



OPEN ACCESS

EDITED BY

Pia Knoeferle,
Humboldt University of Berlin, Germany

REVIEWED BY

Sofiana Lindemann,
Transilvania University of Braşov, Romania
Dalia Elleuch,
University of Sfax, Tunisia
Mohamed Taiebine,
UEMF, Morocco
Chao Sun,
Peking University, China

*CORRESPONDENCE

Masaya Hosoda
✉ hzm127@gmail.com

RECEIVED 11 September 2024

ACCEPTED 26 February 2025

PUBLISHED 09 April 2025

CITATION

Hosoda M (2025) Verbs' implicit causality in
coreference and coherence processing
during L2 comprehension.
Front. Lang. Sci. 4:1494500.
doi: 10.3389/flang.2025.1494500

COPYRIGHT

© 2025 Hosoda. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC
BY\)](#). The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Verbs' implicit causality in coreference and coherence processing during L2 comprehension

Masaya Hosoda*

Department of English, Seijo University, Setagaya, Japan

The existing psycholinguistic research suggests that verbs' implicit causality (IC) elicits two types of bias: a *coreference bias*, which favors re-mentioning the causally implicated entity of the event (*she = Mary* in *Mary annoyed Lisa because she. ..*), and a *coherence bias*, which leads speakers to expect an explanation in the upcoming discourse (*Mary annoyed Lisa* is continued with *Mary sang loudly*). Of these two biases, previous second-language (L2) studies have predominantly focused on coreference bias in contexts where an upcoming explanation is explicitly signaled (*Mary annoyed Lisa because...*). The present study advances the L2 literature by examining both coherence and coreference biases in L2 comprehension. Eye-tracking and story-continuation experiments revealed that L2 learners are fundamentally weaker than native speakers in terms of coherence bias. As a result, an upcoming explanation must be explicitly signaled for IC to trigger coreference bias during online L2 processing. The findings suggest that while the underlying mechanism of IC bias functions similarly in both L1 and L2 comprehension, there is a pronounced L1–L2 difference in the ease with which an implicit explanation relation can be activated through expectation-based processing. The findings are discussed in terms of the source and time course of IC, as well as theoretical accounts of L2 prediction.

KEYWORDS

implicit causality, L2 processing, eye-tracking, coherence relations, coreference processing, expectation-based processing

1 Introduction

Discourse comprehension involves complex cognitive processes that use explicit and implicit cues. Verbs' *implicit causality* (IC) bias is a well-studied implicit cue in psycholinguistic research. For example, in the sentence fragment (1a), the causal connective *because* explicitly signals a coherence relation between clauses, with the dependent clause explaining the event in the matrix clause, which is referred to as the *explanation* relation. As the cause of the event denoted by the verb *annoyed* is imputed to the subject position (NP1 [first noun phrase]) entity (e.g., *Mary* was an annoying person, which was why she annoyed *Lisa*), *she* would typically be interpreted as *Mary*. Conversely, in fragment (1b), with *punished* as the verb, *she* is usually interpreted as *Lisa* because, in this case, the event cause is imputed to the object position (NP2 [second noun phrase]) entity (e.g., *Lisa* did something wrong, which was why *Mary* punished her).

- (1) a. Mary annoyed Lisa because she...
- b. Mary punished Lisa because she...

Previous studies on first-language (L1) comprehension have demonstrated that this *coreference bias* of IC impacts moment-to-moment (i.e., online) comprehension processes, as shown in self-paced reading and eye-tracking tasks (e.g., Koornneef and Vanberkum, 2006; Pykkönen and Järviö, 2010). Notably, IC coreference bias has been observed in various languages (English, German, Spanish, Russian, Japanese, Mandarin, and Korean; Bott and Solstad, 2014; Hartshorne et al., 2013), suggesting that IC plays a language-universal role in guiding coreference processing.

It has been demonstrated that IC induces another discourse-level bias in addition to the coreference bias. Kehler et al. (2008) discovered that speakers provided an explanation continuation (*She sang loudly*) when prompted to continue sentences with an IC verb (*Mary annoyed Bob*) in 60% of the instances. By contrast, 25% of the continuations from sentences with a non-IC verb (*Mary saw Bob*) were explanations. Thus, IC verbs create a stronger-than-normal expectation for an explanation in the subsequent discourse, a phenomenon known as (explanatory) *coherence bias*. The observation of coherence bias indicates that IC is used not only for integrative processing but also for expectation-based processing regarding the upcoming discourse.

Furthermore, IC has increasingly become the focus of second-language (L2) research (Cheng and Almor, 2017; Contemori and Dussias, 2019; Hijikata, 2021; Hosoda, 2023; Kim and Grüter, 2021). Studies have demonstrated that L2 learners use IC bias for online coreference processing during comprehension. Meanwhile, coreference bias may be reduced or absent in L2 learners with limited exposure to IC verbs in L2 or when a disparity occurs in verb usage between L1 and L2 (Cheng and Almor, 2017; Hosoda, 2020). This study focused on IC processing by Japanese learners of English. Although Japanese learners are highly sensitive to the IC coreference bias in Japanese, they often show reduced sensitivity to NP1 coreference bias in English (Hijikata, 2021; Hosoda, 2020). This may be due to differences in how English and Japanese NP1-IC verbs encode causation at the morphemic level. Specifically, NP1 verbs explicitly mark causation with overt morphemes in Japanese (e.g., *Mary ha Bob wo iradata-seru*), whereas causation is integrated into the lexical semantics of the verb without any explicit marker in English (e.g., *Mary annoyed Bob*). Examining Japanese learners' sensitivity to IC bias in English may hence elucidate the cross-linguistic influence of L1 on discourse processes.

Despite these findings, previous L2 studies have predominantly examined coreference bias in contexts with an explicit explanation relation signaled by a causal connective (e.g., *Mary annoyed Bob because*). Considering that IC verbs invoke coherence bias, which leads speakers to expect an explanation in the upcoming discourse, investigating the coreference and coherence biases of IC without an explicit explanation provides a theoretically sound approach to exploring discourse expectations in L2 comprehension.

Various L2 theoretical models suggest differences in expectation-based processing between native speakers and L2 learners (*two-stage model*, Corps et al., 2023; *prediction-byproduction model*, Amos and Pickering, 2020; and *interface hypothesis*, Sorace, 2011). In particular, the *RAGE hypothesis* posits that the differences in L1 and L2 comprehension emerge

directly from L2 learners' reduced ability to generate expectations regarding the upcoming discourse (Grüter et al., 2017; Grüter and Rohde, 2021). Cognitive factors that affect language processing in general (e.g., working memory capacity, attentional resources, task-induced processing, and the quality of L2 lexical representations) may be the cause of the reduced expectation in L2 learners (see Kaan, 2014; Schlenter, 2023 for reviews). Accordingly, L2 learners might struggle more than native speakers in expecting an explanation from IC verbs, potentially showing weaker IC effects when the explanation is not signaled. Given this context, this study aimed to advance our understanding of IC processing in L2 learners by addressing the following unresolved questions:

1. Does IC invoke coreference bias during online L2 processing without the explicit explanation relation?
2. Is the coherence bias of IC, previously observed in native speakers, present in L2 learners?

1.1 Source of the implicit causality bias

Previous IC studies have provided numerous explanations for the mechanisms underlying the IC coreference bias. The major question is whether IC bias is inherent in the linguistic properties of verbs or stems from speakers' general world knowledge. The *verb-based* account argues that IC bias is determined by the verb's semantics and the thematic roles of its arguments; the stimulus (*Mary* in *Mary annoyed Bob*) is more likely to be associated with the event cause than the experiencer (*Bob*; Crinean and Garnham, 2006). The *world-knowledge-based* account argues that IC bias emerges from speakers' inferences regarding the typical causes and effects of the event (Hartshorne et al., 2015). For example, one could infer that *Mary annoyed Bob* is caused by *Mary* based on the typical interpersonal relationship associated with such situations (*Mary sang loudly, which annoyed Bob*).

Although these accounts describe some aspects of the effects of IC, they explain only the coreference bias that occurs in a single sentence. To address this, Bott and Solstad (2014) integrated the verb-based and world-knowledge accounts to develop the *empty-slot theory* (see also Bott and Solstad, 2021; Solstad and Bott, 2022). What distinguishes this theory from the others is that it provides a unified account of the coreference and coherence biases that occur beyond the sentence boundary. Specifically, it posits that the IC bias is fundamentally attributable to the lexical semantics of IC verbs (as in verb-based accounts). Crucially, IC verb semantics triggers the expectation that an explanation will follow in the subsequent discourse because of its explanatory underspecified content, called a *slot*. In *Mary annoyed Bob*, the semantics of the verb *annoyed* has a slot (e.g., *Mary* did something annoying, which explains *Bob's* annoyance with her). Critically, this slot is unfilled because the proper name *Mary* provides no specification. Because speakers follow a general processing strategy to avoid underspecification, this slot causes them to expect the upcoming discourse to provide an explanation (coherence bias). As *Mary* is the main target of this explanation, coreference bias arises, favoring re-mentioning this referent in the produced explanation.

This relationship between the IC verb semantics, explanatory expectation, and coreference bias leads to the following prediction: If speakers expect an explanation from the IC verb and maintain it during comprehension, the expected explanation allows IC to cause coreference bias, even when the explanation is not explicitly signaled. Recall that in the empty-slot theory, coreference bias results from coherence bias toward explanatory expectations. This means that the manifestation of coreference bias in the absence of an explicit explanation requires speakers to expect an explanation. From this perspective, the observation of coreference bias in contexts that lack an explicit explanation serves as a theoretically supported indicator of discourse-level expectations.

