

Causal and Semantic Relations in L2 Text Processing: An Eye-Tracking Study

Shingo Nahatame
University of Tsukuba
Japan

Abstract

This study is an extension of Nahatame's (2018) research that demonstrated the effects of causal and semantic relations between sentences on second language (L2) text processing. Employing eye tracking, this study aimed to examine whether these effects appear during more natural, uninterrupted reading processes and to identify the time course of the effects. In the experiment, Japanese learners of English read two-sentence texts that varied in their causal and semantic relatedness, as evaluated by crowdsourced human judgments and via a computational approach (latent semantic analysis), respectively. Two eye-movement measures were collected and analyzed: first-pass reading times for the second sentence and lookbacks from the second to the first sentence. The results indicated that causal relatedness had a robust impact on both reading times and lookbacks. However, semantic relatedness impacted only reading times, and its effects were modulated by causal relatedness. Theoretical, pedagogical, and methodological implications of this finding were discussed.

Keywords: reading processes, eye tracking, discourse processing, coherence, psycholinguistics, computational linguistics, English as a Foreign Language (EFL)

Efficient word recognition and syntactic parsing are essential for successful reading comprehension (Grabe, 2009; Jeon & Yamashita, 2014). However, readers are also required to process text beyond a single word or sentence, such as making meaningful connections between sentences, to achieve coherent text comprehension (Gernsbacher, 1990; Kintsch, 1988, 1998). Although several types of relations are established in text comprehension (e.g., causal, semantic, referential, spatial, and temporal relations; van den Broek et al., 2011), causal and semantic relations have been most widely investigated in the field of cognitive science. Causal relations refer to the cause-and-effect relationship among the events described in the text, whereas semantic relations refer to the overlap or similarity of meanings among text elements. It is now well established through many empirical studies that causal and semantic text relations play a role in first language (L1) reading comprehension, such that causally and semantically close sentences are read faster and recalled better (e.g., Keenan et al., 1984; Myers et al., 1987; Todaro et al., 2010; Wolfe et al., 2005).

Although second language (L2) reading studies have often focused on lower level processing such as word recognition and syntactic parsing (Chen et al., 2015; Grabe, 2009; Koda, 2005, for a review), some studies have demonstrated that causal and semantic text relations also have an impact on L2 reading (Horiba et al., 1993; Horiba, 1996; Nahatame, 2017, 2018, 2020). In a self-paced reading experiment, Nahatame (2018) demonstrated that Japanese learners of English processed L2 paired sentences that were high in causal and semantic relatedness faster than those that were low in relatedness. Although this finding provides evidence for the role of causal and semantic relatedness in L2 text processing, it has a methodological limitation: the text was presented sentence by sentence. This method prohibits participants from rereading the sentences that they have already processed, which is a typical strategic behavior during natural reading (Hyönä et al., 2003).

The current study aimed to extend Nahatame's (2018) research by overcoming this methodological limitation using eye tracking. Eye tracking has recently received considerable attention from researchers in the field of applied linguistics, L2 acquisition, and bilingualism (Conklin et al., 2018; Godfroid, 2019; Roberts & Siyanova-Chanturia, 2013). Analyzing readers' eye movements enables researchers to assess how much cognitive effort is being allocated to processing the input under more natural reading conditions and to obtain fine-grained information regarding text processing. Leveraging these strengths, the goal of the current study is to provide a better understanding of how causal and semantic relations between sentences influence L2 text processing. The findings of this study hold theoretical, methodological, and pedagogical implications for L2 reading in terms of how readers process connected text depending on sentence relatedness.

The Role of Causal and Semantic Text Relations in Reading

Causal relations

The role of causal relations has been emphasized in several models of text comprehension. Trabasso and colleagues' causal network model provides an account of how readers construct coherent narrative comprehension based on the causal relations among the events described in a text (Trabasso & Sperry, 1985; Trabasso & van den Broek, 1985). Similarly, according to Graesser et al.'s (1994) constructionist theory, readers attempt to develop a coherent text representation by understanding the causal relations among the actions, events, and states described in a text.

Given such theoretical importance, the role of causal relations has been widely examined in empirical studies on L1 reading. Classical studies typically prepare two-sentence texts that vary in the degree of causal relatedness (e.g., Keenan et al., 1984; Myers et al., 1987). For instance, the text "Dick kicked the living room wall in his fury. A mirror fell down and shattered on the floor" has a higher causal relatedness between sentences than the text "Dick came home late after his evening class. A mirror fell down and shattered on the floor" (Myers et al., 1987, pp. 461–462). These studies demonstrated that sentences high in relatedness were processed more quickly and recalled more often than those low in relatedness, suggesting the role of causal relations in text processing and memory. More recent studies have replicated this role with a

wider range of readers and tasks (Barnes et al., 2015; Todaro et al., 2010; Wolfe et al., 2005). Thus, previous studies provided converging evidence that readers strongly rely on causal text relations to establish coherence in comprehension.

Semantic relations

In addition to causal relations, researchers in cognitive science have recognized the role of semantic text relations. Kintsch's (1988, 1998) model of comprehension assumes that the argument overlap across text propositions contributes to constructing a network representation of text ideas. This overlap can be seen as a type of more general semantic relatedness. Additionally, O'Brien, Myers, and colleagues' resonance model supposes that earlier text information becomes more easily available and retrievable during comprehension when there is conceptual overlap, which is regarded as deeper semantic relations, between the earlier information and that being processed (e.g., Albrecht & O'Brien 1993; Myers et al. 1994). In these theories, the activation of semantic associations is considered a relatively passive or automatic process.

Since the late 1990s, studies have adopted computational approaches to assess semantic text relations. Although there are several techniques, latent semantic analysis (LSA), introduced by Landauer and colleagues (Landauer & Dumais, 1997; Landauer et al., 1998), has been the most well-known and widely used method in cognitive science. By applying statistical computations to a large corpus of text documents, LSA represents the meaning of words as vectors in a high-dimensional semantic space derived from the corpus. An important aspect of the LSA theory and mechanism is that it considers both the direct and indirect co-occurrence of words in the text, allowing for the representation of a word's meaning based on its usage and reference to other words within discourse contexts (Landauer & Dumais, 1997). LSA is also able to represent the meanings of text units larger than words, such as sentences and paragraphs, typically by measuring the average of the vectors for the individual words making up the units.

LSA calculates the similarity of the meanings of two textual units as the cosine of the angle between their vectors. The units are considered more semantically related when the cosine is closer to 1 and less so when the cosine is closer to -1 (in practice, 0). Because LSA represents the contextual-usage meaning of words, semantic relations assessed by this technique are much deeper than surface-level contingencies such as exact lexical overlaps and synonym relations.

Studies have demonstrated that semantic relations assessed by LSA can predict human text processing and understanding (e.g., Foltz et al., 1998; Todaro et al. 2010; Wolfe et al., 2005). Wolfe et al. (2005) prepared several pairs of sentences as in Myers et al.'s (1987) study, but they manipulated semantic relatedness between sentences using LSA in addition to causal relatedness. For instance, the text, "Dick kicked the living room wall in all his fury. A mirror fell down and shattered on the floor" is higher in semantic relatedness between sentences than the text, "Dick felt the earthquake hit with sudden force. A mirror fell down and shattered on the floor"; although both are high in causal relatedness. They showed that the second sentences of the texts were processed more quickly and recalled better when they were high in causal and semantic relatedness to the first sentences. However, the effects of semantic relatedness on processing were limited in that they were only observed when causal relatedness was manipulated between participants, and participants read the text in the low-relatedness condition. Thus, Wolfe et al.

