



Causal and Semantic Relations in Second Language Discourse Processing: An Eye-Tracking Study

Shingo Nahatame

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.


July 13, 2020

**Causal and Semantic Relations in Second Language Discourse Processing:
An Eye-Tracking Study**

Shingo Nahatame¹

¹School of Education, Faculty of Human Sciences, University of Tsukuba

Author Note

Shingo Nahatame  <https://orcid.org/0000-0002-8488-8603>

The author declares that there no conflicts of interest with respect to this preprint.

This research was supported by JSPS KAKENHI Grant Number 17K13507.

Correspondence should be addressed to Shingo Nahatame, University of Tsukuba; 1-1-1

Tennoudai, Tsukuba, Ibaraki, Japan. Email: nahatame@human.tsukuba.ac.jp

Abstract

This study investigated how causal and semantic relatedness between sentences affects second language discourse processing as reflected by eye movements. Japanese learners of English read two-sentence texts varying in causal and semantic relatedness and their eye movements were recorded. Linear mixed-effects models of eye movement measures revealed that causal relatedness has a robust impact on both fixation durations and lookback frequency, whereas the effects of semantic relatedness are modulated by causal relatedness and reading skill.

Keywords: reading, eye-tracking, coherence, narrative, latent semantic analysis, natural language processing, second language

Causal and Semantic Relations in Second Language Discourse Processing: An Eye-Tracking Study

To achieve coherent text comprehension, readers need to establish meaningful connections among sentences. Among several types of sentence relations, discourse processing researchers have commonly focused on causal and semantic relations because examining them provides an opportunity to assess different types of processing involved in constructing coherent comprehension (e.g., Wolfe et al., 2005; Todaro et al., 2010). From the viewpoint of the Construction-Integration (CI) model (Kintsch, 1998), semantic relations would derive the bottom-up processing during the construction phase, whereas causal relations would affect how integration proceeds by strengthening or deactivating the concepts.

Wolfe et al. (2005) indicated that readers more quickly process two-sentence texts when the sentences are more causally and semantically related, although the effects of semantic relatedness is likely overwhelmed by those of causal relatedness. Nahatame (2018) obtained similar results for second language (L2) discourse processing. Furthermore, Todaro et al. (2010) suggested that causal relations more likely influence processing by skilled readers, whereas semantic relations more likely influence processing by less skilled readers.

One limitation of these past studies is their employment of a sentence-by-sentence reading paradigm. Given that looking back to the preceding sentences is a typical behavior for maintaining coherence in discourse comprehension (Hyönä et al., 2003), it is important to overcome this limitation. Thus, the current study extends Nahatame's (2018) study on L2 discourse processing by employing eye-tracking to allow for lookbacks during reading and aims to provide a better picture of how causal and semantic relations influence L2 discourse processing.

Method

Participants

The participants were 48 undergraduate and graduate students in Japan, all native speakers of Japanese and intermediate-level learners of English. Their English reading proficiency was assessed using a standardized test to include the variable of L2 reading skill level in the analysis.

Materials

The experimental items were 20 sets of two-sentence texts adopted from Nahatame's (2018) study, which are originally from Wolfe et al. (2005). Each set included four types of the first (prime) sentence and one common second (target) sentence (see Table 1). The prime sentences were manipulated in terms of their causal and semantic relatedness to the target sentence (high/low relatedness). Four lists of text items were created counterbalancing the variables of causal and semantic relatedness.

Table 1

Sample of Experimental Texts

1a.	Mary could not find anything to read in the library.	(CR-High / SR-High)
1b.	Mary wanted to look for recipes for her dinner party.	(CR-High / SR-Low)
1c.	Mary went to the library to look for something to read.	(CR-Low / SR-High)
1d.	Mary was having a dinner party for her office.	(CR-Low / SR-Low)
2.	She went to the bookstore to get new books.	(Target sentence)

Note. CR = Causal Relatedness; SR = Semantic Relatedness.

