

Incremental or delayed processing? L2 learners' active gap-filling in sentence comprehension

Chie Nakamura, Suzanne Flynn, Yoichi Miyamoto & Noriaki Yusa*

Abstract. The contrast between first language (L1) and second language (L2) sentence processing provides valuable insights into how linguistic knowledge and cognitive mechanisms shape real-time comprehension. While L1 speakers are known to engage in incremental parsing—integrating linguistic input as it becomes available—it remains debated whether adult L2 learners adopt similar strategies or rely on delayed, non-structural parsing as proposed by the Shallow Structure Hypothesis (SSH). To investigate this, the present study employed the visual world eye-tracking paradigm to examine how Japanese L2 learners of English process structurally ambiguous *wh*-questions. Results revealed a strong preference for main clause (MC) interpretations, mirroring native speaker behavior. Crucially, eye-tracking data showed early fixation patterns consistent with active gap-filling, providing evidence against SSH. Furthermore, L2 learners demonstrated an even stronger commitment to the MC analysis than native speakers, suggesting that while their syntactic processing is guided by incremental structural parsing, it may also reflect a greater reliance on initial interpretations.

Keywords. L2 sentence processing; incremental parsing; filler-gap dependencies; visual world eye-tracking

1. Introduction. Real-time sentence comprehension involves incremental parsing, in which listeners integrate linguistic cues as they process them. Research on first language (L1) processing has shown that comprehenders actively predict upcoming syntactic structures and resolve dependencies as early as possible (Omaki et al., 2014). This process has been investigated through the processing of *wh*-questions such as (1), where comprehenders must determine the grammatical role of the fronted *wh*-phrase and establish its syntactic relationship with the rest of the sentence.

(1) Where did Lizzie tell someone that she was going to catch butterflies?

Sentence (1) is structurally ambiguous because the *wh*-phrase *where* can be associated with either the main clause (MC) (*tell someone*) or the embedded clause (EC) (*catch butterflies*). In the MC interpretation, *where* refers to the location of Lizzie's telling event, meaning the question asks where Lizzie told someone something. In contrast, in the EC interpretation, *where* refers to the location of the catching event, meaning the question asks where Lizzie was going to catch butterflies.

Previous research has shown that native English speakers, both adults and children, typically favor the MC interpretation, rapidly linking the *wh*-phrase to the main clause verb (*tell*). This preference reflects an active parsing strategy, in which comprehenders attempt to resolve filler-gap dependencies as early as possible. This strategy is described by the *active gap-filling*

* Authors: Chie Nakamura, Waseda University, Tokyo, Japan (cnakamura@waseda.jp), Suzanne Flynn, Massachusetts Institute of Technology (sflynn@mit.edu), Yoichi Miyamoto, Osaka University (y.miyamoto.hmt@osaka-u.ac.jp) & Noriaki Yusa, Miyagi Gakuin Women's University (yusa@mgu.ac.jp).

hypothesis (Omaki et al., 2014), which posits that comprehenders actively seek syntactic structures that allow them to integrate the *wh*-phase as soon as a potential gap site is encountered.

Given that real-time sentence processing, including the processing of filler-gap dependencies, has been shown to be different between native speakers and second language (L2) learners (e.g., Marinis et al., 2005 among others), an open question is whether the active gap-filling hypothesis extends to L2 processing. One of the most influential accounts of L2 processing, the Shallow Structure Hypothesis (SSH) (Clahsen & Felser, 2006) argues that L2 learners rely on more a global parsing strategy, prioritizing overall sentence meaning and semantic plausibility before committing to a specific interpretation (Roberts, 2010). Unlike native speakers, L2 learners are thought to process sentences in a less incremental and more delayed manner, suggesting that they may not immediately resolve filler-gap dependencies upon encountering potential cues. Instead, they may postpone interpretation until additional contextual information becomes available. If this is the case, L2 learners may exhibit a different preference pattern than that of native speakers, possibly favoring the EC interpretation in processing sentences such as (1), as the event of catching butterflies follows the telling event in the sentence.