(2005) concluded that the causal relatedness effects essentially override semantic relatedness during text processing, except when causal relatedness are low and not salient to the reader.

Todaro et al. (2010) conducted a coherence judgment task using materials similar to those used by Wolfe et al. (2005). Participants read paired sentences, then judged the coherence of each text on a 5-point scale. The results showed that causal relatedness had a greater impact on skilled readers' judgments and sentence-reading times, whereas semantic relatedness had a more prominent impact on less skilled readers' judgments and sentence-reading times. Todaro argued that this result suggests that skilled readers focused more on logical relations to establish coherence at a deeper level in comprehension, whereas less skilled readers failed to exclude semantically related but contextually irrelevant concepts from comprehension.

Causal and semantic relations in L2 reading

A few studies have also demonstrated the role of causal and semantic relations in L2 reading (Horiba et al., 1993; Horiba, 1996; Nahatame, 2017, 2018, 2020). Horiba et al. (1993) indicated that L2 readers also use the causal network structure of a text in trying to construct a coherent text representation. Similarly, Horiba (1996) adopted Myer et al.'s (1987) sentence pairs and showed that L2 Japanese learners recalled causally close sentences better than causally distant sentences when reading for studying. These results suggest that the role of causal relatedness in L2 reading is similar to that in L1 reading.

A series of Nahatame's studies examined the simultaneous impact of causal and semantic relations on L2 reading by adopting Wolfe et al.'s (2005) sentence pairs. Nahatame (2017) conducted a coherence judgment task with Japanese learners of English, as in Todaro et al. (2010), and found that L2 readers also regarded sentences high in causal and semantic relatedness as more coherent. In addition, when causal relatedness was manipulated between participants, less skilled readers weighed semantic relatedness more heavily in their judgment than skilled readers, whereas skilled readers weighed causal relatedness more heavily than less skilled readers.

Nahatame's (2018) Experiment 2 examined text memory and processing among Japanese learners of English. The results showed facilitative effects of causal and semantic relatedness on both the processing time and recall of the second sentences. However, there was an interaction effect of causal and semantic relatedness on processing. Although higher causal relatedness generally facilitated the processing time, the effects were more prominent for sentences high in semantic relatedness. On the other hand, the effects of semantic relatedness were only prominent for more closely causally related sentences. Although this finding is not perfectly consistent with previous L1 findings (Wolfe et al., 2005), they have in common in that the effects of semantic relatedness on text processing are limited by causal relatedness.

However, as mentioned earlier, it should be noted that Nahatame's (2018) experiment employed the sentence-by-sentence reading paradigm. While this method enables us to associate processing time measures with individual sentences, it has several issues that must be considered, as noted in Nahatame's study itself. First, it may increase the load on participants' processing systems by not presenting an entire text at once; therefore, the results obtained from employing this type of

paradigm may not be generalizable to typical reading (Rayner, 1998). Second, it prevents readers from returning to previous sentences in the text, which the readers typically do to establish coherence in comprehension (Hyönä et al., 2003). Third, because this method provides only one measure (i.e., sentence-reading times), it is difficult to determine the exact time course of the causal and semantic relatedness effects on text processing. The use of eye tracking described below has the potential to resolve these issues.

The Use of Eye Tracking in Reading Research

Eye tracking has been widely adopted in cognitive science studies because monitoring readers' eye movements provides valuable information regarding moment-to-moment comprehension processes during reading (Just & Carpenter, 1980; Rayner, 1998, 2009). It has also been employed in a growing number of studies in the fields of applied linguistics and L2 studies (Conklin et al., 2018; Godfroid, 2019; Roberts & Siyanova-Chanturia, 2013). There are two basic characteristics of eye movement during reading: *fixations* and *saccades* (Godfroid, 2019). During fixations, the eyes remain on a certain part of the text (e.g., a single word). Saccades are continuous, rapid eye movements between one fixation point and another while reading. Although saccades are usually made from left to right while reading an English text, readers sometimes make backward eye movements from right to left called *regressions*. Short regressions (e.g., to a few letters back within a word) suggest processing difficulty associated with the currently fixated word, whereas long regressions (e.g., to earlier words in the sentences) indicate a comprehension problem in larger parts of text (Roberts & Siyanova-Chanturia, 2013). Longer regressions to a previous sentence are sometimes referred to as *lookbacks*, which have been a valuable measure for discourse processing studies (Hyönä et al., 2003; see also below). Fixations and saccades are used to calculate common eye movement measures such as the number of fixations, fixation durations, saccade length, and regression frequency.

Eye tracking can be more advantageous than other online methodologies, such as think-aloud or moving-window techniques used in self-paced reading (Godfroid, 2019; Hyönä et al., 2003; Rayner, 1998). First, it presents an entire text to participants, therefore allowing them to freely move their eyes during reading, whenever and wherever they want. This enables researchers to examine text processing during natural and uninterrupted reading comprehension. Second, eye tracking provides a variety of processing measures rather than a single, unique measure such as sentence-reading times. Researchers can feel more confident in their findings when “different measures essentially provide converging evidence that an effect is real” (Godfroid, 2019, p. 217). Moreover, when the effect shows up in one measure but not the other, it can be theoretically interesting because these measures are indicative of different text processing stages (Conklin et al., 2018). Some measures reflect the initial or early stages of text processing (e.g., first-fixation duration, skipping rate), whereas others reflect intermediate or late text processing (e.g., regressions, rereading times).

Although eye tracking has typically been employed to study lexical and sentence processing during reading, researchers have also recognized its usefulness for investigating the processing of larger portions of text. Rayner et al. (2006) stated that “the time is ripe for more comprehension studies to use eye movement data to understand discourse processing” (p. 252).

Hyönä et al. (2003) asserted “the potential applicability of eye tracking to study global text processing” (p. 330) by discussing eye movement measures that tap into different aspects of global text processing.

Numerous studies have employed eye tracking to investigate discourse processing (e.g., Hyönä et al., 2003; Kaakinen et al., 2002; Rayner et al., 2006; Rinck et al., 2003; van der Schoot et al., 2012). Many of these studies defined individual sentences—rather than single words or phrases—as the area of interest (i.e., textual unit for eye movement analysis). Typically, they analyzed the first-pass reading time of the target sentence (i.e., the total duration of fixations on the sentence until readers move on to the next sentence or return to the previous sentence), second-pass reading time of the target sentence (i.e., the total duration of fixations on the sentence made after the first-pass reading), and lookbacks from the target to previous sentences.

As for the lookback, Hyönä et al. (2003) defined it as “any fixation on text prior to the most recently-fixated target sentence, including backward and forward fixations as long as they do not return on the target sentence” (p. 331), whereas regressions, the more common term, usually refer to backward fixations within sentences. Studies analyzed the number or percentages of participants showing lookbacks (i.e., whether the individual participant made lookbacks or not), rather than the number and duration of lookbacks, given that the lookbacks do not frequently occur (Rinck et al., 2003; van der Schoot et al., 2012).¹

First-pass reading time reflects the initial processing when readers encounter the target sentences for the first time, whereas second-pass reading time and lookbacks are indicative of later text processing, occurring after the first-pass reading (Rinck et al., 2003; van der Schoot et al., 2012). Previous studies demonstrated that discourse-level features, such as contextual inconsistency, affect these eye-movement measures, supporting the benefit of using eye tracking to examine text processing beyond individual sentences. This motivated me to adopt eye tracking to study the detailed processes of how L2 readers process text across sentences.

