Studies in Second Language Acquisition (2026), **48**, 478–490
doi:10.1017/S0272263125101472

RESEARCH REPORT

Processing of passives using morphological information within the verb in a second language: Slower, but as robust as in a first language

Haerim Hwang and Ganling Han

Department of English, The Chinese University of Hong Kong, Hong Kong
Corresponding author: Haerim Hwang; Email: haerimhwang@cuhk.edu.hk

(Received 05 February 2025; Revised 24 October 2025; Accepted 18 November 2025)

Abstract

This study aims to illuminate the underlying mechanisms of sentence processing in L2 speakers. The phenomenon of interest in the study is the passive structure, which prior research has shown can be challenging for both L1 speakers and L2 speakers to process compared to active structures. Using a visual-world eye-tracking paradigm, this study investigates whether L1-English speakers and L1-Cantonese L2-English speakers employ a morphological cue within the verb to process English actives and passives, and if so, specifically when these cues are integrated into their processing. The results from a growth curve analysis and a divergence point analysis show that the L2-English speakers were slower than the L1-English speakers, but did use the morphological cue to process both actives and passives, even though this cue is absent in their L1 Cantonese. These results suggest that, despite differences in processing speed, the mechanisms underlying L1 and L2 processing are similar.

Keywords: second language processing; underlying processing mechanism; processing speed; passive; visual-world eye-tracking

Introduction

A long-standing debate in the field of second language (L2) processing revolves around the question of whether first language (L1) and L2 speakers rely on different underlying mechanisms to process sentences. An influential hypothesis in this regard is the shallow structure hypothesis, which suggests that L2 speakers have limited access to native-like parsing mechanisms (Clahsen & Felser, 2006) or that their morphosyntactic representation is "less robust in the L2 than in the L1," leading to a greater reliance on lexical and pragmatic information instead (Clahsen & Felser, 2018, p. 701). An alternative approach holds that L2 speakers can effectively use complete morphosyntactic

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (http://creativecommons.org/licenses/by-nc-nd/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use and/or adaptation of the article.

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press

information, making their processing akin to that of L1 speakers. According to this approach, the underlying mechanisms are basically similar between L1 and L2 processing (Cunnings, 2017; Fernandez et al., 2018), and any possible differences observed between the two groups are attributable to L2 learners' lower proficiency levels (Cunnings, 2017) or limitations in working memory or decoding abilities (McDonald, 2006).

Moving beyond a simple comparison of L1 and L2 sentence processing, psycholinguistic research has taken a more nuanced approach, seeking to identify the factors that lead to L1-L2 similarities/differences and to illuminate their mechanisms. Recent research has focused on how different types of information are integrated during processing and how this process unfolds over time (e.g., Henry et al., 2022; Puebla & Felser, 2024). For example, Puebla and Felser's (2024) eye-tracking-while-reading study showed that while both L1-German speakers and L1-Russian L2-German speakers ended up resolving anaphors by selecting a referent outside the sentence, their initial processing profiles revealed interesting patterns: Whereas the L1-German speakers preferred a local, c-commanding antecedent immediately after encountering the anaphor, the L2-German speakers favored a sentence-external referential antecedent, thereby suggesting that L2 speakers rely more on discourse-level than on syntactic-level information. As such, looking at the timing of L2 processing is crucial because it reveals precisely what information L2 speakers utilize, and when, during comprehension (see also Felser, 2019).

Against this backdrop, the present study aims to shed light on the mechanisms underlying L2 speakers' sentence processing, with a particular focus on the timing of their linguistic cue integration. Specifically, we focus on English passives, which are characterized by their relatively low frequency and noncanonical assignment of thematic roles. In conversation, only 2% of all finite verbs are passive (Xiao et al., 2006). In contrast to actives, with the typical agent-verb-patient/theme word order, as shown in (1), passives have an inverted argument structure where the patient becomes the subject, as shown in (2). The agent in passives is often expressed in a *by*-phrase, which is generally optional but may be pragmatically required or constrained depending on context. From a syntactic perspective, a movement operation leads the patient to end up in the clause-initial subject position. This operation is accompanied by morphological and syntactic changes, including the attachment of the past participle morpheme (e.g., *-ed*) to the main verb and the optional inclusion of the agent in a *by*-phrase. During real-time processing of passives, the past participle morpheme can serve as an initial cue to facilitate comprehension.

(1) The cow was washing the horse.

(2) The cow was washed (by the horse).