However, contrary to the predictions of the SSH, some recent studies suggest that L2 learners, particularly those with sufficient exposure to the target language, can develop incremental processing strategies similar to native speakers (Hopp, 2015; McLaughlin et al., 2004; Weber-Fox & Neville, 1996). For example, Omaki and Schulz (2011) provide evidence that L2 learners, like native speakers, actively search for a gap location during sentence processing. Their study suggests that L2 learners do not always rely solely on a global parsing strategy but may instead attempt to resolve filler-gap dependencies as they process a sentence incrementally.

These findings challenge the view that L2 processing is purely meaning-driven by showing that L2 learners, like native speakers, actively build syntactic representations in real time rather than relying only on post-hoc interpretation. They also highlight the influence of task demands and the experimental design on processing outcomes. Notably, many studies that support delayed processing, particularly those aligned with the SSH, often employ offline tasks such as sentence interpretation tasks, grammaticality judgments, or self-paced reading. These methods primarily capture end-state interpretations, rather than the incremental parsing process, and thus may underestimate L2 learners' real-time syntactic processing and predictive parsing abilities.

To reconcile the SSH with evidence for incremental processing, the present study used the visual world eye-tracking paradigm, which allows for the fine-grained tracking of real-time language processing through participants' eye movements during sentence comprehension. Specifically, this study investigated how Japanese-speaking L2 learners of English process *wh*-question, as illustrated in example (1).

2. Methods.

Participants

We recruited 51 native Japanese speakers learning English as their L2, along with 11 native English speakers as a control group. All participants reported normal or corrected-to-normal vision. The proficiency level of the L2 participants ranged from B2 to C1 (Independent user level to Proficient user level) on the Common European Framework of References for Languages (CEFR).

Materials and Design

We used a visual-world eye-tracking paradigm to investigate how L2 learners and native speakers process structurally ambiguous *wh*-questions while their eye movements were recorded. As

participants listened to the wh-question sentences, a corresponding visual display, such as in Figure 1, was presented simultaneously. Each trial began with a context sentence followed by a wh-question to examine how participants resolved structural ambiguity in real time. At the end of the sentence, participants indicated their interpretation by pressing a key corresponding to their chosen answer.

The context sentences varied in event order, creating two conditions. In the EC-MC event condition, the EC event occurred first, followed by the MC event (2a). In the MC-EC event condition, the MC event was presented first, followed by the EC event (2b). This manipulation controlled for recency effects, ensuring that any observed preferences were driven by syntactic processing rather than memory-based biases.

(2) a. EC-MC event condition:

Lizzie caught butterflies in the park using her net. In the afternoon, she saw her friend in the schoolyard and told him about it.

b. MC-EC event condition:

Lizzie saw her friend in the schoolyard. She told him, "I caught butterflies in the park using my net".

Following the context sentence, participants heard a wh-question, such as in (3). The question conditions were manipulated by introducing an unambiguous version. In the Ambiguous question condition (3a), the correct answer could refer to either the location of the telling event (the schoolyard) or the butterfly-catching event (the park). In the Unambiguous question condition (3b), *that* was replaced by *how*, ensuring that the question unambiguously referred to the telling location (the schoolyard). This manipulation was designed to examine how L2 learners process filler-gap dependencies when the gap location is explicitly constrained.

(3) a. Ambiguous question condition:

Where did Lizzie tell someone that she was going to catch butterflies?

Answer: Ambiguous (could refer to the schoolyard or the park)

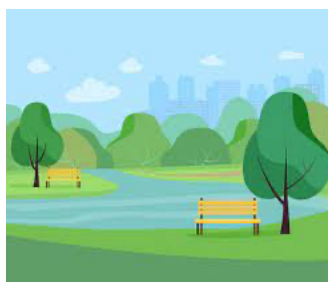
b. Unambiguous question condition:

Where did Lizzie tell someone how she was going to catch butterflies?

