1.98	.048	.00	.05
HPA – Mean	-.01	.01	-.52	.603	-.03	.02
MP – Mean	-.01	.01	-.49	.628	-.03	.02
CO – Mean	-.01	.01	-.98	.329	-.03	.01

Note. The unit of estimate is the rate of recalled IUs.

Table 5 presents the recall rate results of the mixed-effect model. The results indicated that the recall rate was affected by the reader's proficiency (estimate = .008, $SE = .004$, $z = 2.06$, $p = .039$ ($p < .05$)). This result indicates a relationship between recall rates and learners' proficiency.

Table 5: Recall Rates Results from a Linear Mixed-effect Model

	Estimate	SE	<i>z</i>	<i>p</i>	95% CI	
					Lower	Upper

Subject	.004	.002	2.35	.019	.002	.009
Proficiency	.008	.004	2.06	.039	.003	.020

As seen above, predictive inference generation is influenced by story development and learner proficiency.

3.2 Qualitative Analysis Based on the Participants' Comments Regarding the Task

A sample of the participant protocols is shown in Table 6. The participants' comments were originally written in Japanese and translated into English by the author.

Table 6: *Samples of Participants' Predictions from the High-Predictability Target Version*

Story [Target prediction]	The first and last scene	Participants' prediction
<right prediction>		
Story A [SCOLD]	<first scene> Timmy returned from school and went straight to his bedroom to watch television. One channel played a movie that some of his classmates had been discussing. It was an R-rated movie, and he knew he should not watch it. <last scene> He knew it was his mother when he heard a loud noise outside his bedroom door.	His mother will scold him. / He heard his mother scolding him.
<wrong prediction>		
Story B [CATCH]	<first scene> Danny loved to play baseball. He was on one of the town's Little League teams, and he and his friends often played a pickup game after school. Danny was one of the best players in the league and was an outstanding outfielder. <last scene> Danny looked up and opened his glove as he watched the ball fly through the air.	The word "outfielder" means the person out of the field / "the ball flew through the air" means the ball was gone somewhere else.

Passages in the High-Predictability Target version facilitated the generation of easy predictive inference. For example, the target prediction in Story A (shown in Table 6) is *SCOLD*. In the first part of this story, readers can recognize that the boy did something terrible (he was watching an R-rated movie); then, the last scene in this story follows; "He knew it was his mother when he heard a loud noise right outside of his bedroom door." Therefore, most readers could make the following predictions: "His mother will scold him." This information was implied in the first part of the story, not explicitly written down, and most readers successfully connected this information with the story's outcome. Thus, the story's coherence could help readers predict what would happen next. However, some readers had irrelevant predictions, as shown in Story B's predictions. The first scene mentions Danny is a good player, and so the last scene, "Danny opened his glove as he watched the ball fly through the air," would have readers predict that he will catch the ball in a resultative way. However, some readers left comments such as "I'm not sure much about baseball" and mentioned that they believed the word "outfielder" meant the person who was out of the field and "watched the ball fly through the air" meant that the ball went somewhere. These incorrect comments may be due to a lack of topic familiarity, as even readers who have much knowledge about baseball but do not know the English words for the position names in baseball would fail to produce an appropriate prediction even if they read the HPT version because of the lack of language coherence rather than topic familiarity.

4. Discussion

The present study examined text coherence and participants' proficiency in inferring outcomes from written text. The results from the GLMM analysis implied that the encoding of inferences occurred only in the High-Predictability Target version and was affected by the reader's proficiency. The experimental results indicated

that text coherence was essential in predictive inference generation. In addition, a critical discussion point is that EFL readers often find it challenging to deal with English passages because of their limited resources regarding L2 reading-related language knowledge and cognitive resources (e.g., word recognition, syntax level recognition, maintaining the coherence of the sentences, encoding inferences in their mental representations and etcetera.). Therefore, a specific proficiency in lower-level processing is necessary. It implies that for L2 learners, processing and generating inferences in reading are cognitively demanding. We now review the two primary research questions of this study.

RQ1: Do the coherent differences affect EFL learners' encoding of predictive inferences into their mental representation?

The experimental results show that the generation of predictive inferences is limited, depending on text coherence. In the current study, predictive inferences were encoded into mental representations only when the relevant passages had a high predictability target condition. Therefore, the characteristics of the relevant reading passage provided to EFL readers are essential in this regard. When the goal of a lesson focuses on generating predictive inferences for making a given context richer, teachers should sometimes choose appropriate passages or revise passages as required.

RQ2: Do the differences in readers' proficiency affect EFL learners' encoding of predictive inferences into their mental representation?

The experimental results showed that recall rates were higher as high proficiency increased. One of the characteristics of L2 learners is that their differences are significant. Nahatame (2017) suggested that L2 learners usually base their standards of coherence on causal and semantic relations between sentences, but the standards of coherence adopted differ according to their L2 proficiency. He also claimed that less-proficient readers were more concerned with semantic relevance, whereas more-proficient readers tended to be more concerned with causal relevance-generating inferences. Regarding readers' proficiency, this study found that decreasing the burden on L2 readers' lower-level processes might be necessary to improve predictive inference generation.