The Study

Leveraging the benefits of eye tracking described above, the current study aimed to extend Nahatame’s (2018) research by examining whether the influence of causal and semantic relatedness is maintained under more natural reading conditions and identifying the time-course of the influence in L2 text processing. Specifically, it examined the effects of causal and semantic relatedness between paired sentences on the following two eye movement measures: (a) first-pass reading time of the second sentences and (b) lookbacks from the second to first sentences. Accordingly, the following research questions (RQs) were addressed:

1. Does either one of the causal or semantic relatedness between sentences influence the initial stage of L2 text processing, or both?
2. Does either one of the causal or semantic relatedness between sentences influence the later stage of L2 text processing, or both?

Method

Participants

The participants were 48 students recruited from a university in Japan. All participants provided informed consent, and the study design was approved by the Research Ethics Committee at the author's institution on January 30th, 2019 (No. 30-196).

Data from three participants were excluded owing to poor quality of eye-movement recordings throughout the experiment. These participants were identified by examining their scanpaths and eye-tracking replay data, which indicated severe misalignment between fixations and the lines of text. Thus, the following analyses were conducted on data from the remaining 45 participants.

The remaining participants were 35 undergraduate and 10 graduate students; 22 males and 23 females, aged 18–28 years ($M = 21.56$, $SD = 2.26$). All are L1 speakers of Japanese and have studied English as a foreign language (EFL) for more than six years as part of their formal education in Japan. These demographic backgrounds are almost comparable to the participants in Nahatame's (2018) study (37 undergraduate and 8 graduate students; 26 males and 19 females, aged 18–26 years [$M = 21.13$, $SD = 1.95$]), who were all L1 speakers of Japanese and had studied EFL for more than six years in Japan. The current participants were estimated to be beginning- to intermediate-level English learners based on their self-reported results on some standardized English proficiency tests.

L2 reading skill assessment

The participants' English reading proficiency was measured to factor in the variable of L2 reading skill in statistical analysis. It was assessed using the 24 multiple-choice comprehension questions given for six reading passages. This test was created based on the reading subsections of previous editions of the Eiken test, a standardized English proficiency test widely administered in Japan. The reliability was sufficient after deleting three low-discriminability items (Cronbach's $\alpha = .83$). Descriptive statistics for the test scores were as follows: M [95%CI] = 14.38 [13.33, 15.43], $SD = 3.49$, max = 19, min = 4, skewness = -0.99 , kurtosis = 0.55.

Materials

The experimental texts comprised 20 sets of paired sentences adopted from Nahatame's (2018) Experiment 2, which were originally developed as part of Wolfe et al.'s (2005) study. Because the original texts were written for L1 English readers, Nahatame's study specifically selected only those appropriate for English learners in terms of both language and cultural content. These texts were then revised by controlling word frequency and simplifying syntactic structures so that they are comprehensible for English learners (see Nahatame, 2017, 2018 for more details). All sets of experimental texts are shown in Appendix.

Table 1 shows the sample set of experimental texts. Each set includes four types of first (prime) sentences and one common second (target) sentence. The four prime sentences were manipulated

in terms of their causal and semantic relatedness to the target sentence (i.e., high or low). For example, sentences 1a and 1b in Table 1 are high in causal relatedness to the target sentence because they state an event (i.e., finding nothing to read or looking for recipes) that is likely to cause the event described in the target sentence (i.e., going to the bookstore to get new books). On the other hand, the events described in sentences 1c and 1d (i.e., going to the library or having an office dinner party), which are low in causal relatedness to the target sentence, are less likely to cause the event described in the target. Similarly, sentences 1a and 1c are high in semantic relatedness to the target, as the content words of the prime sentences (e.g., “library” and “read”) are more semantically related to the words in the target sentences (e.g., “books”) compared to those of sentences 1b and 1d, which are low in semantic relatedness to the target.

Table 1

A Sample Set 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. Adopted from Nahatame (2017, 2018).

Although the causal manipulation of the experimental texts was previously verified by Wolfe et al.’s (2005) and Nahatame’s (2018) preliminary studies, the current study conducted a similar preliminary survey to confirm the manipulation and obtain a larger dataset for the causal relatedness rating. All procedures were similar to the previous preliminary studies, except that participants were recruited through Amazon Mechanical Turk (AMT). A total of 106 AMT workers residing in the United States and holding US bachelor’s degrees were recruited as skilled English readers. They were given a why-question constructed from the target sentence (e.g., “Why did Mary go to the bookstore to get new books?”) and four answer statements created from the prime sentences (e.g., “Because she could not find anything to read in the library”). They were then asked to judge the quality of each answer statement, ranging from 1 (*very bad answer*) to 6 (*very good answer*). The ratings indicated high reliability (Cronbach’s $\alpha = .90$), and the results showed that sentences high in causal relatedness were rated significantly higher ($M = 4.89$ [4.74, 5.04], $SD = 0.58$) than those low in causal relatedness ($M = 2.23$ [2.08, 2.38], $SD = 0.59$), with a large effect size, $t(78) = 20.38$, $p < .001$, $d = 4.56$.

Similar to Wolfe et al.’s (2005) approach, semantic manipulation was verified by employing LSA to obtain cosines between the prime and target sentences. The current study used the cosines obtained in Nahatame’s (2018) research, which conducted LSA based on the semantic space constructed from the Touchstone Applied Science Associates, Inc. (TASA) corpus. Sentences high in semantic relatedness received significantly higher cosines ($M = 0.29$ [0.26, 0.32], $SD = 0.11$) than those low in semantic relatedness ($M = 0.02$ [0.00, 0.04], $SD = 0.04$), with a large effect size, $t(78) = 14.32$, $p < .001$, $d = 3.20$.

Table 2 shows the descriptive statistics of the subjective ratings for causal relatedness and LSA cosines for semantic relatedness. The following analysis includes these ratings and cosines as continuous variables for causal and semantic relatedness, respectively, rather than the categorical variables of high and low relatedness.

Table 2

Descriptive Statistics of Subjective Ratings for CR and LSA Cosines for SR

	<i>M</i> [95%CI]	<i>SD</i>	<i>Min</i>	<i>Max</i>	skewness	kurtosis
Ratings (CR)	3.56 [3.23, 3.88]	1.46	1.59	5.60	0.03	-1.73
Cosines (SR)	0.16 [0.12, 0.19]	0.16	-0.07	0.55	0.69	-0.42

Note. CR = causal relatedness; SR = semantic relatedness. Ratings range from 1 to 6, and Cosines range from -1 to 1.

Each text set was paired with a simple yes-no comprehension question (e.g., “Was Mary looking for books?” for a sample set in Table 1) to ensure that the participants engaged in text reading. The second sentence was queried so that the same questions could be used among all four variations of experimental texts. A similar number of filler passages were prepared to divert participants’ attention from the experimental text manipulation (i.e., causal and semantic relatedness). Four lists of text items were created by counterbalancing the variables of causal and semantic relatedness, each of which included five experimental passages in one of the four conditions (high or low causal relatedness \times high or low semantic relatedness) along with all filler passages.