The complex nature of passives, coupled with their infrequency in comparison to actives, likely contributes to their processing difficulty, which is demonstrated by slower and more effortful comprehension in both L1 (e.g., Ferreira, 2003) and L2 (e.g., Lee & Doherty, 2019). For example, Lee and Doherty (2019) conducted a complex eye-tracking task where their participants read sentences in either active or passive voice, and then viewed two images and selected the image that depicted the sentence they had read. Although the researchers' goal was to investigate the instruction effect on passive processing using a pretest-posttest design, their pretest results showed that both L1-Spanish speakers and L2-Spanish speakers with diverse L1 backgrounds (e.g., English, Mandarin, Polish) fixated significantly more slowly on the target image for passives than for actives. Furthermore, the L2 speakers demonstrated significantly

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press

lower accuracy in selecting the correct picture for passives (45%) compared to actives (88%), in contrast to the L1-Spanish speakers, who showed a smaller gap (88% vs. 95%). The L2 speakers' low accuracy in this study may be due to their limited language proficiency, which ranged from beginner to low-intermediate levels.

Using an aural forced-choice picture identification task, Crossley et al. (2020) showed a result consistent with that of Lee and Doherty (2019) concerning the processing of passives versus actives. The participants heard a sentence and then clicked on the picture that matched the sentence, with their mouse trajectories recorded. The results showed that L1-English speakers were faster in responding to both active and passive stimuli, and that they made shorter-distance mouse movements with fewer directional changes, compared to L1-Spanish L2-English speakers. However, the two groups exhibited similar difficulties in processing passives, as indicated by their mouse movement trajectories, which were slower, covered longer distances, and showed more spatial competition from the alternative response option in comparison to actives. Furthermore, with increasing proficiency, the L2 speakers converged with the L1 speakers on response time, motion time, and maximum onset velocity. A similar result was observed by Marinis and Saddy (2013) in their study involving children. In a self-paced listening task, both L1-English and L1-Turkish L2-English children exhibited longer reading times at the verb and following positions in both active and passive conditions (e.g., *was kissing the camel*/*was kissed by the camel*) when the sentence they heard did not match the picture they saw, although only the L2-English children showed slower processing for passives compared to actives at the spill-over region.

Notably, the existing L2 processing studies on English passives have employed tasks that cannot provide precise information on when participants actually use a morphological cue to process a sentence, such as eye-tracking-while-reading with picture identification (e.g., Lee & Doherty, 2019), picture identification (e.g., Crossley et al., 2020), and self-paced listening (e.g., Marinis & Saddy, 2013). All these tasks primarily compare picture selection accuracy, fixation/reaction times, or listening times at specific regions across active and passive conditions, rather than pinpointing the exact moment individuals begin to show evidence of processing. Furthermore, in the above-mentioned studies, the L2-English speakers who demonstrated L1-like processing patterns were L1-Spanish (Crossley et al., 2020) or L1-Turkish (Marinis & Saddy, 2013) speakers. In both cases, the L1 has a clear cue within the verb for the passive voice; Spanish verbs include the past participle marker as in English (e.g., [3]), and Turkish verbs contain an independent passive marker (e.g., [4]). Therefore, these speakers' successful L2 processing may have stemmed from the transfer from their L1.

(3) Graciela es abrazada por Pablo y Marcos.
Graciela is–3.SG hugged–3.FEM.SG by Pablo and Marcos
'Graciela is hugged by Pablo and Marcos.'

(Lee & Doherty, 2019, p. 860)

(4) Çamaşır-lar-Ø bir kadın tarafından as-ıl-dı-lar.
cloth-PL-NOM a woman by hang-PASS-PST–3.PL
"The clothes were hung by a woman."

(adapted from Bayram et al., 2019, p. 924)

To address these gaps, the current study explores the following two research questions (RQs) using a visual-world eye-tracking task.

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press

- RQ 1: Do L1-English speakers and L1-Cantonese L2-English speakers utilize a morphological cue within the verb to process actives and passives?
- RQ 2: At what point do L1-English speakers and L1-Cantonese L2-English speakers integrate the morphological cue when processing actives and passives?

Regarding RQ 1, if speakers fixate on the image corresponding to the active or passive event during the verb region, this will suggest that they are able to use a morphological cue to process actives and passives. For RQ 2, a divergence point analysis will be used to analyze the eye-tracking data, which will allow us to pinpoint the exact time point when participants begin to show evidence of integrating the morphological cue to process actives and passives.