Answer: Unambiguous (refers only to the schoolyard, the telling location)



a. Main clause (MC) interpretation



b. Embedded clause (EC) interpretation



c. Distracter

Figure 1. Three pictures presented with (3)

This resulted in a 2 x 2 factorial design, with two context conditions (EC-MC event vs. MC-EC event) and two question conditions (Ambiguous vs. Unambiguous), leading to four experimental conditions. Each participant completed 24 target items and 36 unrelated filler items.

Data Analysis

Key-press responses were first analyzed to assess participants' final interpretations in each condition, comparing the response patterns of native speakers and L2 learners. Following this, the time course of eye-tracking data was examined to track participants' fixations on the MC and EC pictures as they processed the ambiguous wh- question in real time.

3. Results.

Key-Press Responses

Accuracy on filler questions was high for both groups, with native speakers achieving 90.8% and L2 learners achieving 83.4%, indicating that all participants remained engaged and attentive throughout the task. For unambiguous target questions (3b), native speakers demonstrated 85.2% accuracy, while L2 learners performed at 68.4% accuracy. This suggests that although both groups were generally able to process the structure correctly, L2 learners had greater difficulty resolving the filler-gap dependency, even when the correct answer was explicitly marked.

Crucially, for ambiguous target questions (3a), which had no objectively correct answer regarding the two possible event locations, we analyzed the proportion of participants' choices between the MC and EC pictures. Figure 2a illustrates native speakers' choices, revealing a clear preference for the MC interpretation. Native speakers selected the MC picture 81.9% of the time, while choosing the EC picture only 17.3% of the time. Figure 2b presents the responses of L2 learners. Similar to native speakers, L2 learners exhibited a strong preference for the MC interpretation, selecting the MC picture 72.2% of the time and the EC picture only 24.6% of the time.

These results suggest that L2 learners processed the ambiguous wh-question in a qualitatively similar manner to native speakers, favoring the MC interpretation over the EC interpretation.

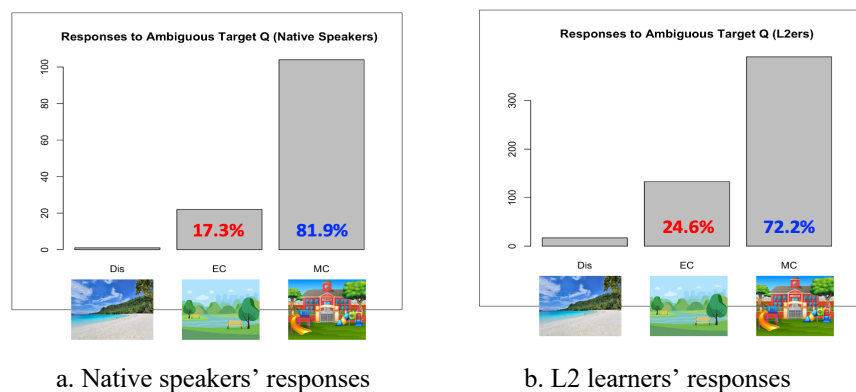


Figure 2. Key-pressing responses for ambiguous questions

Eye-Tracking Results

Figure 3 illustrates the time course of native speakers' eye movements while listening to the target questions in the ambiguous condition (3a). For statistical analysis, the sentence duration was

divided into three distinct time windows: (1) TW1 (1130 ms) – from sentence onset to the mean onset of the MC verb (e.g., *tell* in (3a)) across all items, (2) TW2 (1140 ms) – from the MC verb onset of each item to the mean onset of the EC verb (*catch* in (3a)) across all items, and (3) TW3 (1287 ms) – from the EC verb onset of each item to the mean sentence offset across all items.