1.2 Time course of implicit causality

Another persistent debate in IC literature concerns the time course of IC—when and how IC affects online comprehension processes (Koornneef et al., 2016; Koornneef and Sanders, 2013; Koornneef and Vanberkum, 2006; Pykkönen and Järviö, 2010). A well-known manifestation of online IC effects is the *pronoun-inconsistency effect*, whereby comprehension, operationalized by reading time or eye fixation, is delayed by a pronoun that contradicts the coreference bias of IC. For example, in (2), reading times slow down upon encountering the pronoun *he* because it refers to the NP2 entity (*Bob*), which is inconsistent with the NP1 bias of the verb's IC.

(2) Mary annoyed Bob because he...

Two opposing accounts for this effect have been proposed. The *immediate-focusing* account posits that IC information immediately brings one of the verb's arguments into focus at the expense of the other (McKoon et al., 1993). Hence, this account predicts that an inconsistency effect emerges immediately after encountering a critical pronoun. Oppositely, the *clausal-integration* account posits that IC is used retroactively at the end of the sentence, where interpretations of the two clauses are integrated (Stewart et al., 2000). Accordingly, inconsistency effects are assumed to manifest in the final region of the sentence. Early studies on the time course of IC employed methods such as the probe task or word-by-word self-paced reading. Some studies reported inconsistency effects in the middle of the sentence, a finding consistent with the immediate-focusing account. By contrast, other studies observed these effects at the end of the sentence, thereby supporting the clausal-integration account.

More recently, IC studies have primarily used eye-tracking, a method that captures a wide range of comprehension processes by monitoring various eye movements. Notably, these studies consistently observed the IC effects immediately after or at the introduction of the bias-inconsistent pronoun (Koornneef and Sanders, 2013; Koornneef and Vanberkum, 2006, Experiment 2). The early effects of IC are further supported by research using the visual-world paradigm, which showed that native speakers rapidly fixate on IC-biased referents (Pykkönen and Järviö, 2010).¹

Most relevant to the current purpose, Koornneef et al. (2016) reported early IC effects in contexts that lack an explicit explanation. In their eye-tracking study, stories were presented without a causal connective (e.g., *David apologized to Linda. She, according to the witness, was the one to blame.* [translated from Dutch]). L1-Dutch speakers' reading times were delayed five words after a pronoun that was inconsistent with the IC coreference bias (the underlined *was* in the above example). This finding indicates that native speakers expect an explanation from IC verb semantics during comprehension, which allows IC to bias online coreferential processing (see also Koornneef and Sanders, 2013 for the offline evidence of IC effects without an explicit explanation).

To explain the early IC effects, Koornneef and Sanders (2013) and Koornneef et al. (2016) developed the *incremental integration* account. This model synthesizes components from the immediate-focusing and clausal-integration accounts and recognizes their applicability to different phases of IC processing. Specifically, it suggests that the IC first brings the causally implicated entity into focus (consistent with the immediate-focusing account). Crucially, speakers are assumed to incrementally integrate pronouns with the focused entity on a word-by-word basis. Because the pronoun encounter triggers the use of IC, the IC effects are predicted to emerge immediately after the pronoun, which is consistent with the mid-sentence IC effects frequently observed in the literature.

1.3 Implicit causality in L2 comprehension

Recently, IC coreference bias has become a focal point of investigation in L2 studies (Hijikata, 2021; Hosoda, 2020, 2023; Kim and Grüter, 2021; Wang and Gabriele, 2022). These studies show that IC rapidly influences L2 processes; however, IC effects are often smaller or delayed in L2 comprehension than in L1 comprehension. According to Wang and Gabriele's (2022) self-paced reading experiment, L1-Chinese learners of English showed native-like pronoun-inconsistency effects at or immediately after the critical pronoun, indicating that L2 learners use IC for online coreference processing. However, the story-continuation experiment revealed that learners produced significantly more references inconsistent with the IC bias than native speakers, indicating that learners' sensitivity to the IC coreference bias was weaker than that of native speakers. Kim and Grüter (2021) found online inconsistency effects in intermediate to advanced L1-Korean learners of English using the visual-world paradigm. Specifically, IC affected learners' eye movements 1,000 to 1,500 ms after pronoun offset (approximately two to three words after the pronoun; *Nathan disturbed Owen all the time because he needed help with his homework*). However, these effects were more limited in timing and size than those observed for native English speakers, which persisted from *because* to 1,500 ms after the pronoun offset (*Nathan disturbed Owen all the time because he needed help with his homework*). These findings suggest that, although IC biases online coreferential processing in L1 and L2 comprehension, its effects are slower and weaker in L2 comprehension.

¹ A persistent debate exists regarding whether native speakers proactively use IC to predict the referent online because evidence regarding the

anticipatory looks to the biased referent before the critical pronoun is inconsistent.

To the best of my knowledge, the only study that has reported comparable IC effects between L1 and L2 speakers is that of [Contemori and Dussias \(2019\)](#). Their study showed that highly proficient L1-Spanish and L2-English bilinguals, immersed in L2 from an early age, exhibited native-like IC effects in online self-paced reading and offline story-continuation tasks. Thus, very high L2 proficiency, combined with early L2 exposure, might be necessary for IC processing comparable to that of native speakers.

Notably, as stated in the Introduction, these L2 studies used materials that included an explicit explanation signaled by a causal connective (e.g., *because*). This methodological feature may have contributed to the observed IC effects by making the explanation relation readily available for learners' online processing. The present study directly addressed this issue by comparing contexts with and without an explicit explanation.

The reduced coreference bias among L2 learners can be observed in specific bias directions (NP1 or NP2). For example, [Cheng and Almor's \(2017, 2018\)](#) story-continuation experiments revealed that L1-Chinese learners of English showed a weaker NP2 bias than native English speakers, thereby producing more NP1 references after NP2 verbs. Conversely, the NP1 bias was attenuated in L1-Japanese learners of English. [Hosoda \(2020\)](#) reported that, although Japanese learners generally produced bias-consistent story continuations, the effects of NP1 bias were consistently weaker when the task was conducted in L2-English than L1-Japanese. In a self-paced reading study, [Hosoda \(2023\)](#) found that NP2 bias immediately influenced online coreferential processing, whereas the effects of NP1 bias did not emerge until the sentence-final region (see [Hijikata, 2021](#) for another evidence of the difficulty experienced by Japanese learners with NP1 bias).

A possible explanation for these observations is related to cross-linguistic differences. The weaker NP2 bias among L1-Chinese learners can be attributed to the fact that Chinese has fewer NP2 verbs than English ([Cheng and Almor, 2017](#)). The greater prevalence of NP1 verbs in Chinese may cross-linguistically lead L1-Chinese learners to favor NP1 references.

For the weaker NP1 bias among L1-Japanese learners, a possible explanation is the morphemic differences between English and Japanese verbs. As mentioned in the Introduction, Japanese NP1 verbs have overt causative morphemes (e.g., *-seru* or *-saseru*) that explicitly mark causation (e.g., *konran-saseru* [confuse] and *shitsubou-saseru* [disappoint]). Conversely, English NP1 verbs do not have overt morphemes with causation being implicitly conveyed through the verb semantics. This difference negatively influences the application of the lexical knowledge of NP1 verbs to discourse processes ([Hosoda, 2020, 2023](#)). Specifically, even when L1-Japanese learners knew the meaning of English NP1 verbs (as confirmed by the translation task showing that English verbs were correctly translated into their L1 counterparts), they exhibited reduced NP1 bias in discourse processes, such as referential resolution and the prediction of upcoming referents ([Hijikata, 2021; Hosoda, 2022](#)). This observation differs from that of NP2 verbs, for which causation is not explicitly marked morphemically in either Japanese or English (e.g., *konomu* [like] and *bassuru* [punish]). Consequently, Japanese learners can use NP2 bias in English to a similar extent as native English speakers and as they do in their own L1 ([Hosoda, 2020, 2022](#)). Based on these existing

findings, this study predicted that L1-Japanese learners of English would show reduced effects of coherence and coreference biases for NP1 verbs compared with native speakers.

1.4 Focus of the present study

Existing L2 studies have focused solely on coreference bias in contexts with an explicitly signaled explanation (e.g., *Mary annoyed Bob because he...*). According to the empty-slot theory, the manifestation of coreference bias without an explicit explanation requires speakers to expect an explanation from IC verbs ([Bott and Solstad, 2014, 2021; Solstad and Bott, 2022](#)). Considering L2 accounts positing that L2 learners are more limited in their ability to generate expectations than native speakers ([Grüter and Rohde, 2021; Schlenter, 2023](#)), they may struggle to make an explanatory expectation from IC verbs (coherence bias). In such a scenario, IC will not invoke coreference bias during L2 processing, specifically when the explanation is not explicitly signaled. This study investigated this possibility.

2 Experiment 1

Experiment 1 addressed the following research question (RQ): Is the occurrence of early IC effects during L2 processes, reported in existing L2 literature, extended to a context in which the explanation relation is not explicitly signaled?

To address this RQ, a *full-stop* condition was set, whereby the stimuli were presented without a causal connective (*Mary annoyed Bob. He...*). The experiment used eye-tracking to analyze the time course of IC.

Notably, this study used eye-tracking in reading rather than listening (as in the visual-world paradigm). This is because the listening mode is used much less frequently than reading with the L2 learner population tested in this study, primarily because these learners rarely use L2 in everyday communication. This lower familiarity with listening would have confounded the results. It should also be noted that Experiment 1 did not include native speakers in the control group. It is extremely difficult to recruit a sufficient sample of native English speakers for an in-person eye-tracking experiment who match the L2 participants in age, socioeconomic status, and educational backgrounds in environments where English is rarely used daily. This is a limitation of this study that will be revisited in the General Discussion.