Procedure

The experiment was conducted as part of a larger project that collected eye-movement data of L2 readers of English. Before participating in the eye-tracking experiment, participants completed the English reading test on paper, for which a 35-minute time period was allocated. In the eye-tracking section, the participants’ eye movements were recorded with an EyeLink® 1000 Plus eye tracker (SR Research Ltd., Canada). They sat approximately 70 cm from a 21.5-inch computer screen on which each text was presented in its entirety, in a 20-point font. The texts were double-spaced to easily spot drift in the recording and potentially correct it. A chinrest was used to fix the participants’ head positions. Before beginning the trials, the participants’ eye movements were calibrated and validated using a standard 9-point grid. The participants then completed practice trials to familiarize themselves with the procedure before the experimental session.

Participants were randomly assigned one of the four text item lists and were instructed to read each text at their own pace for comprehension. They pushed the space bar on the keyboard to indicate that they were ready to read; each text then appeared at the center of the screen. They pushed the space bar again when they finished reading. A comprehension question was then presented on the screen, and the participants answered the questions by pressing the slash (/) key for yes and the z key for no. Feedback was given to participants as either “Correct” or

“Incorrect” stated on the screen. Each trial was repeated for all text items presented in a random order.

Eye movement measures

As in past eye-tracking studies on discourse processing (e.g., Kaakinen et al., 2002; Rinck et al., 2003; van der Schoot et al., 2012), this study defined individual sentences as the analysis unit of eye movement. As described above, the following two eye-movement measures were computed (Hyönä et al., 2003): (a) first-pass reading time of the target (second) sentences and (b) the occurrence of lookbacks from the target to the prime (first) sentences. The second-pass reading time of the target sentences was not analyzed because second-pass reading seldom occurred for these sentences. For a similar reason, this study analyzed the occurrence of lookbacks (i.e., whether readers made lookbacks or not, a binary variable) rather than their numbers or duration (Rinck et al., 2003; van der Shoot et al., 2012); such lookbacks only occurred in 34% of all trials in the experiment.

First-pass reading time is the summed duration of all fixations on the target sentences until readers finish processing the text or look back to the previous sentence.² This eye fixation measure reflects the initial processing when the target sentence is first encountered (van der Shoot et al., 2012). This measure was analyzed to answer the RQ1, which concerned the effects of causal and semantic relatedness on the initial stage of L2 text processing. If causal or semantic relatedness to prime sentences influences the initial processing difficulty of the target sentences, the first-pass reading times of the sentences should vary according to relatedness.

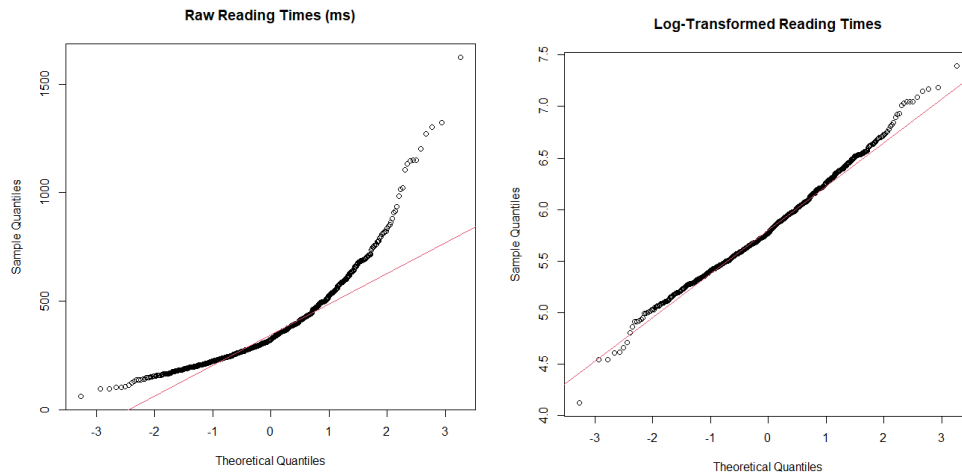
The occurrence of lookbacks indicates whether readers looked back to the prime sentences in the sets after they finished reading the target sentences. Lookbacks were counted only if they lasted longer than 250ms, which approximates the current participants’ average fixation time, given that no meaningful information is likely to be obtained from fixations below this threshold. This eye-fixation measure reflects the strategic action of readers who try to monitor their comprehension and is indicative of later integrative text processing (Rinck et al., 2003; van der Shoot et al., 2012). Lookbacks should occur when readers have difficulty integrating the target sentences with the developing mental representation and try to resolve it. Thus, this measure was analyzed to answer the RQ2, which concerned the effects of causal and semantic relatedness on the later stage of L2 text processing.

Statistical analysis

Given that target sentence length varied slightly among the experimental items, the first-pass reading time (ms) for the target sentence in each item was divided by the number of syllables in that sentence to control for sentence length ($M = 371.35$ [361.35, 381.35], $SD = 182.28$, skewness = 1.94, kurtosis = 6.17). A visual examination indicated that raw reading time data were not normally distributed, hence they were log-transformed to decrease data variability and bring it closer to the normal distribution (see Figure 1).

Figure 1

The Q-Q Plots of Raw and Log-Transformed First-Pass Reading Times (Per Syllable)



The reading time data was analyzed using linear mixed models by running the `lme4` package (Bates et al. 2014), version 1.1.26 in R version 4.0.5. Fixed effects included the variables of causal relatedness (subjective ratings obtained from the AMT study), semantic relatedness (LSA cosines), and their interaction. The models also included the fixed effect of participants' L2 reading proficiency (reading test scores) as a control variable, given that fixations and saccades are contingent on participants' reading skills (Conklin et al. 2018; Rayner, 1998, 2009). Each variable was centered on its mean to reduce collinearity.

The maximal model was constructed with random intercepts for participants and items as well as random slopes for the variables corresponding to all fixed effects, the length (number of syllables) and word frequency³ of target sentences, and the accuracy of comprehension questions. Because the maximal model showed over-parameterization, the model was simplified by excluding random effect parameters with lower variances when doing so led to no significant goodness-of-fit loss (Bates et al., 2015) and the selected models successfully converged. The final best-fitting model was then updated by removing data points with absolute standard residuals in excess of 2.0 *SD* to trim the outliers (Baayen & Milin, 2010; Godfroid 2019), resulting in a loss of less than 5% of the data. The package `lmerTest` (Kuznetsova et al., 2015) was employed to calculate p values for the fixed effects in the models, and the package `MuMin` (Nakagawa & Schielzeth, 2013) was used to obtain a marginal r squared (R^2_{marginal}) and a conditional r squared ($R^2_{\text{conditional}}$) for the models.

The lookback data were analyzed in a manner similar to the first-pass reading time analysis, except that the `glmer` function (generalized linear mixed model) was employed because lookback is a binary variable.

Results

The correct answer rate for the comprehension questions was quite high (91%). Because the preliminary analysis found no significant effects of comprehension question accuracy on any eye movement measures ($ps > .10$), the following analysis was conducted regardless of comprehension question responses.

First-pass reading time (RQ1)

Table 3 displays the results of the final model of the first-pass reading times for the target sentences. There were significant main effects of causal relatedness and L2 reading proficiency, indicating that first-pass reading times decreased as causal relatedness was higher and participants' L2 reading proficiency increased. However, the effects of causal relatedness were qualified by a significant interaction with semantic relatedness, which is visualized in Figure 2.