5. Limitation of This Study

The current study was based on previous studies, whose findings said that EFL readers struggle to generate predictive inferences during reading, and the generation will be delayed (Ushiro et al., 2012). That is why we investigated offline inference generation as the encoding of predictive inference by cued recall task, whose recall rates are the evidence of the reader's mental representation. However, it cannot be denied that using the recall test may have resulted in a larger measurement of simple memory strength than that of predictive inferences. Future research must investigate why online predictive generation is tricky for EFL readers using online measurement, such as think-aloud tasks in qualitative analysis.

6. Conclusion & Educational Implication

This study examined whether variable story coherence could affect EFL readers' generation of predictive inferences. The results showed that the generation of predictive inferences occurred in a limited scene in offline contexts, as previous research has also revealed (Potts et al., 1998; McKoon & Ratcliff, 1986, 1999; Ushiro et al., 2012). The encoding of targeted predictive inferences is closely related to higher-proficiency levels. Teachers should choose appropriate passages or, sometimes, revise passages to allow learners to decrease the burden of language analysis, leading to an easier generation of predictions. The findings in this study make essential contributions to enhancing teachers' understanding of their students' inference generation, which could impact how teachers choose the text that influences predictive inference generation.

Acknowledgements

I would like to express my sincere gratitude to Dr. Yuko HIJIKATA and her seminar colleagues at the University of Tsukuba for their helpful comments.

References

- Cranford, E. A., & Moss, J. (2019). Generating predictive inferences when multiple alternatives are available. *Discourse Processes*, 56, 289–309. <https://doi.org/10.1080/0163853X.2018.1497921>.
- Graesser, A. C., Singer, N., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101, 371–395. <https://doi.org/10.1037/0033-295x.101.3.371>
- Ikeno, O. (1996). The effects of text-structuring-guiding questions on comprehension of texts with varying linguistic difficulties. *JACET Bulletin*, 27, 51–69. <http://ci.nii.ac.jp/naid/110003814657>.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95, 163–182. <https://doi.org/10.1037/0033-295X.95.2.163>

- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 82–91. <https://doi.org/10.1037/0278-7393.12.1.82>
- Muramoto, T. (2000). Daini gengo no bunshou rikai katei ni oyobosu shujyukudo no eikyo –Bunsaininkadai niyoru jikken– [The effects of second-language proficiency on text comprehension]. *Science of Reading*, 44, 43–50.
- Murray, J. D., Klin, C. M., & Myers, J. L. (1993). Forward inferences in narrative text. *Journal of Memory and Language*, 32, 464–473. <https://doi.org/10.1006/jmla.1993.1025>
- Murray, J. D., & Burke, K. A. (2003). Activation and Encoding of Predictive Inferences: The Role of Reading Skill. *Discourse Processes*, 35, 81–102. https://doi.org/10.1207/S15326950DP3502_1
- Nahatame, S. (2014). Suppressing and revising incorrect predictive inferences in EFL reading: An empirical study using the meaningfulness judgment and sentence recognition tasks. *ARELE*, 25, 159–174. https://doi.org/10.20581/arele.25.0_159
- Nahatame, S. (2017). Standards of coherence in second language reading: Sentence connectivity and reading proficiency. *Reading in a Foreign Language*, 29, 86–112.
- Oba, M. (2016). *RLG vocabulary test*. <http://www.mitewakaru.com/RLG/>
- Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurrence of elaborative inference: Lexical decision versus naming. *Journal of Memory and Language*, 27, 399–415. [https://doi.org/10.1016/0749-596X\(88\)90064-2](https://doi.org/10.1016/0749-596X(88)90064-2)
- Trabasso, D. N., & Mensink, M. (2011). Focusing effects from online and offline reading tasks. *Text Relevance and Learning from Text*, 141–164. Charlotte, NC: Information Age Publishing.
- Trabasso, T., van den Broek, P., & Suh, S. Y. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes*, 12, 1–25. <https://doi.org/10.1080/01638538909544717>
- University English Education Society. (2016). *Daigaku eigo kyoiku gakkai shin JACET8000 [New JACET 8000]*. Kiriara Shoten.
- Ushiro, Y., Nahatame, S., Hasegawa, Y., Shimizu, H., Kimura, Y., Tanaka, Y., & Nakagawa, C. (2012). Activation and encoding of bridging and predictive inferences in EFL reading comprehension. *JACET Journal*, 54, 33–52.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.
- Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: dynamic representations in language comprehension. *Cognitive Science*, 28, 611–619. https://doi.org/10.1207/s15516709cog2804_5.

The Author:

Masahiro Maeda (<https://orcid.org/0000-0001-5803-0606>), an Associate Professor in the Faculty of Humanities at Kanazawa Seiryō University, Ishikawa, Japan. While working, his doctoral coursework is progressing at the University of Tsukuba, Ibaraki, Japan graduate school. His speciality is English teaching methodology, especially for L2 learners' reading processes. Email: maedam@seiryo-u.ac.jp