2.1 Method

2.1.1 Participants

Forty-two L1-Japanese university students participated in the study (24 females; $M_{age} = 20.02$; age range: 18–24). They had been learning English in Japan for six or more years. None of them had studied abroad in an English-speaking country. All lived in environments where English is used as a foreign language, meaning that it is not used for communication in daily life. All participants had normal or corrected-to-normal visual acuity. Their

L2 proficiency was estimated to be at the CEFR A2–B1 levels or 28–80 on the TOEFL iBT test (Papageorgiou et al., 2015), based on their scores on the TOEIC IP test ($M = 474.62$, $SD = 127.37$). This was conducted as a placement test for the university English course 1–2 months before the experiment. Data from two participants were excluded owing to major losses in their eye-tracking data.

2.1.2 Materials

2.1.2.1 Implicit causality verbs

This study used 24 IC verbs (12 NP1 and 12 NP2 verbs), originally from the 300 English verbs identified by Ferstl et al. (2011) in a norming study on IC bias in English verbs.² The coreference bias of these verbs was piloted with separate Japanese university students ($N = 20$), who completed the story-continuation task in Japanese and English. The bias was determined using the rate of NP1 references (number of NP1 references/[number of NP1 + NP2 references]). Verbs eliciting 70% or more NP1 references in English and Japanese were categorized as NP1 verbs, whereas those eliciting 30% or fewer NP1 references in both languages were categorized as NP2 verbs. Accordingly, the experimental IC verbs were confirmed to cause the coreference bias of similar strength in the same referential directions in Japanese and English. This suggests that the bias was similar in these languages; hence, cross-linguistic differences in the bias should not interfere with learners' use of IC in L2 English. Furthermore, verbs eliciting fewer than 65% references to NP1 or NP2 referents in Japanese and English were categorized as non-IC verbs, which were used in Experiment 2.³

All verbs were within Rank 3,000 or below on the *New Japan Association of College English Teachers (JACET) List of 8,000 Basic Words* (JACET Basic Word Revision Committee, 2016), which lists the 8,000 English words that Japanese students learn between elementary school and university based on frequency. English words of this rank have been introduced in junior high or high schools in Japan. Therefore, the IC verbs used in this study were assumed to be familiar to the participants.

Notably, this study did not differentiate between the types of IC verbs (action, psychological, and state verbs) for three main reasons. First, the inclusion of a diverse set of IC verbs allowed the use of verbs whose strength and direction of bias were shared in English and Japanese. Second, this study aimed to maximize the number of items to ensure adequate statistical power. Third, this study's primary interest was to examine IC effects in

TABLE 1 Example NP1 stimuli.

Sentence construction	Example
The first and second sentences.	Steve and Hanako were in a meeting about their company's new product. They did not agree on the sales plan and had an argument.
Target sentence (Because-consistent condition)	Steve hurt Hanako because he had always by nature been an aggressive person.
(Because-inconsistent condition)	Hanako hurt Steve because he had always by nature been a sensitive person.
(Full-stop-consistent condition)	Steve hurt Hanako. He had always by nature been an aggressive person.
(Full-stop-inconsistent condition)	Hanako hurt Steve. He had always by nature been a sensitive person.
(Comprehension question)	Did Steve and Hanako have an argument?

a heterogeneous set of verbs, rather than effects attributable to specific verb types.

2.1.2.2 Story stimuli

The IC verbs were used in three-sentence story stimuli constructed following L1 studies on the time course of IC (Koornneef et al., 2016). All stimuli comprised English vocabulary and grammatical structures introduced in junior high schools in Japan, thereby ensuring that the participants did not struggle excessively with lexical or grammatical processing.

Table 1 shows an example of the stimulus. The first sentence sketched a story and introduced two characters (one female and one male) with their names.⁴ The second sentence continued the story while mentioning the characters with *they* to keep them similarly salient.

The third sentence was the target sentence in the format *NP1 IC verb-ed NP2 because he...* The presence of the causal connective was manipulated by replacing *because* with a full stop.

The target sentence established consistent and inconsistent conditions based on the consistency of the pronoun *he* with the coreference bias of IC. Pronoun consistency was manipulated by switching characters in the NP1 and NP2 positions instead of changing *he* to *she* to avoid potential confounding due to word differences. At least five words after the pronoun were held in common across the conditions to accommodate spillover IC effects. After this common region, the consistent and inconsistent stories diverged, resulting in information that made the overall story coherent.

Each stimulus was accompanied by a comprehension question to encourage the participants to read carefully. The question targeted explicit information in the story that was irrelevant to the interpretation of the pronoun. Half of the questions were correctly answered with "Yes" and the rest with "No." Four sets were constructed, crossing 2 (pronoun consistency: consistent, inconsistent) \times 2 (connective: because, full stop) conditions. The assignment of stimuli to the conditions was counterbalanced across the sets.

4 Japanese university students from the same population ($N = 32$) confirmed that the genders of the names were easy to identify.

2 All the materials, datasets, analysis codes, and statistical results in this study are available at https://osf.io/3qnz6/?view_only=a4e46c7da585467e994632f4f57aee4b.

3 Non-IC verbs included stimulus–experiencer verbs (e.g., *surprised* and *encouraged*) and an experiencer–stimulus verb (*forgot*), which could be classified as NP1 and NP2 verbs, respectively. However, norming data from Ferstl et al. (2011) and this study showed that the percentage of NP1 references for these verbs ranged from 42% to 55%, indicating that these verbs do not induce a bias toward NP1 or NP2 references. Based on this evidence, this study categorized these verbs as non-IC verbs. I appreciate the reviewer's comment on this issue.

2.1.2.3 Translation task

A translation task was conducted to confirm whether the participants had semantic knowledge of IC verbs. This task was necessary because IC effects would not emerge unless the participants knew the meaning of the IC verbs. The participants were presented with a matrix clause of the 24 target sentences (e.g., “Mary annoyed Bob.”) and asked to translate it into Japanese.

2.1.3 Apparatus and procedure

The eye movements were recorded using the Tobii Pro Spectrum (300 Hz). The experiment was performed using the Tobii Pro Lab. A chin rest was used to minimize head movement. The participants were tested individually in a silent room. The experimenter explained the procedure to the participants and obtained their written informed consent. Participants sat approximately 60 cm from the 23.8-inch monitor and received written and oral instructions. Subsequently, they read two practice stories (Courier New 20-point font with double spacing) presented in their entirety, each accompanied by a comprehension question.

The eye-tracker was calibrated using a standard nine-point grid. When the error in any gaze position exceeded 1.0°, the eye-tracker was recalibrated until the average error was smaller than 0.5°. The participants read the experimental stimuli, and their eye movements were recorded. The order in which the stimuli were presented was randomized for each participant. After each stimulus, they answered a comprehension question by fixating on the “Yes” or “No” mark displayed on the monitor. Each session consisted of three blocks. The participants had a 5-min break between blocks, and the eye-tracker was recalibrated. After all stimuli were presented, the translation task was administered.

2.1.4 Coding and data treatment

2.1.4.1 Translation task

Two raters coded the responses as “correct” or “incorrect” by matching them with the intended meaning of the IC verb. In particular, the raters carefully checked whether the NP1 verbs were translated with a transitive causative meaning because the differences between Japanese and English are evident in this usage. Other meanings or usages (e.g., passive interpretations) were coded as “incorrect.” Conversely, the raters were flexible regarding wording variations as long as the translations captured the gist of the intended meaning (e.g., for *disappointed*, *shitsubou-saseta* and *gakkari-saseta* were coded as “correct”). If any ambiguity was noted, the response was coded as “incorrect.” The inter-rater agreement rate was 96%. All disagreements were resolved through discussion.

2.1.4.2 Eye-tracking measures

This study conducted a region-by-region, rather than word-by-word, analysis. This is because function words are often skipped and fixated on 35% of the time, whereas content words receive fixations 85% of the time (Carpenter and Just, 1983).

Three regions were set, each comprising two words (“Steve hurt Hanako because/he had/always by/nature been/an aggressive person.”). Region 1 included the critical pronoun and one word after it (“.../he had/always by/nature been/...”). Region 2 consisted of two and three words after the pronoun (“.../he had/always by/nature been/...”). Region 3 included four and five

words after the pronoun (“.../he had/always by/nature been/...”). Four eye-tracking measures were computed to capture the IC effects in the initial and late processes. *First-fixation* and *first-gaze durations* were computed as measures of the initial processes (e.g., initial access to the word’s semantic information). First-fixation duration was the duration of the first fixation on a region. First-gaze duration was the sum of the fixation durations on one region before the participant moved forward or returned to another region.

Regression-path and *right-bound durations* were computed to measure late processes (e.g., processing difficulty associated with the integration of a target word into the previous context). The regression-path duration summed all fixation durations when the participant fixated on a region until they moved on to the following region. This means that the measure included all fixation durations of looking back at the prior region of the text after the target region had been fixated on.⁵ The right-bound duration was the sum of all fixation durations on a region before the participant left the region in the forward direction. Right-bound duration differed from regression-path duration as it did not include fixation durations while looking back at prior regions.

The data from these measures were converted into per-syllable measures to account for differences in word length. Subsequently, the data were log-transformed to correct for right skewness. Among the trials, 1% were excluded because of tracker losses or eye blinks. The data from incorrectly translated items (16%) were also excluded. Fixation durations that were more than 2 SD from the mean in any experimental condition were treated as missing data (< 10% for all measures).

2.1.5 Statistical analysis

Linear mixed-effects models were used to analyze eye-tracking measures. The fixed effects were bias direction (NP1 and NP2), pronoun consistency (consistent and inconsistent), connective (because and full stop), all sum-coded, and Direction × Consistency × Connective interaction. For random effects, the best-fitting and most parsimonious structures were selected through a backward model comparison. Specifically, a random slope was successively removed from the maximal model (with random intercepts of participants and items as well as the by-participant random slopes of consistency, bias direction, connective, and Direction × Consistency × Connective interaction), unless its removal significantly decreased the model fit. This was done by referring to the AIC (Akaike Information Criterion) and *p* values of the model comparisons using the *anova* function. The selected models had random intercepts of participants and items and a by-item random slope of consistency (formula: eye-tracking measure ~ direction * consistency * connective + [1 | participant] + [1 + consistency | item]).