To determine whether there was a statistically significant difference in participants' proportion of fixation on the MC object and the EC object, we conducted a linear mixed-effects (LME) analysis, using fixation proportion to these two objects as the dependent variable (Baayen, et al., 2008). The analysis was performed separately for each of the three time windows (TW1, TW2, and TW3). In TW1, there was no significant difference between looks to MC picture and EC picture. In TW2, statistical analysis confirmed that native speakers looked significantly more at the MC picture than the EC picture ($p > .001$). This pattern persisted in TW3, where native speakers continued to fixate significantly more on the MC picture than the EC picture ($p > .001$), suggesting a strong preference for the MC interpretation in real-time processing.

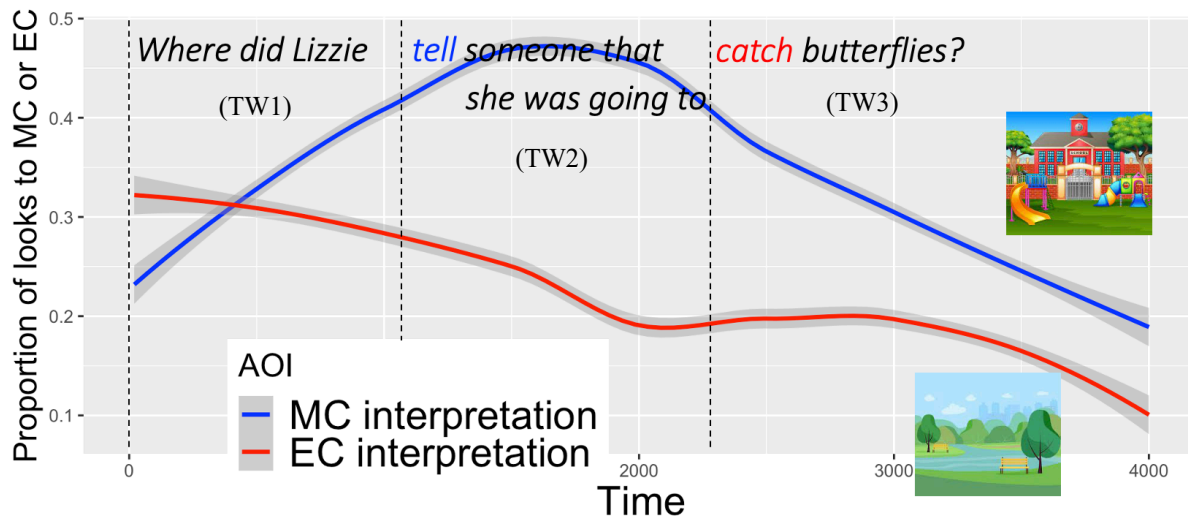


Figure 3. Time course of eye-movements on MC and EC pictures (Native speakers)

Figure 4 illustrates the results for L2 learners. Similar to native speakers, L2 learners fixated on the MC picture more frequently throughout the duration of the question. Statistical analysis was conducted across the same three predefined time windows. The results showed that L2 learners fixated significantly more on the MC picture than the EC picture in all time windows ($ps > .001$), indicating a clear preference for the MC interpretation throughout the processing of the ambiguous wh- questions.