Crucially, the L1 of the L2-English participants invited to this study is Cantonese, which has no verbal morpheme for passives, as illustrated in (5). Cantonese does obligatorily employ the marker *béi* "by" before the agent of an action, in a phrase that comes before the verb, to indicate passive voice (Matthews & Yip, 2011). However, its English equivalent, the *by*-phrase, comes after the verb if it is present. These crosslinguistic differences may lead to different processing patterns between L1-English speakers and L1-Cantonese L2-English speakers in their processing of English passives: Whereas L1-English speakers may be able to employ the morphological cue in the verb to process passives, L1-Cantonese L2-English speakers may either fail to process passives at all or struggle to readily use this cue and wait for the *by*-phrase. Alternatively, these L2-English speakers may be able to utilize the morphological cue within the verb to process passives, like L1-English speakers, overcoming any possible L1 influences. By making use of these crosslinguistic differences, we aim to illuminate whether and when the morphological cue, which is absent in the L2 speakers' L1, is integrated during their L2 processing.

(5) Yíhchìhn gó go beisyū béi gīngléih cháau-jó.
before that CL secretary by manager sack-PFV
"The previous secretary was sacked by the manager."
(Matthews & Yip, 2011, p. 168)

Method

Participants

Thirty L1-English speakers and 43 L1-Cantonese L2-English speakers took part in this study. The study was carried out in accordance with ethical guidelines at the authors' institution, and informed consent was secured from each participant before the study began.

We excluded one L1 speaker and seven L2 speakers due to eye-tracking calibration issues and two further L2 speakers due to their scores falling below the chance level on the fill-in-the-blank task used to screen for offline knowledge (see Section 2.2.2). This procedure led to the exclusion of one L1 speaker and nine L2 speakers in total, resulting in a final sample of 29 L1 speakers and 34 L2 speakers (see Table 1). All L2 participants had Cantonese as their native and dominant language. The average age at which they began learning English was 3.21 years, and they had an average of 1.84 months of immersion experience in an English-dominant country. At the time of testing, they self-estimated their current English exposure to be, on average, 27.82%, which means English constituted roughly one-quarter of their overall language exposure relative to

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press

Table 1. Background information of participants

	L1-English speakers (n = 29)	L1-Cantonese L2-English speakers (n = 34)
Age at testing	25.38 (SD = 9.39)	21.11 (SD = 2.45)
Age of L2 onset	NA	3.21 (SD = 1.39)
Length of residence in an English-dominant country (in months)	NA	1.84 (SD = 3.42)
Current L2 exposure (in percent)	NA	27.82 (SD = 12.45)
Proficiency score (out of 50)	39.28 (SD = 6.30)	34.32 (SD = 5.68)

other languages. Their mean score on a cloze proficiency test was 34.32 out of 50 (see Section 2.3), and a box plot analysis of the proficiency scores did not identify any participant as an outlier.

Materials

Visual-world eye-tracking task

The eye-tracking task contained 12 audio stimuli in four conditions—Control, Active, Passive with an agent, and Passive without an agent—alongside 36 fillers. These 12 items were distributed in a Latin-square design, ensuring that each participant received each item in only one of four conditions (i.e., 4 conditions × 3 items per participant). The current study, however, addresses only the results from two conditions differing in *Voice*: the Active condition and the Passive condition with an agent, as outlined in (6). This means that each participant contributed six stimuli to the analysis —three from the Active condition and three from the Passive condition. Appendix A contains a full list of all critical items included in this study.

(6) Critical conditions in the present study
   a. Active: The cow was washing **and cleaning** the horse last Saturday.
   b. Passive: The cow was washed **and cleaned** by the horse last Saturday.

Twelve verbs (i.e., *chase*, *clean*, *cover*, *hug*, *kick*, *kiss*, *pull*, *punch*, *push*, *protect*, *tickle*, and *wash*) were selected for our stimuli because their meaning denotes actions rather than psychological states, making them suitable for visual representation within the visual-world eye-tracking task.[1] Each stimulus employed two different verbs, drawn from the set of 12, and so each of the verbs appeared twice throughout the entire task (e.g., [6]). The reason for using two different verbs within a single stimulus was to expand the time window of interest within the stimulus. This allowed for a more thorough investigation into time-related changes in eye movements among our participants.

[1]A reviewer points out that the naturalness of passive over active voice for some of these verbs can be influenced by various factors, such as "how important the agent is and how likely the event is to be reported from the perspective of the object." Further research should examine how individual verbs influence the processing of passives.

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press

Because this study is concerned with morphosyntactic-level processing and we did not want to place an undue burden on our L2 participants, all stimulus sentences used simple lexical items. Additionally, we used only third person singular subjects for all sentences, with the auxiliary verb *was* followed by a regular verb with the *-ing* morpheme for the Active condition or the regular *-ed* morpheme for the Passive condition.