⁵ Because regression-path duration is the most inclusive eye-tracking measure, some studies deemed it as an early measure (Clifton et al., 2007; Cunnings and Sturt, 2018). This study treated it as a late measure because it indexes processing difficulty, which occurs in later stages of comprehension.

2.2 Results

2.2.1 Translation task

The participants produced more erroneous translations for NP1 verbs (error rates = 22%) than for NP2 verbs (error rates = 10%; $\beta = 1.92$, $SE = 0.92$, $z = 2.07$, $p = 0.039$), indicating that they had less semantic knowledge of NP1 than NP2 verbs. In particular, in 65% of the erroneous trials, the participants interpreted NP1 verbs in a passive sense (e.g., *Mary ha Bob ni shitsubou-saserareta* ["Mary was disappointed by Bob"] for *Mary disappointed Bob*), rather than in a transitive causative sense. The implications of this finding are presented in the General Discussion.

2.2.2 Eye-tracking measures

The scores for the comprehension questions were generally high ($M = 90.44$, $SD = 12.98$), confirming that the participants carefully read the stimuli (no participant showed a correct percentage of 70% or less). Table 2 presents the descriptive statistics of the eye-tracking measures in the *because* and full-stop conditions. Figure 1 illustrates the mean duration of the eye-tracking measures.

This section focuses on the results concerning the pronoun-inconsistency effects that are relevant to the RQs.⁶ None of the eye-tracking measures showed significant effects associated with pronoun consistency in Region 1 ($p > 0.050$). However, significant Direction \times Consistency \times Connective interactions emerged for the regression-path duration in Regions 2 ($\beta = 0.04$, $SE = 0.02$, $t = 2.07$, $p = 0.039$) and 3 ($\beta = 0.04$, $SE = 0.02$, $t = 2.07$, $p = 0.039$). Follow-up tests revealed that the regression-path duration was delayed by the bias-inconsistent pronoun when the NP2 verb was followed by *because* (Region 2: $\beta = -0.25$, $SE = 0.11$, $t = -2.33$, $p = 0.026$; Region 3: $\beta = -0.26$, $SE = 0.11$, $t = -2.36$, $p = 0.025$). None of the eye-tracking measures were affected by pronoun consistency when *because* was absent or the verb had an NP1 bias ($p > 0.050$).⁷

2.3 Discussion

The analyses revealed that the connective condition and IC bias directions interactively determined the occurrence of early IC effects during L2 processes. Specifically, NP2 verbs elicited significant inconsistency effects for the regression-path duration in Regions 2 and 3. Notably, these effects were observed only when the explanation relation was signaled by *because*. The measures showed no inconsistency effect in the full-stop condition or after NP1 verbs.

The null effect of NP1 bias matches the evidence from prior research, showing that L1-Japanese learners less effectively use NP1

bias than NP2 bias (Hijikata, 2021; Hosoda, 2023). However, prior L2 research on the time course of IC has never contrasted the connective and full-stop contexts. The present findings provide initial evidence that an upcoming explanation must be explicitly signaled for IC to invoke a mid-sentence influence on online L2 processing.

The empty-slot theory posits that the occurrence of coreference bias in the absence of an explicit explanation requires speakers to expect an explanation from IC verb semantics. This account explains the present finding that the L2 participants struggled to expect explanations from IC verbs in the full-stop condition. This was addressed in Experiment 2.

3 Experiment 2

The main RQ of Experiment 2 was as follows: Are L2 learners' expectations of an explanation from IC verbs lower than those of native speakers? To address this, a story-continuation task was used, in which the participants created continuations of the story sentence, including an IC (*Mary annoyed Bob*) or non-IC (*Mary saw Bob*) verb. As mentioned in the Introduction, native speakers favor explanation continuations for IC-verb stories relative to non-IC verb stories. Experiment 2 compared this coherence bias between L2 learners and native speakers. Based on the results of Experiment 1, this study predicted that L2 learners would show a lower coherence bias than native speakers. Because this study was specifically interested in whether an explanation was expected from IC verbs, the *because* condition in which an upcoming explanation is signaled was omitted (only the full-stop condition was set).

In addition, Experiment 2 asked the following RQ: When the explanation relation is not signaled, is L2 learners' reference to the IC-biased referent reduced compared with that of native speakers? Weakened coherence bias should lead to a reduction in coreference bias considering the assumption that coreference bias results from coherence bias (Bott and Solstad, 2014, 2021; Solstad and Bott, 2022). To investigate this, the L1 and L2 groups were compared in terms of consistency between the entities mentioned in the continuation and IC coreference bias (henceforth, *reference-IC bias consistency*).

It must be noted that the story-continuation task is an offline measure; continuations are produced after a sentence has been comprehended. Therefore, the data do not provide direct information on whether the participants expect an explanation during comprehension. Despite this limitation, this study used the story-continuation task because it is the most widely adopted method to test the expectation of coherence relations in L1 (Bott and Solstad, 2021; Kehler et al., 2008) and L2 (Grüter et al., 2017) studies. Employing this task was necessary to ensure the comparability of the results with those in the literature.

3.1 Method

3.1.1 Participants

In the L1 group, 56 monolingual English speakers were recruited through Amazon Mechanical Turk (27 females; M_{age}

⁶ All statistical results are provided in the Supplementary material.

⁷ L2 proficiency test scores (centered) were added to the maximal converging models to examine the potential effects of the participants' L2 proficiency on the results. The model comparisons revealed that the addition of L2 proficiency did not significantly contribute to the model fit, meaning that the difference in the participants' L2 proficiency did not influence the results.

TABLE 2 Eye-tracking measures (in ms per syllable) as a function of IC-Bias direction, pronoun consistency, and region.

Measure and condition		Because						Full-stop					
		Region 1		Region 2		Region 3		Region 1		Region 2		Region 3	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
First-fixating duration													
NP1	Con	156	76	151	66	212	92	147	75	154	56	201	88
	Inc	148	74	171	74	202	80	173	104	155	70	201	89
NP2	Con	176	72	201	83	166	72	180	80	186	75	171	81
	Inc	175	80	196	76	171	71	173	91	190	81	172	78
First-gaze duration													
NP1	Con	205	87	199	83	335	168	202	107	199	90	315	151
	Inc	212	108	214	102	303	144	225	125	196	84	294	125
NP2	Con	217	95	249	107	217	85	205	99	221	80	222	98
	Inc	209	98	247	92	231	113	201	106	221	91	228	95
Regression-path duration													
NP1	Con	286	183	284	188	397	213	336	214	300	229	405	244
	Inc	280	167	258	140	391	251	350	223	313	252	381	237
NP2	Con	285	175	287	138	257	127	285	164	304	195	335	285
	Inc	243	137	406	284	335	199	294	189	279	121	271	171
Right-bound duration													
NP1	Con	233	102	219	89	369	177	260	124	235	110	364	184
	Inc	258	135	236	113	344	171	278	136	240	111	324	148
NP2	Con	246	109	270	115	243	103	242	113	244	95	242	113
	Inc	221	100	297	127	255	106	244	118	263	110	224	91

NP1, NP1 context; NP2, NP2 context; Con, consistent condition; Inc, inconsistent condition.

= 21.80; age range: 19–42). All participants had completed a university-level education. Ten participants who failed to appropriately complete an instructional manipulation check that aimed to ensure that they had read the instructions carefully were excluded. The analysis excluded another six participants because they did not provide an appropriate answer to “catch” items, which aimed to ensure that the participants carefully comprehended the stimuli. These procedures were necessary because participants in the online survey often “satisfice,” meaning that they try to complete the task without concentrating. The remaining 40 participants were included in the final analysis. They were paid US \$7.00 for their participation.

In the L2 group, 40 Japanese university students participated (24 females; M_{age} = 18.71; age range: 18–20). None of them had been tested in Experiment 1 or had study-abroad experience in an English-speaking country. Their L2 proficiency was estimated to be at the CEFR A2–B1 levels or 28–80 on the TOEFL iBT test, based on their scores on the TOEIC IP test (M = 459.38, SD = 101.69), which was administered 3 weeks before the experiment. These scores suggested that the L2 participants’ proficiency was almost similar to that in Experiment 1.

3.1.2 Materials

3.1.2.1 Verbs and the story-continuation task

In addition to the 24 IC verbs used in Experiment 1, 24 non-IC verbs from the norming study reported in Experiment 1 were used (see Method in Experiment 1). These verbs generated fewer than 65% of the references to the NP1 or NP2 referents in Japanese and English. This confirmed that non-IC verbs in English and Japanese do not cause a coreference bias toward the NP1 or NP2 direction.

The context stimuli were constructed using the 48 experimental verbs in the format of *NP1 verb-ed NP2*. Examples of the stimuli are presented in Table 3. The stimuli had two familiar English names for different genders. Half of the stimuli had a female NP1, whereas the other half had a male NP1.

In this task, the participants were instructed to write a continuation of the context stimulus that naturally came to mind in English. They were asked not to worry about grammar or spelling mistakes while avoiding humor. The exact instruction (presented in English and Japanese to the L1 and L2 participants, respectively) was as follows: “You will create continuations of stories in English. Please read the English sentences carefully to understand the story. Then, write a continuation of the story that comes to mind.”