The causal manipulation was ensured by the subjective ratings obtained from more than 100 skilled English readers recruited through Amazon Mechanical Turk. The semantic manipulation was verified by employing latent semantic analysis (LSA) on the TASA corpus to obtain cosines between the prime and target sentences (Wolfe et al., 2005; Todaro et al., 2010). Note that the following analysis included these ratings and cosines as independent variables instead of the high/low relatedness category.

Procedure and Eye Movement Measures

The participants were randomly assigned one of the four experimental lists. Participants were instructed to read each text on the computer screen for comprehension and to answer a yes/no comprehension question after each text. Their eye movements during reading were recorded with the EyeLink® 1000 Plus eye tracker (SR Research Ltd., Canada).

This study collected two kinds of eye movement measures (Hyönä et al., 2003): (a) first-pass reading time of the target sentences, which is indicative of initial processing of the discourse, and (b) the occurrence of lookbacks to the prime sentences after finishing reading the target sentences, which reflects the late stage of discourse processing.

Statistical Analysis

First-pass reading times were log-transformed and then analyzed using linear mixed-effects models by running the lme4 package version 1.1.18.1 on R version 3.6.1. Fixed effects included the causal relatedness (subjective ratings obtained from the AMT participants), semantic relatedness (LSA cosines), L2 reading proficiency (reading test scores), and the interactions of these variables. Each factor was centered on its mean. The lookback data were

analyzed in the same way as the first-pass reading time analysis, except that the *glmer* function (generalized linear mixed-effects model) was employed given that the data of lookbacks is binary.

The maximal model was constructed with random intercepts for participants and items as well as random slopes for several variables (e.g., word frequency of target sentences and accuracy of comprehension questions). Because the maximal model showed over-parameterization, the model was simplified by identifying the random effect parameters with the lower variances and removing them as long as it resulted in no significant loss of goodness of fit.

Results

First-Pass Reading Time

Table 2 displays the results of the final model of the first-pass reading times for the target sentences. There were significant main effects of the causal relations and reading proficiency. However, these effects were qualified by the significant interactions between causal relatedness and semantic relatedness and between semantic relatedness and reading proficiency.

Table 2

Summary Table for the Results From a Mixed Effects Model of First-Pass Reading Times

<i>Parameters</i>	<i>Fixed Effects</i>				<i>Random Effects</i>			
	Estimate	<i>SE</i>	<i>t</i>	<i>p</i>	By Participant		By Item	
					Variance	<i>SD</i>	Variance	<i>SD</i>
Intercept	4.70	0.05	102.94	<.001*	0.04	0.19	0.02	0.16
SR	-0.10	0.07	-1.38	.168	—	—	—	—
CR	-0.05	0.01	-6.97	<.001*	—	—	—	—
Pro	-0.03	0.01	-3.34	.002*	—	—	—	—
SR × CR	-0.12	0.05	-2.20	.028*	—	—	—	—
SR × Pro	-0.05	0.02	-2.70	.007*	—	—	—	—
CR × Pro	0.00	0.00	0.21	.832	—	—	—	—
SR × CR × Pro	0.01	0.01	0.67	.503	—	—	—	—

Note. CR = causal relatedness; SR = semantic relatedness = SR; Pro = reading proficiency. Model formula: $\log(\text{FPR}) \sim \text{SR} * \text{CR} * \text{Pro} + (1 | \text{Subject}) + (1 | \text{item})$. $R^2_m = .08$; $R^2_c = .40$.

Figure 1 visualizes the interaction effect between causal and semantic relatedness. As shown in the left panel, overall, the first-pass reading times for the target sentences were shorter when they were preceded by more causally related sentences. However, the effects became more prominent when the sentences were also semantically related. On the other hand, the facilitation

effects of semantic relatedness were only observed when these sentences were more causally related (see the right panel of Figure 1).

Figure 1

The plots of causal and semantic relation effects in the model of first-pass reading times.