As shown in the left panel of Figure 2, although the first-pass reading times of the target sentences were generally shorter when preceded by more closely causally related sentences, the effects were increasingly prominent for more semantically related sentences. On the other hand, as shown in the right panel of Figure 2, higher semantic relatedness resulted in decreased reading times only for more causally related sentences. Table 4 presents the predicted data values at specific levels of causal and semantic relatedness to confirm these plot interpretations.⁴

Table 3

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>			
	<i>Estimate</i>	<i>SE</i>	<i>t</i>	<i>p</i>	<i>By Subject</i>		<i>By Item</i>	
					<i>Variance</i>	<i>SD</i>	<i>Variance</i>	<i>SD</i>
Intercept	5.82	0.05	119.76	< .001*	0.04	0.20	0.03	0.17
CR	-0.06	0.01	-8.40	< .001*	_____	_____	_____	_____
SR	-0.12	0.07	-1.76	.078	_____	_____	_____	_____
RP	-0.03	0.01	-3.08	.004*	_____	_____	_____	_____
CR × SR	-0.10	0.05	-2.06	.039*	_____	_____	_____	_____

Note. CR = causal relatedness; SR = semantic relatedness; RP = reading proficiency. All factors were centered on their mean values. Model formula: $\log(\text{FPRT}) \sim \text{cCR} * \text{cSR} + \text{cRP} + (1 | \text{subject}) + (1 | \text{item})$. $R^2_{\text{marginal}} = .11$; $R^2_{\text{conditional}} = .53$.

Figure 2

Plots of the Interaction Effect Between Causal Relatedness (CR) and Semantic Relatedness (SR) in the Model of First-Pass Reading Times. The Shaded Band Indicates a 95% Point-Wise Confidence Interval.

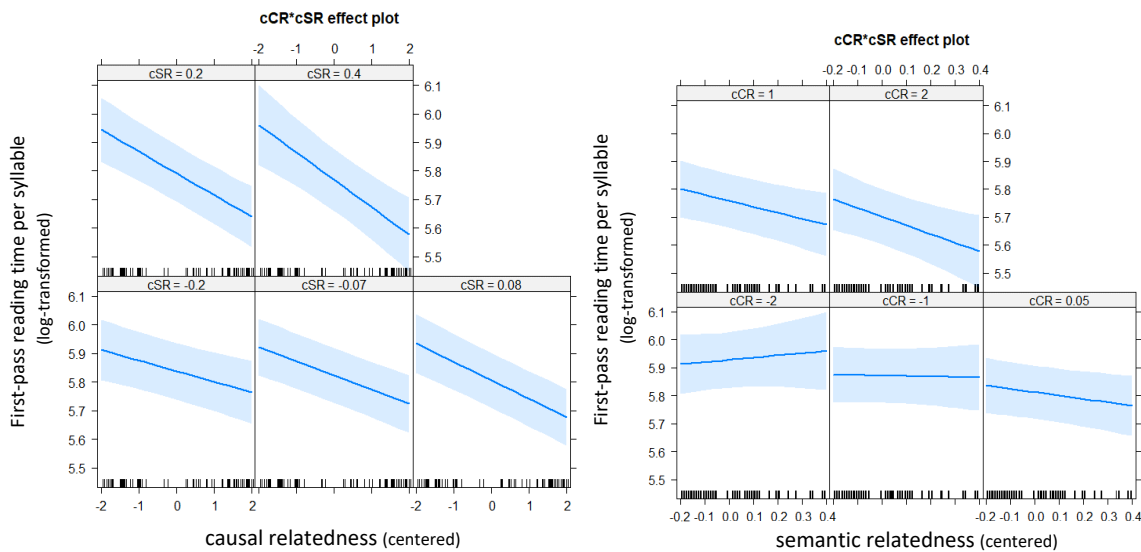


Table 4

First-Pass Reading Times (per Syllable, Log-Transformed) Predicted by the Interaction Effect Between Causal and Semantic Relatedness in the Model

CR	SR				
	-0.20	-0.07	0.08	0.20	0.40
-2.00	5.91	5.92	5.93	5.94	5.96
-1.00	5.88	5.87	5.87	5.87	5.86
0.05	5.84	5.82	5.80	5.79	5.76
1.00	5.80	5.77	5.74	5.72	5.67
2.00	5.76	5.72	5.68	5.64	5.58

Note. CR = causal relatedness (centered); SR = semantic relatedness (centered).

Lookbacks (RQ2)

Table 5 presents the results of the final model for lookbacks from the target to the prime sentences. There was a significant main effect of causal relatedness, indicating that participants were less likely to make lookbacks when the target sentences were preceded by more closely causally related sentences, as illustrated in Figure 3. No other effects were significant, including those related to semantic relatedness.

Table 5

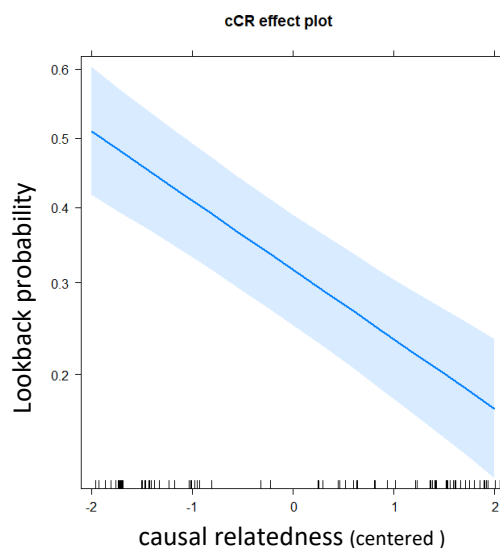
Summary Table for the Results From a Mixed-Effects Model of Lookbacks

<i>Parameters</i>	<i>Fixed Effects</i>				<i>Random Effects</i>			
	<i>Estimate</i>	<i>SE</i>	<i>z</i>	<i>p</i>	<i>By Subject</i>		<i>By Item</i>	
					<i>Variance</i>	<i>SD</i>	<i>Variance</i>	<i>SD</i>
Intercept	-0.77	0.17	-4.67	< .001*	0.50	0.70	0.19	0.44
CR	-0.41	0.06	-7.18	< .001*	—	—	—	—
SR	0.74	0.52	1.42	.157	—	—	—	—
RP	0.04	0.04	1.02	.310	—	—	—	—
CR × SR	-0.44	0.39	-1.12	.263	—	—	—	—

Note. CR = causal relatedness; SR = semantic relatedness; RP = reading proficiency. All factors were centered on their mean values. Model formula: Lookbacks ~ cCR * cSR + cRP + (1 | subject) + (1 | item). $R^2_{marginal} = .08$; $R^2_{conditional} = .24$.

Figure 3

Plots of the Effect of Causal Relatedness (CR) in the Model of Lookback Probability. The Shaded Band Indicates a 95% Point-Wise Confidence Interval.



Discussion

The results observed in this eye-tracking experiment clearly indicate whether and how L2 readers process text depending on the causal and semantic relatedness between sentences. The following section discusses the results in terms of the two RQs.

Effects on the initial text processing (RQ1)

RQ1 focused on the effects of causal and semantic relatedness on the initial stage of L2 text processing. The results of first-pass reading times indicated that both causal and semantic relatedness influenced processing when encountering the target (second) sentences. As the target sentences were more causally and semantically related to the prime sentences, L2 readers processed the target sentences faster during the first reading. However, these effects of causal and semantic relatedness interacted with each other, such that the effects of causal relatedness were greater for sentences high in semantic relatedness, whereas semantic relatedness facilitated processing only for sentences high in causal relatedness.