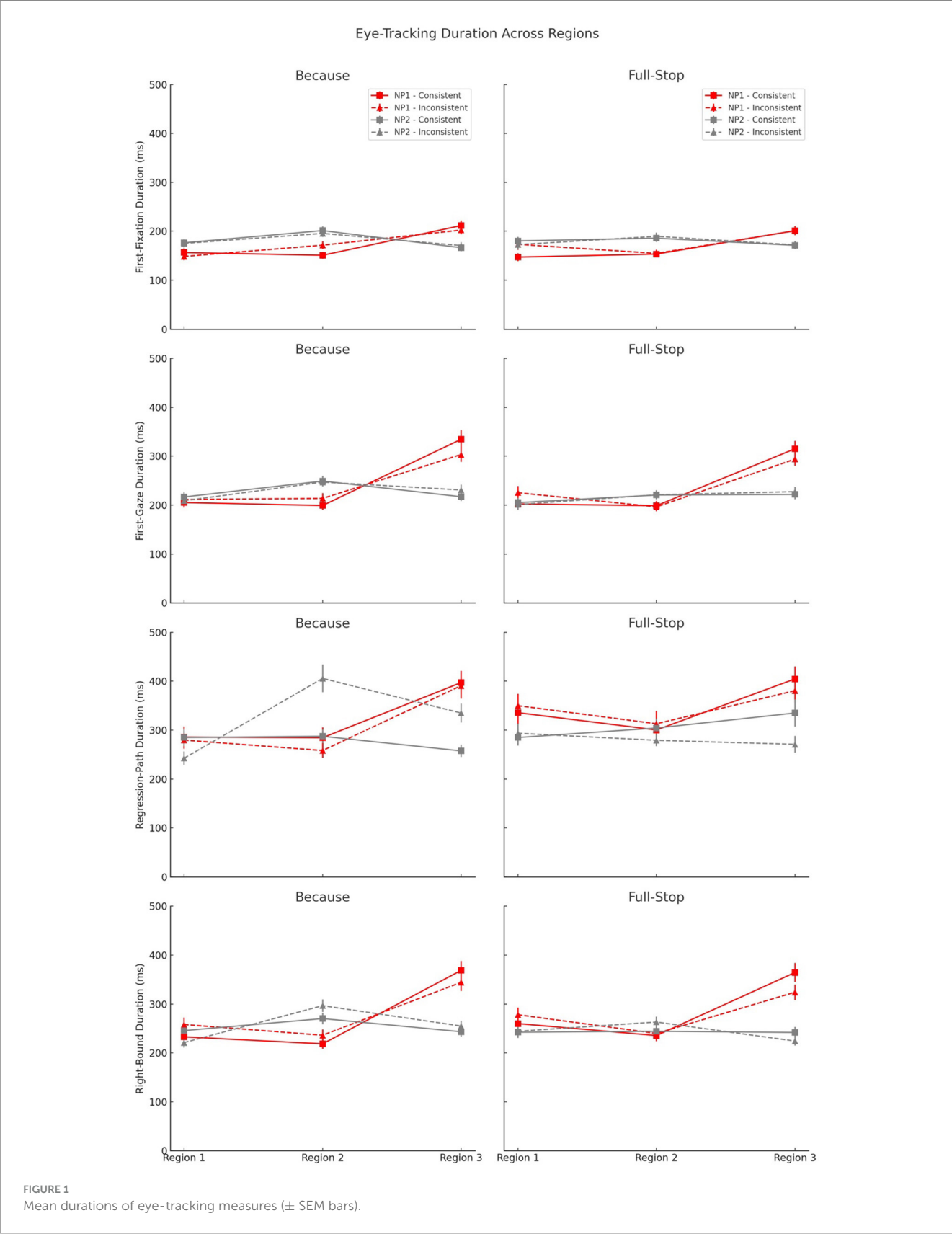


TABLE 3 Example story stimuli in experiment 2.

Condition	Stimulus
NP1 Bias	Mary annoyed Bob.
NP2 Bias	Mary punished Bob.
Non IC	Mary saw Bob.

Two material sets were constructed with the 48 experimental items and 48 fillers used in another experiment for the interpretation of the relative clause. The gender of the NPs was counterbalanced across the sets. The order of presentation of the items was randomized for each participant.

3.1.2.2 Translation task

The translation task was administered to the L2 participants in the same format as in Experiment 1. The items included 24 IC verbs and 24 non-IC verbs.

3.1.3 Procedure

The L1 participants were recruited through Amazon Mechanical Turk and directed to the survey website Qualtrics. They completed a demographic questionnaire and an instructional manipulation check. Subsequently, a story-continuation task was performed. One to four participants were tested simultaneously in the L2 group. The participants completed the story-continuation task in English, and the translation task was administered.

3.1.4 Coding and data treatment

3.1.4.1 Translation task

Two raters coded the responses (including the author) as either “correct” or “incorrect,” using the same procedure as in Experiment 1. The inter-rater agreement rate was 92%. All disagreements were resolved through discussion.

3.1.4.2 Story-continuation task

Prior to coding, continuations were excluded if they did not make sense (<1% of the data) or if L2 participant provided an incorrect translation of the corresponding item (14% of the L2 data).

3.1.4.3 Coherence bias

Two raters coded the continuations (including the author) for whether they explained the cause of the event in the stimulus (“explanation”) or not (“non-explanation”). The raters used two tests based on previous story-continuation studies on coherence relations to identify explanation continuation (Grüter et al., 2017; Kehler et al., 2008). First, the raters considered whether the continuation answered the *why* question of the stimulus event. Second, they checked whether the causal connective *because* could felicitously relate stimulus and continuation without changing the gist of the story. The continuation was coded as “non-explanation” if it did not meet either of these criteria.

For example, the continuation *Bob broke the rule* for the stimulus *Mary punished Bob* is coded as “explanation” because it answers the *why* question (*Why did Mary punish Bob?*) and can be

felicitously related to the stimulus by *because* (*Mary punished Bob because Bob broke the rule*). Conversely, *Mary worked for the same company as Bob* would be coded as “non-explanation” because it does not answer the *why* question but describes details of the event (i.e., an elaboration relation). Continuations were coded as “non-explanation” when they were interpreted as the result (e.g., *Bob felt sorry*), temporal succession (e.g., *Mary taught him what to do*), or an unexpected outcome of the event (e.g., *But Bob continued to make trouble*).

To identify the coherence relation, the raters drew on transition words (e.g., connectives and adjectives) if they were present in the continuation (e.g., *Mary punished Bob. Because he broke the rule*). However, the raters carefully checked whether the story continued reasonably using the relations denoted by such words. For example, even though *because* was used, some continuations can more felicitously be interpreted as elaborations rather than explanations (e.g., *Mary punished Bob. Because Bob was her coworker.*). In this case, the continuation was coded as “non-explanation.” When two or more types of coherence relations could be inferred, the continuation was coded as “unclear” and excluded from the analysis (5% of the data). Data were also excluded when the continuations did not make sense (<1% of the data). The inter-coder agreement was 90%. All disagreements were resolved through discussion.

3.1.4.4 Coreference bias

Two raters (including the author) annotated the continuations produced in the IC (NP1 and NP2) contexts for which NP1 or NP2 entity was mentioned. When the pronoun was used, the raters used its gender as a cue. However, the raters carefully confirmed whether the referent referred to by the pronoun made sense considering the rest of the story. Continuations were coded as “unclear” when they had any ambiguity or neither the NP1 nor NP2 entity was mentioned (e.g., mentioning an entity that is not present in the stimuli or mentioning both NPs with a conjoined noun phrase [e.g., *Mary and Bob*] or a plural pronoun [*they*]). “Unclear” continuations (7% of the data) were excluded from the analysis. The inter-rater agreement rate was 92%. All disagreements were resolved through discussion.

3.1.5 Statistical analysis

3.1.5.1 Coherence bias

The continuation data (explanation vs. non-explanation) were submitted to a mixed-effects logistic regression model to test coherence bias, which caused more explanation continuations in the IC (NP1 and NP2) contexts than in the non-IC context. The fixed effects were bias direction (NP1, NP2, and non-IC [reference level]), group (L1 and L2), and the Direction × Group interaction. The bias direction was dummy-coded with the other sum-coded. This is because the difference between non-IC verbs, serving as the reference level, and IC (NP1 and NP2) verbs (i.e., simple effects) was relevant to the RQs, rather than the main effect of bias direction. The random-effects structure was determined through a backward model comparison using the same procedure as in Experiment 1. The selected model had random intercepts of participants and items as well as the by-participant random slope of the bias direction and by-item random slope of the group (formula:

TABLE 4 Means and standard deviations of rates of explanation continuations.

Group	NP1 context		NP2 context		Non-IC context	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
L2	0.24	0.43	0.52	0.50	0.24	0.43
L1	0.57	0.50	0.72	0.45	0.23	0.42

continuation \sim direction * group + [1 + direction | participant] + [1 + group | item]).

3.1.5.2 Coreference bias

The reference-IC bias consistency was computed by matching the referent mentioned in the continuation with the IC coreference bias (consistent vs. inconsistent). This consistency was analyzed using a mixed-effects logistic regression model with the fixed effects of bias direction (NP1 and NP2) and group (L1 and L2), both sum-coded, and the Direction \times Group interaction. The random-effects structure was determined using a backward model comparison, as in other analyses. The selected model had random intercepts of participants and items as well as the by-participant random slope of the bias direction and by-item random slope of the group (formula: consistency \sim direction * group + [1 + direction | participant] + [1 + group | item]).

3.2 Results

3.2.1 Translation task

L2 participants produced more erroneous translations for NP1 verbs (error rates = 27%) than NP2 (error rates = 9%; $\beta = -2.14$, $SE = 0.95$, $z = -2.56$, $p = 0.024$) and non-IC (error rates = 10%; $\beta = -1.77$, $SE = 0.80$, $z = -2.21$, $p = 0.027$) verbs. Similar to Experiment 1, L2 participants often interpreted NP1 verbs with a passive meaning (61% of the erroneous trials). The error rates were not significantly different between the NP2 and non-IC verbs ($\beta = -0.40$, $SE = 0.88$, $z = -0.45$, $p = 0.651$).

3.2.2 Coherence bias

The descriptive statistics of the explanation rates are presented in Table 4.⁸ As illustrated in Figure 2, significant Direction \times Group interactions emerged in NP1 and NP2 contexts (NP1: $\beta = -0.93$, $SE = 0.18$, $z = -5.15$, $p < 0.001$; NP2: $\beta = -0.70$, $SE = 0.17$, $z = -3.99$, $p < 0.001$).