The time window of our interest spanned 911 ms on average (SD = 81 ms) from the offset of the first morphological cue that would indicate either an active or passive structure (i.e., *-ing* or *-ed*) to the offset of the second main verb. In terms of duration of this region, the Active condition (M = 892 ms; SD = 76 ms) and the Passive condition (M = 930 ms; SD = 85 ms) were not significantly different in a t-test (p = .980). Each audio stimulus was paired with a visual scene (see Figure 1) showing two images that served as the target and competitor depending on the condition, where the agent and patient roles were interchanged. For example, the target image for the Active condition (6a) was the left picture in Figure 1, with the right picture serving as the competitor; for the Passive condition (6b), these roles were swapped. The areas of interest corresponded to each image in the visual scene, outlined by grey lines in Figure 1.

Fill-in-the-blank task

To ensure our participants possessed the necessary grammatical knowledge of passives, as indicated by above-chance-level accuracy, we administered a separate offline task. In this task, participants filled in blanks to complete sentences by selecting the correct main verb form among four options. While the original offline task included 12 critical items across four conditions, like the eye-tracking task did (see Visual-world eye-tracking task), as well as 12 filler items, the current analysis focuses on the two conditions of the Active and Passive (see Appendix B). Unlike the eye-tracking task, every stimulus in this task contained only one main verb that needed to be filled in.

Procedure

Participants first filled out a language background questionnaire on Google Forms (see Appendix C), when they signed up for the experiment. Approximately one week later, they came to the eye-tracker lab and completed a set of tasks in the following order: a visual-world eye-tracking task, a fill-in-the-blank task, a listening span task, a cloze test

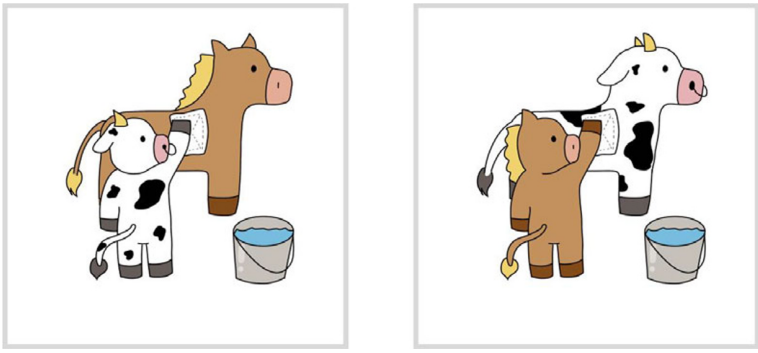


Figure 1. Example visual stimuli in the visual-world eye-tracking task.

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press

(Brown, 1980), and a picture verification task. The eye-tracking task was created using the SR Research Experiment Builder. The cloze test, created on PythonAnywhere (https://www.pythonanywhere.com/), served as an independent measure of language proficiency in this study. This test required participants to fill in 50 gaps within a passage on "man and progress," and their scores, ranging from 0 to 50, reflected the number of accurate responses. All other tasks were designed and presented on PCIbex Farm (https://farm.pcibex.net/). The findings from the listening span and picture verification tasks, which were part of a separate study examining a different phenomenon, will not be addressed in this paper.

In our main visual-world eye-tracking task, participants were instructed to listen to sentences while viewing scenes (e.g., Figure 1) on a computer screen and then select the matching image by clicking on it with a mouse after the offset of each sentence. The hardware setup for this task consisted of the EyeLink 1000 plus system, along with a 19-inch computer screen with a resolution of 1,280 × 1,024 pixels. The system recorded participants' right eye movements at a sampling rate of 1,000 Hz. The task began with an eye calibration procedure using a 13-point array. After completing a practice session with three trials, participants proceeded to the main session. In both practice and main sessions, a fixation cross first appeared for 500 ms at the center of the screen. If any eye-tracking errors were detected at this point, a calibration process was reinitiated. A 2,000-ms preview of the visual scene followed, and then an audio stimulus was played through the computer speaker. Once participants clicked the mouse for a comprehension question for a given trial, the next trial began with a fixation cross. The stimuli were presented in a pseudo-randomized order to avoid the participants encountering the same condition in consecutive trials. The placement of target images within the visual scenes was also counterbalanced to ensure that they appeared equally often on the left and right sides of the screen.