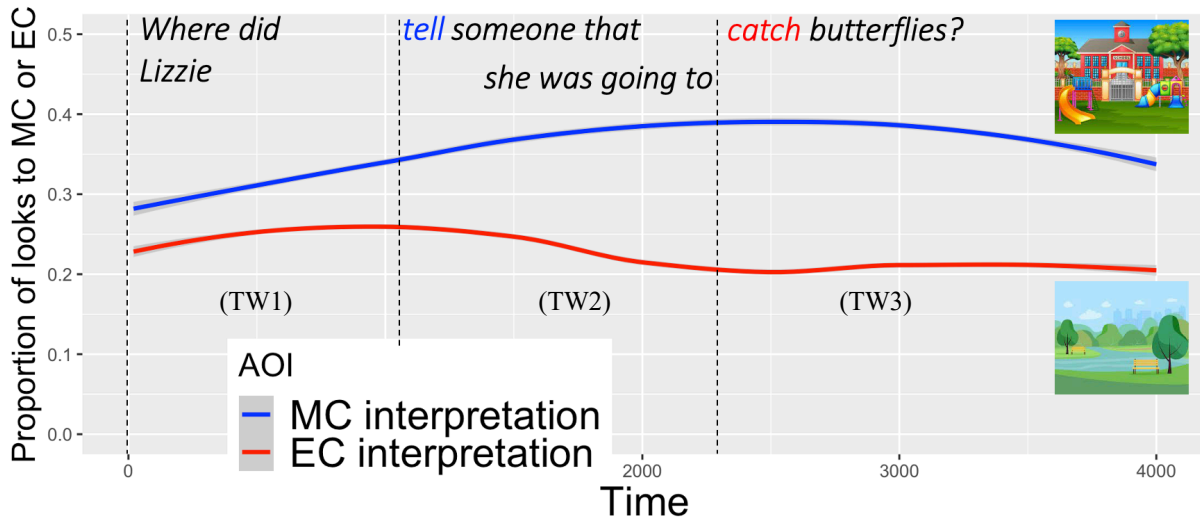


Figure 4. Time course of eye-movements on MC and EC pictures (L2 learners)

4. Discussion.

The results of the current study provide evidence that L2 learners actively resolve filler-gap dependencies during sentence processing. Both the key-pressing data and eye-tracking results indicate that, like native speakers, L2 learners prioritize the first verb phrase they encounter when processing sentences when building filler-gap dependencies during sentence processing in their second language. Interestingly, the time course of the eye-movement analysis revealed that L2 learners exhibited this preference even more strongly than native speakers, as reflected in their more immediate and sustained fixations on the MC picture. This pattern may reflect two key characteristics of L2 processing.

First, L2 learners may develop a stronger expectation for the first verb (e.g., *tell* in (3a)) even before its onset, relying on predictive processing based on surface-level cues. Given that the matrix verb *tell* was used across all target items, L2 learners might have adopted a heuristic strategy to anticipate this verb, leading them to commit to the MC analysis even earlier than native speakers.

Second, once L2 learners commit to an initial interpretation, they are less likely to entertain alternative possibilities, likely due to cognitive resource and working memory constraints. Unlike native speakers, who flexibly adjust their interpretations as new linguistic input unfolds, L2 learners appear to adopt a more rigid processing strategy that minimizes reanalysis. This suggests that while both groups engage in prediction, L2 learners may prioritize efficiency over flexibility, leading to a stronger but potentially less adaptable initial commitment.

Overall, our findings suggest that incremental processing is a universal mechanism in language comprehension, where both native speakers and L2 learners process sentences step-by-step, building interpretations incrementally as they hear them, rather than waiting for the entire sentence to unfold. Moreover, real-time eye-tracking data revealed that while L2 learners' fixation patterns showed structural guidance similar to native speakers, they exhibited an even stronger commitment to MC analysis, suggesting that their syntactic processing is guided by structure but leans more heavily toward an initial interpretation.

Crucially, these findings challenge the SSH, which posits that L2 learners delay commitment to syntactic structures until the sentence is fully processed. Instead, the results demonstrate that

L2 learners engage in incremental processing, actively resolving filler-gap dependencies as they encounter relevant cues. This suggests that, rather than being fundamentally different from native speakers, L2 learners rely on similar structural processing mechanisms, albeit with stronger reliance on early cues and reduced flexibility in revision.

It is worth noting that in Japanese, question particles appear at the end of the sentence, providing the necessary information to interpret fronted wh-phrases (Kuroda, 1965; Saito, 2017). In terms of processing, Nakamura et al. (2025) report that native Japanese speakers tend to wait until the end of the sentence to determine the interpretation, as Japanese grammar requires the question particle at the end to fully determine the interpretation of fronted wh-phrases (but see Aoshima et al., 2004, for evidence that Japanese speakers actively posit gap sites). This contrasts with the real-time processing behavior of Japanese L2 learners of English in the current study, who exhibited an early preference for MC interpretation. This divergence between L1 Japanese processing and L2 English processing suggests that L2 learners do not fully transfer their typical L1 processing strategy but instead adapt to the structural properties of English, where wh-dependencies are typically resolved incrementally.