Follow-up tests revealed that L1 participants showed coherence bias (more explanations in the IC context than the non-IC context) in NP1 and NP2 contexts (NP1: $\beta = 1.97$, $SE = 0.43$, $z = 4.55$, $p < 0.001$; NP2: $\beta = 2.88$, $SE = 0.44$, $z = 6.53$, $p < 0.001$). L2 participants showed coherence bias only in the NP2 context ($\beta = 1.49$, $SE = 0.33$, $z = 4.56$, $p < 0.001$) with no significant difference between the NP1 and non-IC contexts ($\beta = 0.13$, $SE = 0.35$, $z = 0.39$, $p = 0.700$).

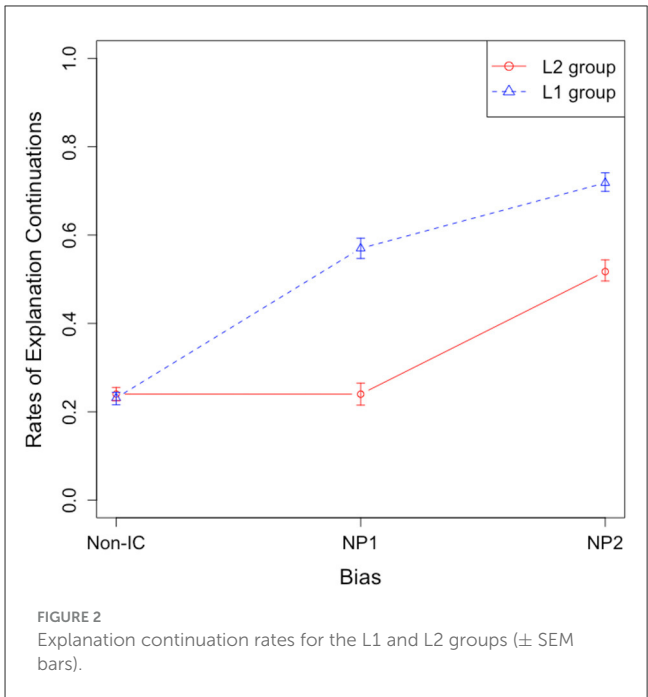


TABLE 5 Means and standard deviations of reference-IC bias consistency rates.

Group	NP1 context		NP2 context	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
L2	0.41	0.49	0.79	0.41
L1	0.62	0.49	0.80	0.40

The model showed no significant L1–L2 group difference in the non-IC context ($\beta = 0.11$, $SE = 0.13$, $z = 0.90$, $p = 0.371$). Conversely, NP1 and NP2 contexts elicited more explanations from L1 than L2 participants (NP1: $\beta = -0.82$, $SE = 0.12$, $z = -6.73$, $p < 0.001$; NP2: $\beta = -0.53$, $SE = 0.11$, $z = -4.96$, $p < 0.001$).

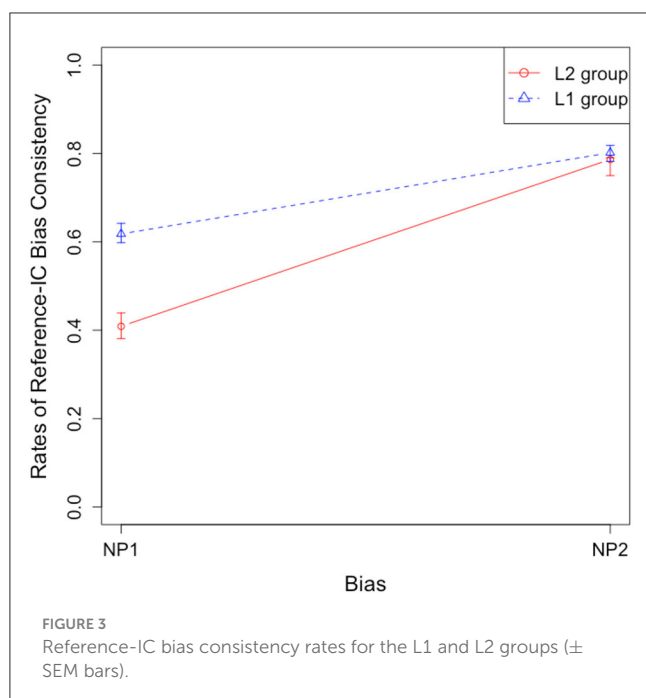
3.2.3 Coreference bias

Table 5 presents the descriptive statistics of reference-IC bias consistency rates. Figure 3 illustrates the results. Overall, the consistency was lower among L2 participants than L1 participants and in the NP1 context than the NP2 context, as indicated by the significant main effects of group ($\beta = -0.29$, $SE = 0.11$, $z = -2.52$, $p = 0.012$) and bias direction ($\beta = -0.82$, $SE = 0.17$, $z = -4.93$, $p < 0.001$), respectively.

Additionally, a significant Direction \times Group interaction was observed ($\beta = -0.25$, $SE = 0.12$, $z = -2.02$, $p = 0.044$). The NP1 context elicited higher consistency from the L1 group than the L2 group ($\beta = -0.51$, $SE = 0.14$, $z = -3.53$, $p < 0.001$). Conversely, the NP2 context showed no significant L1–L2 group difference ($\beta = -0.07$, $SE = 0.16$, $z = -0.45$, $p = 0.651$).⁹

⁸ All statistical results are provided in the Supplementary material.

⁹ L2 proficiency test scores (centered) were added to the maximal converging models of L2 data. The addition of L2 proficiency did not significantly improve the model fit in either coherence or coreference bias.



3.3 Discussion

The results confirmed that explanatory expectations from IC verbs were more reduced in L2 learners than in native speakers. Unlike L1 participants, who showed higher explanation rates in both NP1 and NP2 contexts than in the non-IC context,¹⁰ L2 participants showed this coherence bias only in the NP2 context. Moreover, the explanation rates were consistently lower in L2 than L1 participants. These results indicated that IC verbs invoked coherence bias in L2 participants only in the limited condition (after the NP2 verb) and to a limited extent compared with L1 participants.

Experiment 2 also found that L2 participants were limited in terms of coreference bias, as reflected by their lower reference-IC bias consistency than that of L1 participants in the NP1 context. Because L2 participants failed to expect an explanation from the NP1 verbs, the explanation relation was not sufficiently operative to allow them to refer to the causally implicated referent, thereby creating the intergroup differences specifically in the NP1 context.

Finally, the NP2 context showed no significant L1–L2 differences in terms of coreference bias. This differs from the results for coherence bias, where L2 participants were limited relative to L1 participants in both NP1 and NP2 contexts. It might be that its recency to the end of the stimulus sentence made the NP2 entity more salient than the NP1 entity in the participants' mental models when the continuation was created.¹¹ This might have increased the

NP2 references to mask possible intergroup differences in the NP2 context. This idea was supported by the fact that the main effect of bias direction was significant (as reported in the first paragraph of the coreference bias results), indicating that the participants generally favored NP2 references over NP1 references.¹¹

4 General discussion and conclusion

This study investigated how verbs' IC affects L2 learners' coreference and coherence processing. Experiment 1 revealed that NP2 bias affected learners' online coreference processing. Furthermore, this effect was observed only when an upcoming explanation was explicitly signaled. Subsequently, Experiment 2 showed that L2 learners were limited compared with native speakers in terms of coherence bias, whereby speakers expect explanations from IC verbs. Additionally, NP1 verbs failed to cause either coreference or coherence biases in learners, which matched the absence of the NP1 bias effect on online L2 processing, as shown in Experiment 1.

Prior to discussing the findings, we must be careful regarding the comparison of L1 and L2 comprehension in terms of the time course of IC because Experiment 1 did not contrast L2 learners with native speakers. However, Experiment 2 directly compared L1 and L2 participants and found that the IC effects in the L2 group were significantly more limited than those in the L1 group. Additionally, the difficulty experienced by L2 learners in the offline task was unlikely to be alleviated in the online task, considering that the online task poses higher cognitive demands: When performing an online task, the incoming linguistic information must be continuously processed within a limited time constraint, whereas linguistic knowledge can be retroactively or strategically used in an offline task. Considering these observations, it is reasonable to presume that using IC for online processing is more difficult for L2 learners than native speakers. From this perspective, the following sections discuss the findings in terms of the source and time course of IC.

4.1 Source of implicit causality bias in L2 comprehension

The empty-slot theory posits that the coreference bias of IC emerges as an epiphenomenon of coherence bias toward explanatory expectations (Bott and Solstad, 2014, 2021; Solstad and Bott, 2022). Thus, the finding of Experiment 1 that IC failed to cause coreference bias in the absence of an explicitly signaled explanation suggests that L2 learners were essentially limited in expecting an explanation from IC verbs. This idea was supported by Experiment 2, which showed that the L2 group exhibited a

This means that the difference in the participants' L2 proficiency did not influence the results.

¹⁰ One may argue that the explanation rates for the L1 group were not high (70% at best). However, these explanation rates were generally similar to those in L1 research in which native speakers provided explanations about 60% of the time after IC verbs (Kehler et al., 2008).

¹¹ In the non-IC context, the analysis showed no L1–L2 group difference in NP2 reference rates ($\beta = -0.18$, $SE = 0.23$, $z = -0.78$, $p = .437$), indicating that the L1 and L2 participants were affected by the recency of the NP2 entity, if any, to a similar extent. Considering this, the reduced coreference bias in the L2 participants in the NP1 context cannot be explained by the idea that the L2 participants were more strongly affected by the recency than the L1 participants.

significantly weaker preference for explanation continuations after IC verbs than the L1 group. These findings suggest that L2 learners' fundamental weakness in IC processing is in coherence bias; they were significantly weaker than native speakers in generating discourse explanatory expectations from IC verb semantics. Owing to this limited explanatory expectation, the upcoming explanation must be explicitly signaled for learners to use the IC coreference bias online.