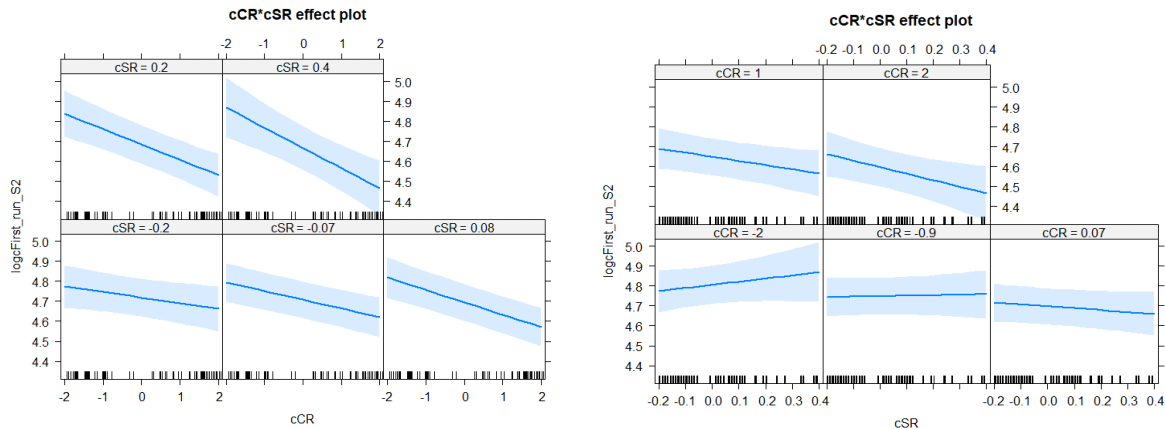
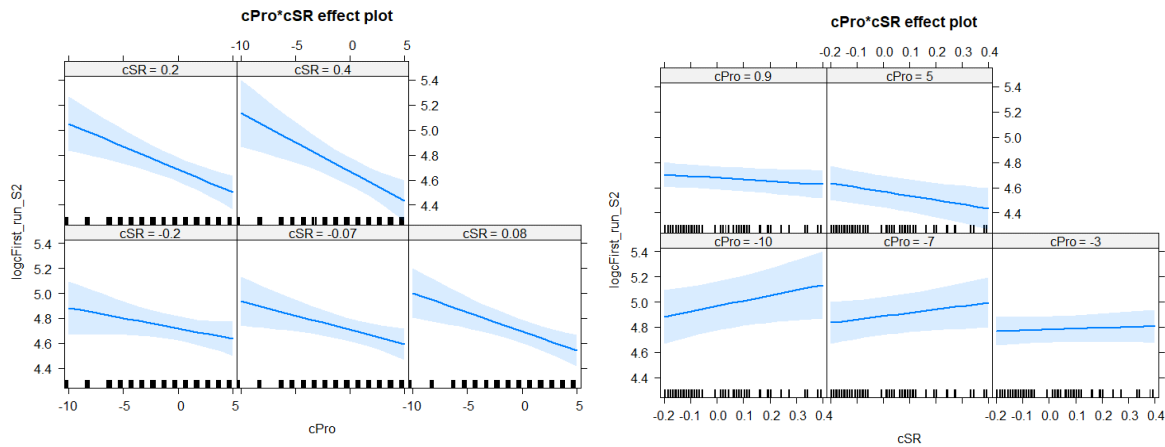


Figure 2 displays the interaction effect between semantic relatedness and reading proficiency. Not surprisingly, as illustrated in the left panel, first-pass reading times were shorter for more proficient readers in general; however, the effects became greater as the sentences are higher in semantic relatedness. More importantly, it seems that the effects of semantic relatedness varied according to reading proficiency level (see the right panel of Figure 2). The relatedness had little influence on the processing times or tended to decrease them among proficient readers, whereas it likely increased the less proficient readers' processing times.

Figure 2

The plots of semantic relation and reading proficiency effects in the model of first-pass reading times.



Lookbacks

Table 3 presents the results of the final model for the lookbacks from the target to prime sentences. There was a significant main effect of the causal relatedness. The negative estimates of the causal relatedness effects indicated that participants were less likely to make lookbacks when the target sentences were preceded by causally close sentences, as illustrated in Figure 3.

Furthermore, there was a significant interaction between semantic relatedness and reading proficiency, although the effect was just below the significance level. As shown in the left panel of Figure 4, proficient readers were more likely to make lookbacks than less proficient readers when the sentences were more semantically related. Similarly, proficient readers were more likely to make lookbacks for the more semantically related sentences than less related sentences (see the right panel of Figure 4). However, the opposite tendency was observed for less proficient readers. They were less likely to make lookbacks for the more semantically related sentences than less related sentences.