This cross-linguistic difference is particularly important because it demonstrates that the observed effects cannot be attributed to structural similarities between learners' L1 and L2. Our finding adds an important dimension to previous studies that examined English L2 learners whose L1 is structurally similar to English, such as Spanish L1-English L2 speakers (e.g., Omaki & Schulz, 2011). In those cases, both languages share word order similarities, making it unclear whether L2 learners' processing strategies were shaped by L1 transfer or by universal mechanisms of sentence comprehension. The current study helps clarify this issue by demonstrating that L2 learners can adopt new processing strategies that differ from their L1, reinforcing the idea that incremental processing is not solely driven by native language transfer but rather by adaptation to target language syntax.

While the current study focused on the processing of filler-gap dependencies, future studies should continue to explore a broader range of syntactic phenomena to gain a more comprehensive understanding of how learners process complex syntactic information and adapt their parsing strategies. Additionally, investigating cross-linguistic differences in syntactic processing across various language pairs would provide valuable insights into the adaptability of incremental processing across diverse linguistic contexts.

References

- Aoshima, S., Phillips, C., & Weinberg, A. (2004). Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language*, 51(1), 23-54.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390-412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(1), 3-42. <https://doi.org/10.1017/S0142716406060024>
- Felser, C., & Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research*, 23(1), 9-36. <https://doi.org/10.1177/0267658307071600>
- Hopp, H. (2015). Individual differences in the second language processing of object-subject ambiguities. *Applied Psycholinguistics*, 36(2), 129-173. <https://doi.org/10.1017/S0142716413000180>

- Kuroda, S.-Y. (1965). Generative grammatical studies in the Japanese language [Doctoral dissertation, Massachusetts Institute of Technology]. Dspace@MIT. <https://dspace.mit.edu/handle/1721.1/13078>
- Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, 27(1), 53–78. <https://doi.org/10.1017/S0272263105050035>
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, 7(7), 703–704. <https://doi.org/10.1038/nn1264>
- Nakamura, C., Flynn, S., Miyamoto, Y., & Yusa, N. (2025, June). Comparing the processing of wh-question in Japanese-English bilinguals: A cross linguistic eye-tracking study. [Paper presentation]. 15th International Symposium on Bilingualism, Donostia–San Sebastián, Spain.
- Omaki, A., Davidson White, I., Goro, T., Lidz, J., & Phillips, C. (2014). No fear of commitment: Children’s incremental interpretation in English and Japanese wh-questions. *Language Learning and Development*, 10(3), 206–233. <https://doi.org/10.1080/15475441.2013.858178>
- Omaki, A., & Schulz, B. (2011). Filler-gap dependencies and island constraints in second-language sentence processing. *Studies in Second Language Acquisition*, 33(4), 563–588. <https://doi.org/10.1017/S0272263111000313>
- Omaki, A., Lau, E. F., Davidson White, I., Dakan, M., Lidz, J., & Phillips, C. (2014). Hyper-active gap filling. *Frontiers in Psychology*, 6, Article 384. <https://doi.org/10.3389/fpsyg.2015.00384>
- Roberts, L. (2010). Syntactic processing in the second language: The timing of ambiguity resolution and reanalysis. *Studies in Second Language Acquisition*, 32(4), 535–562.
- Saito, M. (2017). Japanese wh-phrases as operators with unspecified quantificational force. *Language and Linguistics*, 18(1), 1–25. <https://doi.org/10.1075/lali.18.1.01sai>
- Weber-Fox, C., & Neville, H. J. (1996). Maturation constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8(3), 231–256. <https://doi.org/10.1162/jocn.1996.8.3.231>