These findings indicate that the source of IC bias do not vary qualitatively between native speakers and L2 learners. In either case, coreference bias results from the explanatory coherence bias, which stems from IC verb semantics. Therefore, in the L2 context, this study supports the accounts (including the empty-slot theory) that attribute the source of IC bias to verb semantics (Bott and Solstad, 2014, 2021; Crinean and Garnham, 2006).

Furthermore, this study discovered that a significant L1–L2 distinction was located in the ease with which an implicit explanation relation can be activated through expectation-based processing. Hence, this study supports growing evidence indicating that L2 comprehension—compared to L1 comprehension—is characterized by reduced expectations (Cheng and Almor, 2018; Grüter et al., 2017; Kim and Grüter, 2021; Lew-Williams and Fernald, 2010).

The reduced expectation effects observed among the L2 participants may be related to the quality of their lexical representations. Specifically, the underspecified quality of L2 lexical representations constrained the retrieval of L2 semantic information, negatively affecting the generation of semantically driven expectations (Kaan, 2014; Kim and Grüter, 2021). In support of this reasoning, recent reviews of L2 prediction have indicated that the accuracy or consistency of lexical representations is a major factor in the reduced prediction effects in L2 (Kaan, 2014; Schlenter, 2023). In relation to the empirical evidence of IC, Kim and Grüter (2021) cited L2 learners' underspecified lexical representations as a potential cause of their weaker online IC effects compared with the stronger and more persistent IC effects in native speakers. Recall that, in the present study, L2 participants were confirmed by the translation task to know the meanings of IC verbs. Thus, their reduced expectations likely resulted from the *utilization* of the existing knowledge of IC verbs rather than a *lack* of that knowledge. Specifically, although L2 learners knew the meaning of IC verbs (as confirmed by the translation task), their representations of the knowledge of IC verbs were less detailed than those of native speakers. Consequently, the learners utilized their lexical knowledge of IC verbs less effectively than native speakers to expect an explanation.

From a theoretical standpoint, this perspective aligns with the lexical bottleneck hypothesis. Originally, this account posits that low-quality lexical representations constrain the integration of lexical and syntactic information (Hopp, 2013, 2018). However, the present study argues that it can be extended to incorporate expectation-based processing, in line with Kim and Grüter's (2021) argument that underdeveloped lexical representations can impede the effective use of lexical semantics, which is crucial for generating expectations. Accordingly, less developed lexical representations are assumed to limit learners' IC-based expectations of how the global discourse unfolds.

An alternative (not mutually exclusive) explanation for the reduced expectation effects involves neurocognitive components such as working memory and attentional resources. Specifically, in L2 comprehension, a substantial percentage of attentional resources in working memory is devoted to local-level processing (e.g., the recognition of individual words and parsing of the current clause). This leaves few resources for retrieving semantic information from the prior discourse and integrating multiple pieces of information, both of which are necessary for discourse-level expectations (e.g., in the case of IC coherence bias, combining verb semantics with the knowledge of event cause). Particularly, forming explanatory expectations from NP1 verbs is assumed to demand extensive attentional resources for the current L2 participants because these verbs are morphemically different from their counterparts in their L1. Supporting this idea, Experiment 2 found that L2 participants showed no coherence bias originating from NP1 verbs.

Weaker explanatory expectations in the L2 group also align with L2 accounts pointing to L1–L2 differences in predictive processing (Amos and Pickering, 2020; Corps et al., 2023; Grüter et al., 2017; Hopp, 2018). Specifically, the RAGE hypothesis states that the different processing patterns observed between native speakers and L2 learners result from the learners' reduced ability to generate expectations. Previous empirical studies have supported this view in terms of word- or phrase-level expectations (Grüter and Rohde, 2021; Lew-Williams and Fernald, 2010). The current findings corroborate the idea of reduced L2 expectations and extend it to clause- or sentence-level expectations by demonstrating that L2 learners are less sensitive to the IC coherence bias than native speakers.

In this regard, Grüter et al. (2017) reported seemingly contrasting findings to those of this study. In their research, native speakers and L2 learners expected different types of coherence relations from sentences with perfective and imperfective verbs. Notably, verb aspects are overtly marked by auxiliary verbs and morphemes (e.g., *John handed/was handing a book to Bob*), rendering the distinction between the different aspects explicit. By contrast, both IC coherence and coreference biases lack explicit linguistic markers and are (in agree with the “both IC coherence and coreference biases”) implicated in the semantics of the verb. Owing to its less deterministic nature, IC bias may be more challenging for L2 learners to use than verb aspects.

4.2 Time course of implicit causality in L2 comprehension

Experiment 1 found that only contexts in which NP2 verbs were followed by an explicitly signaled explanation (i.e., the NP2-*because* condition) elicited early IC effects during online L2 comprehension. The absence of online effects from the NP1 bias aligns with previous IC research on L1-Japanese learners (Hijikata, 2021; Hosoda, 2023) and may be attributed to cross-linguistic differences. In Japanese, NP1 verbs include overt causative morphemes (e.g., *konran-saseru* and *shitsubou-saseru*) that explicitly signal causation, whereas English NP1 verbs lack such morphemes (e.g., *confuse* and *disappoint*), thereby leaving causation implicitly encoded. This disparity likely hindered

learners' ability to use IC information from NP1 verbs to guide online coreference processing. Another notable observation from the translation task was that L2 participants often incorrectly translated NP1 verbs by assigning them passive instead of transitive causative meanings. Although these mistranslated items were excluded from the analysis, this finding suggests that L2 participants' representations of causation expressed by English NP1 verbs were less developed.

Together, the findings of this study suggest that an explicit explanation relation and similar linguistic features in L1 are necessary for IC to influence online L2 coreference processing. When the explanation relation was not signaled (in the full-stop condition), learners use the IC only in the offline task. When verbs are differently encoded in L1 (NP1 verbs), IC is available in neither the online nor offline task.

This conclusion indicates that the difficulty L2 learners experience in applying IC bias to discourse processes primarily stems from cross-linguistic interference at the word level. As discussed above, the key difference between Japanese and English with respect to IC bias lies in the presence or absence of an overt causative morpheme in NP1 verbs. NP2 verbs do not explicitly mark causation in either language, and both languages share similar discourse-level coherence relations, which can be explicitly signaled by connectives (e.g., *nazenara* in Japanese, *because* in English) or inferred from context. According to the empty-slot theory (Bott and Solstad, 2014, 2021; Solstad and Bott, 2022), Japanese learners of English are less sensitive to the explanatory empty slot due to their underspecified representations of causation in English NP1 verbs, which can be attributed to cross-linguistic interference. As a result, they are less likely to expect that events denoted by NP1 verbs will be followed by an explanation in the upcoming discourse, manifesting as the absence of coherence bias. Moreover, the underspecified representations likely increase the cognitive demands required to compute the causally implicated referent during real-time comprehension. Consequently, online IC effects were observed only for NP2 bias, while NP1 bias effects did not emerge—even in the *because* condition, where the explanation relation was explicitly signaled.

These findings indicate that IC's entry into online L2 comprehension is determined by both discourse coherence relations and learners' L1 backgrounds. Apparently, the time course of IC in L2 cannot be explained simply by traditional focusing or integration account. The focusing account proposes that IC enters comprehension processes immediately after the bias-consistent pronoun (McKoon et al., 1993), whereas, according to the integration account, IC exerts effects only at the end of the sentence (Stewart et al., 2000). Neither account adequately explains the modulation of discourse or learner factors in the time course of IC. Thus, this study agrees with the incremental integration account, which integrates the focusing and integration phases to describe the full spectrum of IC effects (Koornneef et al., 2016; Koornneef and Sanders, 2013). This model maintains that the focusing phase determines the timing at which IC begins to affect comprehension, whereas the integration phase determines how linguistic factors and speakers' individual differences modulate the IC effects.

This account explains the present findings that the integration phase of IC is conditioned by coherence relations and L1 linguistic

properties. Specifically, L2 learners can use IC bias online under the condition that the explanation relation is explicitly signaled and the IC verbs are similarly represented in L1 and L2. Regarding the focusing phase, the current study's findings, however, do not provide definitive conclusions. This is because under the current methodology, online IC effects are observable only at or after the pronoun is encountered as IC effects are operationalized by the consistency between the pronoun and IC bias. Therefore, whether the starting point of the IC effect is the participants' proactive prediction of the referent *prior to* the pronoun or their incremental integration *after* the pronoun cannot be discerned.

Consequently, I avoid making definitive claims regarding when IC begins to affect L2 learners' comprehension. It is, still, certain that L2 learners do not wait until the end of the sentence to start using IC information; Experiment 1 has found the pronoun inconsistency effects in Regions 2 and 3 that are located in the middle of the sentence. Learners used IC immediately after encountering the pronoun *at the latest*, in line with the accumulating evidence of early IC effects during L2 comprehension (Contemori and Dussias, 2019; Hosoda, 2023; Kim and Grüter, 2021; Wang and Gabriele, 2022). Accordingly, this study contributes to the expanding L2 literature showing that L2 learners can use IC coreference bias not only retroactively but also incrementally during comprehension. The novelty of this study is that it specifies the conditions under which such online IC effects occur in L2 processing as well as when and why L2 learners fail to use IC for coreference and coherence processing. These discoveries provide a finer-grained picture of the cognitive mechanism of IC processing in L2.

To conclude, several limitations of this study should be discussed to guide future research. First, the sample size was relatively small, resulting in a narrow range of L2 proficiency among participants. A critical limitation is that both experiments tested basic-level L2 learners. If reduced expectations are a characteristic of L2 learners' comprehension, as posited by the RAGE hypothesis, we would expect weakened expectations to persist even among highly proficient learners who typically possess richer L2 lexical representations and are less constrained by attentional resources. Further empirical investigations are required to clarify this aspect.