These results are highly compatible with those of Nahatame's (2018) sentence-by-sentence reading experiment. Thus, the current study confirms that the processing of L2 text indeed depends on both causal and semantic relatedness between sentences and that causal relatedness plays a more dominant role in the processing. Nevertheless, it is worth noting that the current eye-tracking investigations extended this finding to more natural, uninterrupted L2 reading processes.

Effects on the later text processing (RQ2)

RQ2 focused on the effects of causal and semantic relatedness on the later stage of L2 text processing. Importantly, the results of lookbacks indicated the different time course of the causal and semantic relation effects, which could not be identified in Nahatame's (2018) experiment. In the current study, when target sentences were less causally related to the previous sentences, participants not only spent more time on the target sentences but also looked back to the previous sentences. This suggests that the effects of causal relatedness were persistent from the initial to the later text processing, supporting the significant role of causal relations in L2 reading comprehension (Horiba et al., 1993; Horiba, 1996; Nahatame, 2018). On the other hand, this effect on lookbacks suggests that the increased sentence-reading times for less causally related sentences observed in Nahatame's (2018) study is partly attributable to the fact that the sentence-by-sentence reading paradigm did not allow for lookbacks. I return to this point in the following section to discuss the methodological implications of this finding.

Semantic relatedness, however, did not have an impact on lookbacks after reading the target sentences. This suggests that the effects of semantic relatedness are limited compared to causal relatedness, in terms of not only the prominence but also the time course of the effects. The absence of semantic relatedness effects on later and strategic text processing seems in line with the view that semantic associations to text information are generated quickly and automatically (Kintsch, 1988, 1998; Myers et al., 1994) and their effects may decrease with delay because surface-level text representation, where semantic associations are activated and stored, is short-lived (Kintsch et al., 1990).

Conclusion and Implication

The current study aimed to extend Nahatame's (2018) research on the role of causal and semantic relatedness in the processing of L2 paired sentences using eye tracking. Nahatame's study, using the sentence-by-sentence reading paradigm, found the effects of both causal and semantic relatedness on processing time and that causal relatedness effects are more prominent than semantic relatedness effects. The eye-tracking experiment reported here confirmed these findings; more importantly, it extended them to a more natural reading situation where the complete text was visually available. The current findings support the view that L2 readers engage in this kind of discourse processing if they have no major difficulty with lower reading processing (Horiba, 1996; Nahatame, 2018).

The current study also contributes to the literature by demonstrating the different time courses of causal and semantic relatedness effects. That is, both causal and semantic relatedness of paired sentences influence processing when encountering later sentences, but only causal relatedness has an impact on lookbacks to previous sentences afterward. This concurs with the studies that demonstrated the significant role of causal relations in L2 reading (Horiba et al., 1993; Horiba, 1996; Nahatame, 2018); however, the current study is distinct from these past investigations in that it offers unique information about the time course of the effects, compared to semantic relatedness.

Given that causal relatedness among sentences more strongly and persistently influences L2 text processing, this study suggests that L2 textbook designers consider the cause-and-effect relationships of sentences if they want to create reading texts that are more smoothly processed and less likely to cause lookbacks during reading. Text readability and processability can be also improved, though perhaps to smaller extent, by increasing semantic overlap between sentences. L2 textbook designers and reading instructors should bear in mind that, in addition to linguistic features such as lexical properties and syntactic structures, sentence relatedness also influences learners' text processing (Nahatame, 2018).

From a methodological perspective, the current findings indicate that the causal and semantic relatedness effects on reading times obtained using sentence-by-sentence presentation are valid in that these effects were indeed observed when the entire text was presented to readers. They also suggest, however, that the effects obtained by the sentence-by-sentence reading method may lack ecological validity to some extent, as discussed in Rinck et al. (2003). Even if participants want to reread the previous sentences as observed for the less causally related sentences in the current study, the sentence-by-sentence presentation does not allow for such rereading; instead, it requires participants to rely on text memory. This can be advantageous for researchers in that it may increase the size of the effect of interest by combining the effects caused by initial processing difficulty and those caused by the recovery actions from that difficulty. However, the results obtained by this method are "not necessarily representative of the processes occurring during everyday reading of text" (Rinck et al., 2003, p. 85). Therefore, the current study suggests that L2 reading researchers, ideally, adopt eye tracking when studying the processing of connected text. Given that eye tracking is not always easily available to many L2 researchers, at a minimum, they should keep in mind the lack of validity of the sentence-by-sentence reading method when interpreting the results obtained by this method.

Limitations and Directions for Future Research

This study has several limitations that present promising avenues for future research. First, this study is limited in the number of participants involved, which did not allow us to examine whether the effects of causal and semantic relatedness interact with participants' individual differences, such as reading skills (Nahatame, 2017; Todaro et al., 2010). Although eye-tracking experiments require substantial time and effort to conduct, increasing the sample size enables us to consider some confounding variables of individual differences while maintaining adequate power of analysis.

Second, this study only focused on the processing of text rather than the product of text comprehension (e.g., text memory after reading). However, what can be inferred from the processing sometimes differs from what can be inferred from the product (Horiba, 2013). Thus, it is beneficial to analyze both processing and product data within a single research study by conducting, for example, eye tracking and recall task. Particularly, such studies are needed to better understand the time course of semantic relatedness effects, given that the current study found the effects only on initial text processing but not on later processing, whereas some previous studies indicated the effects of semantic relatedness on recall performance (Nahatame, 2018; Wolfe et al., 2005).

Third, although LSA has traditionally been used to assess text semantic relatedness, recent studies on computational linguistics have proposed similar but more sophisticated methods such as latent Dirichlet allocation (LDA; Blei et al., 2003) and Word2Vec (Mikolov et al., 2013). Because some of these new methods may be able to better predict human performance on psycholinguistic tasks than LSA (Mandera et al., 2017), it would be interesting to examine whether the effects of semantic relations vary depending on these different methods. In addition, given that the semantic relation effects (on first-pass reading times in the interaction with causal relatedness) were significant but not sufficiently clear in the current study, it is valuable to investigate these effects in future studies.

Finally, although the current study examined the processing of paired sentences to extend Nahatame's (2018) study, future studies should explore longer texts such as those L2 learners usually encounter in everyday life. This will extend our knowledge of the effects of causal and semantic relatedness to those beyond adjacent sentences (Crossley et al., 2019; Nahatame, 2020; Wole et al., 2005), offering more practical pedagogical implications.

Notes

¹ Although Rinck et al. (2003) and van der Schoot et al. (2012) used the term "regressions" in their studies, it indicates the same reading behavior that Hyönä et al. (2003) defined as "lookbacks." In addition, Hyönä et al. (2003) discriminated between "lookback" and "lookback time" (i.e., duration of lookbacks).

² The current study adopted the extended definition of first-pass reading times proposed by Hyönä et al. (2003). According to this definition, when readers look away before finishing reading all parts of the target sentence and then return to the sentence, the processing time for the remaining part of the sentence is added to the first-pass reading time for the initial part of the sentence (called extended first-pass reading time). This avoids calculating first-pass reading times that are unreasonably short. Nevertheless, this extended definition was only applied to less than 4% of all trials in the current study.

³ Word frequency was derived from the JACET 8000 (The JACET Basic Word Revision Committee, 2003). This is a frequency word list developed for Japanese learners of English.

⁴ I did not conduct a subsequent analysis of the interaction effect because this study focused on the pattern of the interaction effect—sufficiently shown in the plot and the table of predicted values—rather than the particular value where the effect of a given variable is significant or not (Dawson, 2014). Additionally, this study avoided the problems of multiple testing caused by dividing the data into subsets for subsequent analysis.