Analysis

We first removed trials that had incorrect mouse-click responses (L1: 2.87%; L2: 0.98%). To analyze the eye-tracking data, we chose 20-ms time bins, which, at 1,000 Hz, provided 20 data points per bin. Using SR Data Viewer's automatic processing, we extracted counts of fixations on the target and competitor within these 20-ms time bins. Next, we transformed them to empirical logits (elogits), using the basic "log" function in R, to address their bounded nature (e.g., Barr, 2008). This transformation was based on the sum of the actual fixation counts per bin, instead of relying on the value 20, as track loss can occasionally reduce the observed fixation counts below 20. Following that, we computed target advantage scores by subtracting the elogit-transformed fixation counts for the competitor from those for the target.

The time window of interest was newly defined for efficient data interpretation, starting at the end of the first *-ing* morpheme (for the Active condition) or *-ed* morpheme (for the Passive condition) at 0 ms and extending to the offset of the second main verb plus an extra 200 ms to account for the typical lag in eye movement (e.g., Matin et al., 1993). In this new time window, 0 ms represents the earliest point where participants could accurately identify the target picture based on the available morphological cue.

We conducted two complementary statistical analyses—a growth curve analysis (e.g., Henry et al., 2022) and a divergence point analysis (e.g., Ito & Knoeferle, 2023; Stone et al., 2021)—in R (R Core Team, 2025). The growth curve analysis enabled us to probe statistical interaction effects between *Group* (L1; L2) and *Condition* (Active;

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press

Passive) over time. The divergence point analysis identified the point at which looks at the target versus the competitor started to diverge, and further indicated whether the onset of this divergence was significantly earlier in one group compared to another. Thus, the two approaches together provided nuanced information impossible to obtain with either method alone.

We first constructed a growth curve analysis model on the target advantage scores within the time window of our interest, including *Group* and *Condition* as fixed effects and *Participant* and *Item* as random effects. The two fixed effects were contrast-coded (−0.5 assigned to L1 and Active; 0.5 assigned to L2 and Passive) and then centered. To effectively model the time-dependent changes in our data, we included both the linear time term (i.e., poly1), which represents the linear trend in fixation counts, and the quadratic time term (i.e., poly2), which captures the acceleration/deceleration of this trend or differences in curvature (e.g., Henry et al., 2022).[2] Our initial approach involved a model with a maximal random-effects structure, but due to a convergence error, we simplified the model by removing the slope of *Group* over *Participant* and the slope of *Group* and *Condition* over *Item* (final model formula: lmer(target_advantage ~ (poly1 + poly2) * Group * Condition + (1 + Condition | participant) + (1 | item)).

To perform the divergence point analysis, we used a non-parametric bootstrapping technique, which allowed us to estimate when a difference in fixation counts across the two groups first appeared and to establish temporal confidence intervals (CIs) around this onset. We repeatedly resampled 2,000 times in total with replacement from our initial dataset to create new datasets (e.g., Ito & Knoeferle, 2023; Stone et al., 2021). Following Stone et al.'s approach, a buffer of 200 ms was appended to the critical time window, giving extra time to detect when divergence potentially occurs beyond the targeted time window. The bootstrapping procedure was applied twice to test differences between the L1 and L2 groups for the Active condition and for the Passive condition.

To pinpoint when fixation patterns began to differ, we specifically applied *t*-tests to fixation counts within every time bin. Following Ito and Knoeferle (2023), divergence was defined as the initial point in time where a minimum of four consecutive time points showed statistically significant results. We then compiled a distribution representing how the divergence time points varied between the two groups for each condition. When the 95% CIs for this distribution do not contain zero, it indicates a statistically reliable group difference.

Results

Both groups of participants achieved high response accuracy in the fill-in-the-blank task (L1: $M = 97.13\%$, $SD = 16.75$; L2: $M = 88.73\%$, $SD = 31.71\%$) and high mouse-click accuracy in the eye-tracking task (L1: $M = 97.13\%$, $SD = 16.75$; L2: $M = 99.02\%$, $SD = 9.88\%$). These results suggest that our participants had intact knowledge of actives and passives, and that they were able to comprehend the given sentences.

Figure 2 displays the elogit-transformed fixations on the target and competitor, broken down by group and condition. These fixations were calculated for 20-ms time bins, with each bin aggregating 20 data points, given a sampling rate of 1,000 Hz.

[2]To address a reviewer's concern about the inclusion of a quadratic term in our analysis, we ran a likelihood ratio test to compare the full linear growth curve model, including both linear and quadratic terms, to the reduced model without the quadratic term. The fuller model provided a significantly better fit to the data: $\chi^2(6) = 96.65$, $p < .001$.