Second, this study did not compare online IC processing of L2 learners with that of native speakers. Although online IC processing can be assumed to pose greater difficulty for L2 learners than for native speakers (see the second paragraph of the General Discussion), further research comparing L1 and L2 groups is necessary to corroborate this view.

Third, although eye tracking can measure the time course of comprehension processes, it does not directly reveal the specific cognitive operations that participants engaged in during the task. For example, L2 learners might strategically code-switch or code-mix while processing L2 materials, which could confound their processing time. To capture L2 learners' IC processing more thoroughly, future research should complement the present findings with neurophysiological measures, such as event-related potentials.

Finally, the extent to which L2 participants generated explanatory expectations during real-time comprehension is unclear because Experiment 2 used the offline story-continuation

task. Future research should supplement this study's findings with online measures.

Nevertheless, the fact that L2 learners showed reduced expectation effects even in the offline task does not necessarily undermine the conclusions of this study. Considering that the online task involves higher cognitive demands, it seems unlikely that L2 learners would generate explanatory expectations online if they already fail to do so in the offline task.

Data availability statement

Publicly available datasets were analysed in this study. This data can be found here: https://osf.io/3qnz6/?view_only=a4e46c7da585467e994632f4f57aee4b.

Ethics statement

The studies involving humans were approved by Hokkaido University of Education. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

MH: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

References

- Amos, R. M., and Pickering, M. J. (2020). A theory of prediction in simultaneous interpreting. *Biling.: Lang. Cogn.* 23, 706–715. doi: 10.1017/S1366728919000671
- Bott, O., and Solstad, T. (2014). "From verbs to discourse: a novel account of implicit causality," in *Psycholinguistic Approaches to Meaning and Understanding Across Languages*, eds. C. Fabricius-Hanse, B. Hemforth, and B. Mertins (Cham: Springer), 213–251.
- Bott, O., and Solstad, T. (2021). Discourse expectations: Explaining the implicit causality biases of verbs. *Linguistics* 2, 361–416. doi: 10.1515/ling-2021-0007
- Carpenter, P. A., and Just, M. A. (1983). "What your eyes do while your mind is reading," in *Eye Movements in Reading: Perceptual and Language Processes*, eds. K. Rayner (San Diego, CA: Academic Press), 275–307.
- Cheng, W., and Almor, A. (2017). The effect of implicit causality and consequentiality on nonnative pronoun resolution. *Appl. Psycholinguist.* 38, 1–26. doi: 10.1017/S0142716416000035
- Cheng, W., and Almor, A. (2018). A Bayesian approach to establishing coreference in second language discourse: Evidence from implicit causality and consequentiality verbs. *Bilingualism: Lang. Cognit.* 22, 456–475. doi: 10.1017/S136672891800055X
- Clifton, C., Staub, A., and Rayner, K. (2007). "Eye movements in reading words and sentences," in *Eye Movements: A Window on Mind and Brain*, eds. R. V. Gompel, M. Fisher, W. Murray, and R. L. Hill (Amsterdam: Elsevier), 341–371.
- Contemori, C., and Dussias, P. E. (2019). Prediction at the discourse level in Spanish–English bilinguals: an eye-tracking study. *Front. Psychol.* 10:956. doi: 10.3389/fpsyg.2019.00956
- Corps, R. E., Liao, M., and Pickering, M. J. (2023). Evidence for two stages of prediction in non-native speakers: a visual-world eye-tracking study. *Bilingualism: Lang. Cognit.* 26, 231–243. doi: 10.1017/S1366728922000499
- Crinean, M., and Garnham, A. (2006). Implicit causality, implicit consequentiality and semantic roles. *Lang. Cogn. Process.* 21, 636–648. doi: 10.1080/01690960500199763
- Cummings, I., and Sturt, P. (2018). Retrieval interference and semantic interpretation. *J. Memory Lang.* 102, 16–27. doi: 10.1016/j.jml.2018.05.001
- Ferstl, E. C., Garnham, A., and Manouilidou, C. (2011). Implicit causality bias in English: a corpus of 300 verbs. *Behav. Res. Methods* 43, 124–135. doi: 10.3758/s13428-010-0023-2
- Grüter, T., and Rohde, H. (2021). Limits on expectation-based processing: use of grammatical aspect for co-reference in L2. *Appl. Psycholinguist.* 42, 51–75. doi: 10.1017/S0142716420000582
- Grüter, T., Rohde, H., and Schafer, A. J. (2017). Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguist. Approach. Bilingual.* 7, 199–229. doi: 10.1075/lab.15011.gru
- Hartshorne, J. K., O'Donnell, T. J., and Tenenbaum, J. B. (2015). The causes and consequences explicit in verbs. *Lang. Cognit. Neurosci.* 30, 716–734. doi: 10.1080/23273798.2015.1008524
- Hartshorne, J. K., Sudo, Y., and Uruwashii, M. (2013). Are implicit causality pronoun resolution biases consistent across languages and cultures? *Exp. Psychol.* 60, 179–196. doi: 10.1027/1618-3169/a000187
- Hijikata, Y. (2021). The time course of the effects of implicit causality bias on anaphora resolution by Japanese learners of English. *Ann. Rev. English Lang. Educ. Japan (ARELE)* 32, 1–16. doi: 10.20581/arele.32.0_1
- Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Lang. Res.* 29, 33–56. doi: 10.1177/0267658312461803

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported by JSPS Grant-in-Aid for Early-Career Scientists (Nos. 20K13127 & 24K16137).

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/flang.2025.1494500/full#supplementary-material>

- Hopp, H. (2018). The bilingual mental lexicon in L2 sentence processing. *Second Lang.* 17, 5–27. doi: 10.11431/secondlanguage.17.0_5
- Hosoda, M. (2020). Establishing coreference in Japanese EFL learners' using verbs' implicit causality: A sentence completion study. *ARELE* 31, 193–208.
- Hosoda, M. (2022). "Expecting coherence relations from verbs' implicit causality: a comparison between L2 learners and native speakers," in *Proceedings of the 47th Japan Society of English Language Education Conference in Hokkaido*, 130–131.
- Hosoda, M. (2023). Time course of verbs' implicit causality during L2 comprehension: An extended replication of Hijikata 2021 Japanese EFL learners. *ARELE* 34, 97–112.
- JACET Basic Word Revision Committee (2016). *The New JACET list of 8000 Basic Words*. Tokyo: Kiriara Shoten.
- Kaan, E. (2014). Predictive sentence processing in L2 and L1: what is different? *Linguist. Approach. Bilingual.* 4, 257–282. doi: 10.1075/lab.4.2.05kaa
- Kehler, A., Kertz, L., Rohde, H., and Elman, J. L. (2008). Coherence and coreference revisited. *J. Semant.* 25, 1–44. doi: 10.1093/jos/ffm018
- Kim, H., and Grüter, T. (2021). Predictive processing of implicit causality in a second language: a visual-world eye-tracking study. *Stud. Second Lang. Acquisit.* 43, 133–154. doi: 10.1017/S0272263120000443
- Koornneef, A., Dotlačil, J., van den Broek, P., and Sanders, T. (2016). The influence of linguistic and cognitive factors on the time course of verb-based implicit causality. *Quart. J. Exp. Psychol.* 69, 455–481. doi: 10.1080/17470218.2015.1055282
- Koornneef, A. W., and Sanders, T. J. (2013). Establishing coherence relations in discourse: The influence of implicit causality and connectives on pronoun resolution. *Lang. Cogn. Process.* 28, 1169–1206. doi: 10.1080/01690965.2012.699076
- Koornneef, A. W., and Vanberkum, J. (2006). On the use of verb-based implicit causality in sentence comprehension: evidence from self-paced reading and eye tracking. *J. Mem. Lang.* 54, 445–465. doi: 10.1016/j.jml.2005.12.003
- Lew-Williams, C., and Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *J. Mem. Lang.* 63, 447–464. doi: 10.1016/j.jml.2010.07.003
- McKoon, G., Greene, S. B., and Ratcliff, R. (1993). Discourse models, pronoun resolution, and the implicit causality of verbs. *J. Exp. Psychol.* 19, 1040–1052. doi: 10.1037//0278-7393.19.5.1040
- Papageorgiou, S., Tannenbaum, R. J., Bridgeman, B., and Cho, Y. (2015). *The Association Between TOEFL iBT® Test Scores and the Common European Framework of Reference (CEFR) Levels*. Princeton, NJ: Educational Testing Service.
- Pyykkönen, P., and Järvikivi, J. (2010). Activation and persistence of implicit causality information in spoken language comprehension. *Exp. Psychol.* 57, 5–16. doi: 10.1027/1618-3169/a000002
- Schlenter, J. (2023). Prediction in bilingual sentence processing: How prediction differs in a later learned language from a first language. *Biling.: Lang. Cogn.* 26, 253–267. doi: 10.1017/S1366728922000736
- Solstad, T., and Bott, O. (2022). On the nature of implicit causality and consequentiality: the case of psychological verbs. *Lang. Cognit. Neurosci.* 37, 1311–1340. doi: 10.1080/23273798.2022.2069277
- Sorace, A. (2011). Pinning down the concept of 'interface' in bilingualism. *Linguist. Approach. Bilingual.* 1, 1–33. doi: 10.1075/lab.1.1.01sor
- Stewart, A. J., Pickering, M. J., and Sanford, A. J. (2000). The time course of the influence of implicit causality information: focusing versus integration accounts. *J. Mem. Lang.* 42, 423–443. doi: 10.1006/jmla.1999.2691
- Wang, T., and Gabriele, A. (2022). "Individual differences modulate sensitivity to implicit causality bias in both native and nonnative processing," in *Studies in Second Language Acquisition, First View* (Cambridge: Cambridge University Press), 1–29.