Data Availability Statement

The materials that were used in this study are openly available in IRIS at <https://www.iris-database.org/iris/app/home/detail?id=york:939993>

The data that support the findings of this study and the code used to perform the statistical analyses are available at the Open Science Framework (OSF) at <https://osf.io/mcq6u/>

References

- Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1061–1070. <https://doi.org/10.1037/0278-7393.19.5.1061>
- Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28. <https://doi.org/10.21500/20112084.807>
- Barnes, M. A., Ahmed, Y., Barth, A., & Francis, D. J. (2015). The relation of knowledge-text integration processes and reading comprehension in 7th- to 12th-grade students. *Scientific Studies of Reading*, *19*(4), 253–272. <https://doi.org/10.1080/10888438.2015.1022650>
- Bates, D. M., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv:1506.04967, 1–27.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *Linear mixed-effects models using eigen and S4* (R package version 1.1-18-1). <http://project.org/package=lme4>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.
- Chen, X., Dronjic, V., & Helms-Park, R. (2015). *Reading in a second language: Cognitive and psycholinguistic issues*. Routledge.
- Conklin, K., Pellicer-Sánchez, A., & Carrol, G. (2018). *Eye-tracking: A guide for applied linguistics research*. Cambridge University Press.

- Crossley, S. A., Kyle, K., & Dascalu, M. (2019). The tool for the automatic analysis of cohesion 2.0: Integrating semantic similarity and text overlap. *Behavior Research Methods*, *51*(1), 14–27. <https://doi.org/10.3758/s13428-018-1142-4>
- Dawson, J. F. (2014). Moderation in management research: What, why, when, and how. *Journal of Business and Psychology*, *29*(1), 1–19. <https://doi.org/10.1007/s10869-013-9308-7>
- Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textual coherence with latent semantic analysis. *Discourse Processes*, *25*(2–3), 285–307. <https://doi.org/10.1080/01638539809545029>
- Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Lawrence Erlbaum Associates.
- Godfroid, A. (2019). *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. Routledge.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395. <https://doi.org/10.1037/0033-295X.101.3.371>
- Horiba, Y. (1996). The role of elaborations in L2 comprehension: The effect of encoding task on recall. *The Modern Language Journal*, *80*(2), 151–164. <https://doi.org/10.1111/j.1540-4781.1996.tb01155.x>
- Horiba, Y., van den Broek, P., & Fletcher, C. R. (1993). Second language readers' memory for narrative texts: Evidence for structure-preserving top-down processing. *Language Learning*, *43*(3), 345–372. <https://doi.org/10.1111/j.1467-1770.1993.tb00618.x>
- Horiba, Y. (2013). Task-induced strategic processing in L2 text comprehension. *Reading in a Foreign Language*, *25*(2), 98–125.
- Hyönä, J., Lorch, Jr. R. F., & 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.
- Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, *64*(1), 160–212. <https://doi.org/10.1111/lang.12034>
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2002). Perspective effects on on-line text processing. *Discourse Processes*, *33*(2), 159–173. https://doi.org/10.1207/S15326950DP3302_03
- Keenan, J. M., Baillet, S. D., & Brown, P. (1984). The effect of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 115–126. [https://doi.org/10.1016/S0022-5371\(84\)90082-3](https://doi.org/10.1016/S0022-5371(84)90082-3)
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. [https://doi.org/doi:10.1016/S0166-4115\(08\)61551-4](https://doi.org/doi:10.1016/S0166-4115(08)61551-4)
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language*, *29*(2), 133–159. [https://doi.org/10.1016/0749-596X\(90\)90069-C](https://doi.org/10.1016/0749-596X(90)90069-C)

- Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). lmerTest: Tests in linear mixed effects models. R package version 2.0-29. <http://CRAN.R-project.org/package=lmerTest>
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, *104*(2), 211–420. <https://doi.org/10.1037/0033-295X.104.2.211>
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2–3), 259–284. <https://doi.org/10.1080/01638539809545028>
- Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, *92*, 57–78. <https://doi.org/10.1016/j.jml.2016.04.001>
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv, 1301–3781.
- Myers, J. L., O'Brien, E. J., Albrecht, J. E., & Mason, R. A. (1994). Maintaining global coherence during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 876–886. <https://doi.org/10.1037/0278-7393.20.4.876>
- Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language*, *26*(4) 453–465. [https://doi.org/10.1016/0749-596X\(87\)90101-X](https://doi.org/10.1016/0749-596X(87)90101-X)
- Nahatame, S. (2017). Standards of coherence in second language reading: Sentence connectivity and reading proficiency. *Reading in a Foreign Language*, *29*(1), 86–112.
- 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. <https://doi.org/10.1111/modl.12466>
- Nahatame, S. (2020). Revisiting second language readers' memory for narrative texts: The role of causal and semantic text relations. *Reading Psychology*, *41*(8), 753–777. <https://doi.org/10.1080/02702711.2020.1768986>
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. <https://doi.org/10.1080/17470210902816461>
- Rayner, K., Chace, K., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*(3), 241–255. https://doi.org/10.1207/s1532799xssr1003_3
- Rinck, M., Gamez, E., Diaz, J. M., & De Vega, M. (2003). Processing of temporal information: Evidence from eye movements. *Memory & Cognition*, *31*(1), 77–86. <https://doi.org/10.3758/BF03196084>

- Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition*, 35(2), 213–235. <https://doi.org/10.1017/S0272263112000861>
- The JACET Basic Word Revision committee (Ed.). (2003). *JACET List of 8,000 Basic Words*.
- 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. <https://doi.org/10.1080/01638530903253825>
- Trabasso, T. & Sperry, L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language*, 24(5), 595–611. [https://doi.org/10.1016/0749-596X\(85\)90048-8](https://doi.org/10.1016/0749-596X(85)90048-8)
- Trabasso, T. & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, 24(5), 612–630. [https://doi.org/10.1016/0749-596X\(85\)90049-X](https://doi.org/10.1016/0749-596X(85)90049-X)
- van den Broek, P., Bohn-Gettler, C., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 123–140). Information Age Publishing.
- 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. <https://doi.org/10.1007/s11145-011-9337-4>
- 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. <https://doi.org/10.1080/0163853X.2005.9651678>

Appendix

Experimental passages used in this study

Note. These passages are based on the materials used in Wolfe et al.'s (2005) study. CR = Causal Relatedness; SR = Semantic Relatedness; CQ = Comprehension Question.

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

Target sentence: She went to the bookstore to get new books.

CQ: Was she looking for books? (Yes)

- | | | |
|----|---|---------------------|
| 2. | a. Frank lost control of his new sports car. | (CR-High / SR-High) |
| | b. Frank was looking at the screen of his new cell phone. | (CR-High / SR-Low) |
| | c. Frank bought a new sports car. | (CR-Low / SR-High) |
| | d. Frank bought a new cell phone. | (CR-Low / SR-Low) |

Target sentence: He got into an accident with a truck.

CQ: Did he have an accident with a bicycle? (No)

3. a. Frank thought Mary would be a great wife for him. (CR-High / SR-High)
 b. Frank knew that Mary had a lot of money. (CR-High / SR-Low)
 c. Frank met Mary at a divorce support group. (CR-Low / SR-High)
 d. Frank met Mary at the bank. (CR-Low / SR-Low)

Target sentence: He decided to propose marriage to Mary.