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press

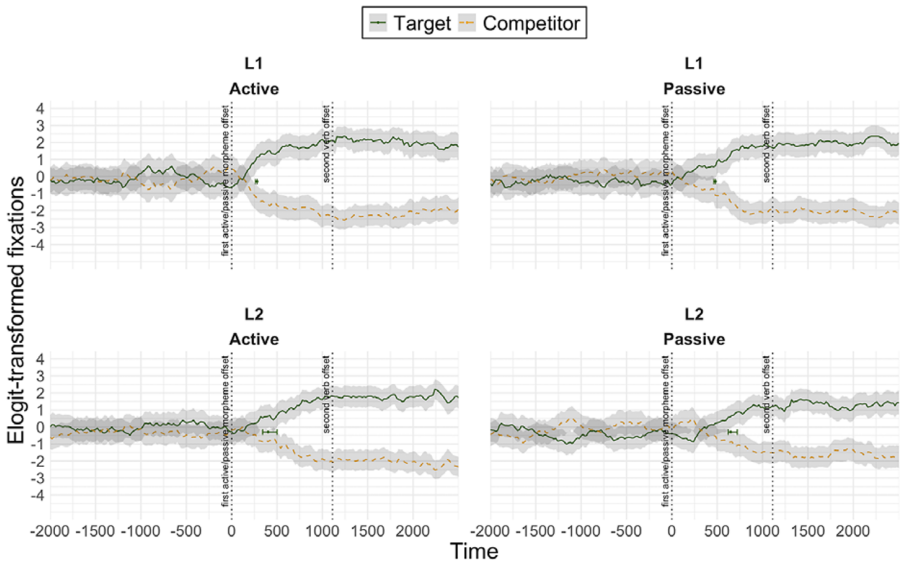


Figure 2. Empirical-logit transformed fixations on the target and competitor by group and condition. *Notes*: The shaded areas indicate 95% confidence intervals. The dotted vertical lines show the critical time window, which extends from 0 ms to 1,111 ms. The points and error bars represent bootstrap means and 95% confidence intervals, which were estimated based on the divergence point analysis.

Table 2. Output from growth curve analysis model

	b	*SE*	*t*	*p*
(Intercept)	1.90	0.36	5.27	< .001
Linear term	**10.12**	**0.22**	**46.54**	**< .001**
Quadratic term	**−1.38**	**0.22**	**−6.37**	**< .001**
Group	**−1.08**	**0.51**	**−2.13**	**.037**
Linear term	−0.76	0.44	−1.75	.080
Quadratic term	**2.33**	**0.44**	**5.34**	**< .001**
Condition	−0.86	0.67	−1.29	.206
Linear term	0.15	0.44	0.35	.725
Quadratic term	**1.74**	**0.43**	**3.99**	**< .001**
Group × Condition	−0.27	0.86	−0.32	.751
Linear term	−0.14	0.87	−0.16	.872
Quadratic term	**−2.90**	**0.87**	**−3.32**	**< .001**

Notes: Significant effects are shown in bold; R^2_m = .052; R^2_c = .235.

The growth curve analysis of the target advantage scores during the time window of interest showed a significant effect of *Group* on both the intercept (p = .037) and the quadratic term (p < .001), as summarized in Table 2. The analysis also found a significant effect of *Condition* and its interaction with *Group*, but only on the quadratic term (all ps < .001). However, these effects, in the absence of main effects or effects observed on the linear term, are difficult to interpret because they may reflect less relevant variance.[3] Overall, these results suggest that the processing patterns of the L1-English speakers and L1-Cantonese L2-English speakers differed.

[3]We thank the associate editor for bringing our attention to this issue.

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press

The divergence point analysis revealed that the onset of fixation differences between the target and competitor for the Active condition was 278.34 ms (95% CI = [260, 280]) in the L1 speakers and 400.46 ms (95% CI = [340, 500]) in the L2 speakers; and for the Passive condition, the onset of fixation differences between the target and competitor was 478.80 ms (95% CI = [460, 480]) in the L1 speakers and 647.55 ms (95% CI = [620, 720]) in the L2 speakers (see Figure 2). These observed differences were statistically reliable, with the L2 speakers showing a delay of 122.12 ms (95% CI = [60, 240]) in the Active condition and 168.75 ms (95% CI = [140, 240]) in the Passive condition.

Discussion

Regarding RQ 1, the L1-English speakers and the L1-Cantonese L2-English speakers showed similar processing patterns, with both groups demonstrating the ability to employ a morphological cue within the verb to process both active and passive structures in English. For RQ 2, however, these two groups presented timing differences (see below). This section discusses our results in light of the crosslinguistic differences, the possible effects of the age of L2 acquisition onset, and the temporal dynamics of morphological cue integration.