Table 3

Summary Table for the Results From a Mixed Effects Model of Lookbacks

Parameters	Fixed Effects				Random Effects			
	Estimate	SE	z	p	By Participant		By Item	
					Variance	SD	Variance	SD
Intercept	-0.77	0.16	-4.91	<.001*	0.49	0.70	0.14	0.37
SR	0.89	0.52	1.71	.088	—	—	—	—
CR	-0.39	0.06	-6.89	<.001*	—	—	—	—
Pro	0.04	0.04	1.12	.261	—	—	—	—
SR × CR	-0.21	0.38	-0.55	.583	—	—	—	—
SR × Pro	0.31	0.16	1.98	.047*	—	—	—	—
CR × Pro	-0.01	0.02	-0.31	.754	—	—	—	—
SR × CR × Pro	-0.11	0.11	-0.95	.340	—	—	—	—

Note. CR = causal relatedness; SR = semantic relatedness = SR; Pro = reading proficiency. Model formula: Lookbacks ~ SR * CR * Pro + (1 | Subject) + (1 | item). $R^2_m = .09$; $R^2_c = .23$.

Figure 3

The plots of causal relation effects in the model of lookbacks.

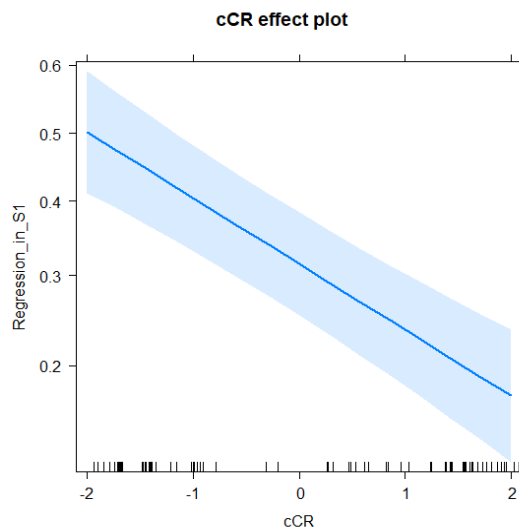
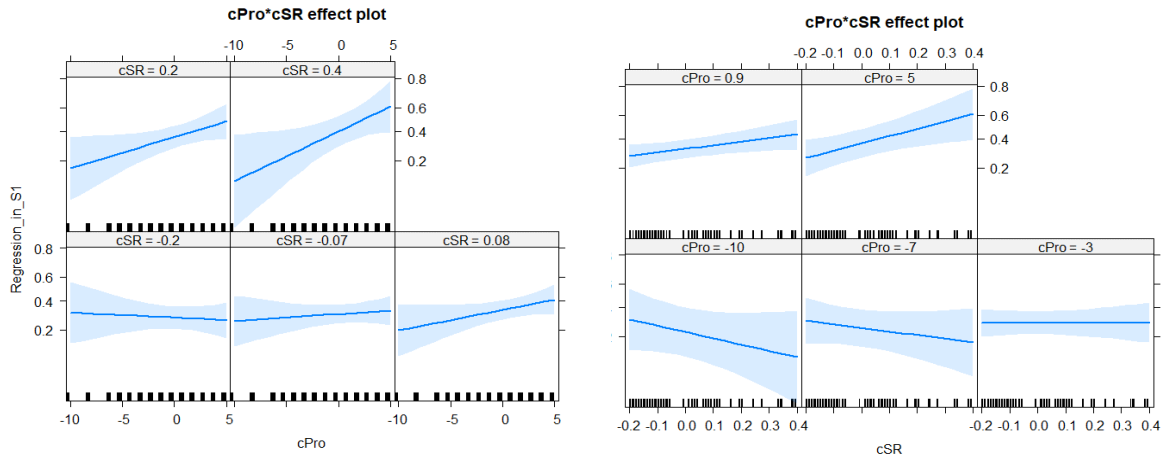


Figure 4

The plots of semantic relation and reading proficiency effects in the model of lookbacks.