CQ: Did he ask Mary to marry him? (Yes)

4. a. Mary was very tired at school because she stayed up all night. (CR-High / SR-High)
 b. Mary did not go home until 4:00 a.m. (CR-High / SR-Low)
 c. Mary had to finish her homework assignment by the following week. (CR-Low / SR-High)
 d. Mary met a handsome, tall guy at the party. (CR-Low / SR-Low)

Target sentence: She fell asleep during class.

CQ: Did she stay awake during class? (No)

5. a. Frank shot with the gun at the police cars. (CR-High / SR-High)
 b. Frank drank too much before driving home. (CR-High / SR-Low)
 c. Frank was very upset when he saw the police cars. (CR-Low / SR-High)
 d. Frank went to a party with some friends. (CR-Low / SR-Low)

Target sentence: He was arrested and put in jail.

CQ: Did he have to go to jail? (Yes)

6. a. Mary ate some food that had a spoiled taste. (CR-High / SR-High)
 b. Mary was nervous about giving a speech in class. (CR-High / SR-Low)
 c. Mary wanted to try the food at the new restaurant. (CR-Low / SR-High)
 d. Mary went to school today. (CR-Low / SR-Low)

Target sentence: She felt sick to her stomach.

CQ: Did she feel good? (No)

7. a. Mary found that her baby was very sick. (CR-High / SR-High)
 b. Mary's husband fell to the floor. (CR-High / SR-Low)
 c. Mary recently became a nurse. (CR-Low / SR-High)
 d. Mary's husband was working in his room. (CR-Low / SR-Low)

Target sentence: She called her family doctor immediately.

CQ: Did she call her friends? (No)

8. a. Frank remembered that it was his wedding anniversary. (CR-High / SR-High)
 b. Frank was two hours late on his way home. (CR-High / SR-Low)
 c. Frank was weeding the garden in the back yard. (CR-Low / SR-High)
 d. Frank was preparing for a presentation at work. (CR-Low / SR-Low)

Target sentence: He decided to buy his wife some flowers.

CQ: Did he decided to buy a ring? (No)

9. a. Frank found typing errors in his memo. (CR-High / SR-High)
 b. Frank's coffee was very cold this morning. (CR-High / SR-Low)

- c. Frank was dictating an important memo. (CR-Low / SR-High)
 d. Frank asked for a cup of hot coffee. (CR-Low / SR-Low)
Target sentence: He was very angry at his secretary.
CQ: Did he feel angry? (Yes)
10. a. Frank kicked the living room wall, angrily. (CR-High / SR-High)
 b. Frank felt a large earthquake begin. (CR-High / SR-Low)
 c. Frank and Mary talked about the carpet in the living room. (CR-Low / SR-High)
 d. Frank noticed that the neighborhood was very quiet. (CR-Low / SR-Low)
Target sentence: A mirror fell down and was broken.
CQ: Did the mirror break? (Yes)
11. a. Mary's fingers slipped while she was slicing the steak. (CR-High / SR-High)
 b. Mary did not know the top of the fence was spiked. (CR-High / SR-Low)
 c. Mary sliced the steak with a sharp knife. (CR-Low / SR-High)
 d. Mary erected a tall fence in the back yard. (CR-Low / SR-Low)
Target sentence: She cut her hand badly.
CQ: Did she injure her hand? (Yes)
12. a. Mary received a package meant for the previous resident of her home. (CR-High / SR-High)
 b. Mary received the wrong CD from the music store. (CR-High / SR-Low)
 c. Mary checked to see if the letter had been delivered. (CR-Low / SR-High)
 d. Mary had ordered a CD from the music store. (CR-Low / SR-Low)
Target sentence: She put the mail back into the mailbox.
CQ: Did she receive an email? (No)
13. a. The cashier could not give change for Frank's hundred-dollar bill. (CR-High / SR-High)
 b. Frank's girlfriend insisted on treating him to a movie. (CR-High / SR-Low)
 c. Frank went to pay his bill. (CR-Low / SR-High)
 d. Frank arrived at the ticket counter with his girlfriend. (CR-Low / SR-Low)
Target sentence: He put his money back into his wallet.
CQ: Did he lose his wallet? (No)
14. a. Frank collapsed while he was seeing patients in his office. (CR-High / SR-High)
 b. Frank was hit on the head with a hard baseball. (CR-High / SR-Low)
 c. Frank was a doctor in a big city. (CR-Low / SR-High)
 d. Frank had a great seat at the baseball game. (CR-Low / SR-Low)
Target sentence: He was carried to a hospital in an unconscious state.
CQ: Did he carry his children to a hospital? (No)
15. a. Mary was told that people had found a shark in the sea. (CR-High / SR-High)
 b. Mary ate too much at lunch with her friends. (CR-High / SR-Low)
 c. Mary arrived at the beach in her swimming costume. (CR-Low / SR-High)

- d. Mary sat on a chair at the cafeteria. (CR-Low / SR-Low)
Target sentence: She decided to wait before going for a swim.
CQ: Did she swim soon? (No)
16. a. Frank's eyes were sensitive to strong light. (CR-High / SR-High)
 b. Frank wanted to create a romantic mood. (CR-High / SR-Low)
 c. Frank went into the well-lit bedroom. (CR-Low / SR-High)
 d. Frank remembered that it was his wedding anniversary today. (CR-Low / SR-Low)
Target sentence: He turned the lamp down low.
CQ: Did he break the lamp? (No)
17. a. Mary wanted to have a special dinner. (CR-High / SR-High)
 b. Mary passed her medical license exam. (CR-High / SR-Low)
 c. Mary loved to cook elaborate meals. (CR-Low / SR-High)
 d. Mary was studying to be a doctor. (CR-Low / SR-Low)
Target sentence: She went to a fancy restaurant.
CQ: Did she visit a restaurant? (Yes)
18. a. Frank had to pass his last exam to graduate. (CR-High / SR-High)
 b. Frank wanted to get a black belt in karate. (CR-High / SR-Low)
 c. Frank majored in psychology at his university. (CR-Low / SR-High)
 d. Frank had a black belt in karate. (CR-Low / SR-Low)
Target sentence: He began to study for his final test.
CQ: Did he start being ready for the test? (Yes)
19. a. Mary didn't take anything to wear in the pool. (CR-High / SR-High)
 b. Mary lost her luggage at the international airport. (CR-High / SR-Low)
 c. Mary was invited to swim with her friends. (CR-Low / SR-High)
 d. Mary went to Alaska during her spring vacation. (CR-Low / SR-Low)
Target sentence: She borrowed a swimming costume at the gym.
CQ: Was she at the gym? (Yes)
20. a. Mary turned on her table fan to cool the room. (CR-High / SR-High)
 b. Mary let out a very loud sneeze. (CR-High / SR-Low)
 c. It was hot in the room where Mary was working on her homework. (CR-Low / SR-High)
 d. Mary had allergy attacks during the summer. (CR-Low / SR-Low)
Target sentence: The papers on her desk blew onto the floor.
CQ: Did the papers fall onto the floor? (Yes)

About the Author

Shingo Nahatame, Ph.D., is an Assistant Professor of Education at the Faculty of Human Sciences, University of Tsukuba, Japan. His research interests include cognitive processes of

reading in English as a foreign language. His current research focuses on examining discourse processing among Japanese learners of English using eye tracking. He has published in *Language Learning*, *The Modern Language Journal*, and *Reading in a Foreign Language*.

E-mail: nahatame.shingo.gp@u.tsukuba.ac.jp. Website: <https://sites.google.com/site/snahatame/>