The L2 speakers' ability to process passives within the verb region is noteworthy because it is not attributable to processing routines developed from their L1 Cantonese, which is markedly different from English. Cantonese lacks morphological affixes in general and employs the obligatory *béi* marker to introduce the agent of a passive structure, whereas English employs several inflectional affixes, including the past participle, which signals a passive structure in addition to the optional *by*-phrase. Despite these crosslinguistic differences, our L1-Cantonese L2-English speakers were able to process passives early on within the verb region, before encountering the preposition *by*.[4] This is important because *by* might have been a more reliable cue for them, given the presence of its equivalent in Cantonese. The finding that the L1-Cantonese L2-English speakers employed morphological information unavailable in their L1 during real-time L2 processing, therefore, suggests that the underlying mechanisms driving L1 and L2 processing are similar (see also Cunnings, 2017; Fernandez et al., 2018).

It should be noted that the L2 participants in this study began learning English at a young age (M = 3.21; SD = 1.39), as an anonymous reviewer pointed out. This age of onset is comparable to that of the child L1-Turkish L2-English speakers in Marinis and Saddy's (2013) study, who began learning their L2 at the mean age of 3.23. The participants' early age of onset in both studies could have facilitated the efficient use of morphological information in their L2 processing of passives. If this is the case, adult L1-Cantonese L2-English speakers with a later age of onset, such as 10 years old, might not be able to use a morphological cue within the verb to process passives, which warrants further investigation.

When it comes to the timing of the morphological cue integration, the L2-English speakers in this study showed slower processing than the L1-English speakers, the result that is in line with that in Crossley et al. (2020) and Marinis and Saddy (2013). The divergence point analysis revealed that, compared to the L1 group, the L2 group's

[4]All L2 participants also spoke Mandarin. Since Mandarin's passives are similar to Cantonese's, except for the optionality of the agent phrase in Mandarin, we hypothesize that their Mandarin knowledge did not impact our results. Future work should test this hypothesis.

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press

integration of morphemes was delayed by 122.12 ms for actives and by 168.75 ms for passives. By identifying differences in the onset time of morphological cue integration between the two groups, our approach to further examining the *Group* effect found from the growth curve analysis supports the recent emphasis on understanding when and what information L2 speakers utilize (e.g., Felser, 2019). Using this approach, importantly, we also revealed that although the L2 speakers were slower than the L1 speakers, they were quick enough to process both actives and passives within the verb region.

In our view, the delayed L2 processing is attributable to cognitive and psycholinguistic factors, such as lexical retrieval difficulties (Hwang & Kim, 2025), lower proficiency levels (Cunnings, 2017), or limitations of working memory capacity or cognitive resources in general (Kaan et al., 2015). While we designed our experimental items to be simple and easy to understand, it is possible that our L2 participants still experienced lexical competition between their L1 and L2. Additionally, the L2 speakers in this study had significantly lower proficiency scores than the L1 speakers ($t = -3.28$, $p < .001$, Cohen's $d = 0.83$), and all had learned English mostly in a non-English-dominant region and all retained Cantonese as their dominant language. Any or all of these factors may have led to delayed processing compared to the L1 speakers. Also, an anonymous reviewer observed that the L2 speakers exhibited a relatively wide range in the duration of their stays in an English-dominant country, with an average of 1.84 months and an *SD* of 3.42 months (see Table 1). Hence, it is conceivable that L2 speakers with very high proficiency, greater naturalistic L2 input in an English-dominant country, English dominance, and/or excellent working memory could display target-like processing patterns in terms of speed. In future work, we plan to measure these factors and investigate their role in L2 processing with a larger population.

On the other hand, the offline results from the visual-world eye-tracking task and the fill-in-the-blank task revealed an interesting difference. The accuracy of the L2 speakers was higher in the visual-world eye-tracking task ($M = 99.02\%$; $SD = 9.88\%$) than in the fill-in-the-blank task ($M = 88.73\%$; $SD = 31.71\%$), which might be due to the different nature of comprehension and production.[5] Comprehension is a data-driven process, where speakers assign meaning to linguistic forms, while production relies on self-activated mechanisms to transform pre-verbal intentions into spoken/written language through grammatical encoding processes (e.g., Alonso, 2016). This distinction supports the common understanding that comprehension comes before production in language development and that individuals typically understand more than they are able to express (Clark, 1993). It is thus likely that the L2 speakers in this study had not yet automated their production to the same degree as their comprehension or as L1 speakers.