Discussion

The analyses of eye movement measures showed that causal relatedness between sentences had a robust impact on first-pass reading times as well as lookbacks during L2 reading, regardless of reading skill level. This suggests that readers strongly rely on the causal relations between sentences to establish coherence in discourse representations even when they process the texts in L2, which is in line with the view of past studies (e.g., Nahatame, 2018).

Furthermore, the use of eye-tracking in the current study enabled us to extend the previous findings in some regards. First, this study demonstrated that causal relatedness effects are still observed during more natural reading processes, which might be distorted by the single-sentence presentations. Second, although the sentence-by-sentence reading experiments indicated that less causally related sentences induce some kind of additional processing during reading, the current eye-tracking study specified what these additional processes are. That is, readers not only devote extra effort at processing the sentences that are less causally related to the previous sentences, but also make lookbacks to the previous sentences. This suggests that causal relations exert their influence at both early and later processing of L2 discourse comprehension.

Semantic relatedness also had an impact on both first-pass reading times and lookbacks, but the effects were more limited than causal relatedness effects. The facilitation effects of semantic relatedness on first-pass reading times were only prominent when the sentences were high in causal relatedness. This confirms the previous findings that the semantic relatedness effects on processing times are overwhelmed by causal relatedness effects (Nahatame, 2018; Wolfe et al., 2005). However, the effects of semantic relatedness in the current study were more complicated than those observed in the past studies in that they also interacted with reading skill level. Higher semantic relatedness had little impact on or slightly promoted the early processing

of the target sentences among the skilled readers, whereas it inhibited the processing among the less skilled readers. These results might be explained by considering the different characteristics of skilled and less skilled readers. According to Todaro et al. (2010), the difference between these readers lies in the ability to prune or adjust the network of activated semantic propositions that has been created from the construction phase. That is, less skilled readers are not as adept at suppressing contextually irrelevant concepts in this network as are skilled readers (Gernsbacher, 1990). Given such a difference, it is reasonable to assume that both skilled and less skilled readers in the current study activated the semantic associations, but skilled readers were successful at suppressing the contextually irrelevant information, whereas less skilled readers failed to do so, resulting in the inhibition effects of high semantic relatedness on their processing.

The similar interaction effect on lookbacks supports the above interpretations. Skilled readers more frequently make lookbacks for the sentences high in semantic relatedness. Such frequent lookbacks can be explained by assuming that skilled readers reanalyzed the previous sentences to suppress superficial semantic relations and to consider the logical connections trying to establish coherence at a deeper level. On the other hand, less skilled readers do not engage in such strategic processes for the semantically related sentences, suggesting that they might have difficulty defining what information should be processed to a deeper level and integrated to the developing text representation.

Nevertheless, the effects associated with semantic relatedness on lookbacks were just below the significance level. Given this, it is possible that the relatedness is more likely to influence the initial phrase than the later phase of discourse processing. This is consistent with the notion that semantic associations to text information are generated and activated quickly during the initial phase of comprehension (Kintsch, 1998; Todaro et al., 2010).

In conclusion, causal relations exert their influence on not only the initial processing but also later integrative processing of the L2 discourse. Semantic relations also have an impact on L2 discourse processing, but the influence is more limited and complicated. The semantic relation effects are more likely to be observed on the initial phrase than the later phase of discourse processing and modulated by other factors, such as causal relatedness and reading skill.

References

- Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Lawrence Erlbaum Associates.
- Hyönä, J., Lorch, R.F., Jr., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313–334). Elsevier Science.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Nahatame, S. (2018). Comprehension and processing of paired sentences in second language reading: A comparison of causal and semantic relatedness. *Modern Language Journal*, *102*(2), 392–415.
- Todaro, S., Millis, K., & Dandotkar, S. (2010). The impact of semantic and causal relatedness and reading skill on standards of coherence. *Discourse Processes*, *47*(5), 421–446.
- Wolfe, M. B. W., Magliano, J. P., & Larsen, B. (2005). Causal and semantic relatedness in discourse understanding and representation. *Discourse Processes*, *39*(2-3), 165–187.