One limitation of this study is its reliance on a small number of items, which may impact the generalizability of the results; future research is thus needed to build on and extend our findings. Another limitation concerns the task order. While administering the eye-tracking task first allowed us to isolate online processing without interference from the fill-in-the-blank task, it may have inadvertently introduced learning or priming effects that influenced participants' responses to the fill-in-the-blank task. Future studies should insert a one-week interval between the tasks to minimize such effects.

[5]We thank the reviewer for bringing our attention to this issue.

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press

In conclusion, this study makes both theoretical and methodological contributions. Theoretically, its findings offer insights into the underlying mechanisms involved in L2 speakers' real-time processing of complex morphosyntactic structures, particularly in terms of the temporal dynamics of morphological information integration. Methodologically, employing both growth curve analysis and divergence point analysis, rather than using either method in isolation, demonstrates a useful approach that enables meticulous analysis of time-series data in language processing research.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0272263125101472.

Competing interests. The authors declare none.

Data availability statement. The experiment in this article earned Open Data and Open Materials badges for transparent practices. All materials, data, and analysis scripts are available at https://osf.io/7fpex.

References

Alonso, R. A. (2016). *Crosslinguistic influence in second language acquisition*. Multilingual Matters. https://doi.org/10.2307/jj.27080052

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*, 457–474. https://doi.org/10.1016/j.jml.2007.09.002

Bayram, F., Rothman, J., Iverson, M., Kupisch, T., Miller, D., Puig-Mayenco, E., & Westergaard, M. (2019). Differences in use without deficiencies in competence: Passives in the Turkish and German of Turkish heritage speakers in Germany. *International Journal of Bilingual Education and Bilingualism*, *8*, 919–939. https://doi.org/10.1080/13670050.2017.1324403

Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *The Modern Language Journal*, *64*, 311–317. https://doi.org/10.2307/324497

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3–42. https://doi.org/10.1017/S0142716406060024

Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, *40*, 693–706. https://doi.org/10.1017/S0272263117000250

Clark, E. C. (1993). *The lexicon in acquisition*. Cambridge University Press. https://doi.org/10.1017/CBO9780511554377

Crossley, S., Duran, N. D., Kim, Y., Lester, T., & Clark, S. (2020). The action dynamics of native and non-native speakers of English in processing active and passive sentences. *Linguistic Approaches to Bilingualism*, *10*, 58–85. https://doi.org/10.1075/lab.17028.cro

Cunnings, I. (2017). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, *20*, 659–678. https://doi.org/10.1017/S1366728916000675

Felser, C. (2019). Structure-sensitive constraints in non-native sentence processing. *Journal of the European Second Language Association*, *3*, 12–22.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164–203. https://doi.org/10.1016/S0010-0285(03)00005-7

Fernandez, L., Höhle, B., Brock, J., & Nickels, L. (2018). Investigating auditory processing of syntactic gaps with L2 speakers using pupillometry. *Second Language Research*, *34*, 201–227. https://doi.org/10.1177/02676583177223

Henry, N., Jackson, C. N., & Hopp, H. (2022). Cue coalitions and additivity in predictive processing: The interaction between case and prosody in L2 German. *Second Language Research*, *38*, 397–422. https://doi.org/10.1177/0267658320963151

Hwang, H.,& Kim, K. (2025). Effects of lexical frequency in predictive processing: Higher frequency boosts first language speed and facilitates second language prediction. Language Learning. Advance online publication. https://doi.org/10.1111/lang.12718

Ito, A., & Knoeferle, P. (2023). Analysing data from the psycholinguistic visual-world paradigm: Comparison of different analysis methods. *Behavior Research Methods*, *55*, 3461–3493. https://doi.org/10.3758/s13428-022-01969-3

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press

Kaan, E., Ballantyne, J., & Wijnen, F. (2015). Effects of reading speed on second-language sentence processing. *Applied Psycholinguistics*, *36*, 799–830. https://doi.org/10.1017/S0142716413000519

Lee, J. F., & Doherty, S. (2019). Native and nonnative processing of active and passive sentences: The effects of processing instruction on the allocation of visual attention. *Studies in Second Language Acquisition*, *41*, 853–879. https://doi.org/10.1017/S027226311800027X

Marinis, T., & Saddy, D. (2013). Parsing the passive: comparing children with specific language impairment to sequential bilingual children. *Language Acquisition*, *20*, 155–179. https://doi.org/10.1080/10489223.2013.766743

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, *53*, 372–380. https://doi.org/10.3758/BF03206780

Matthews, S., & Yip, V. (2011). *Cantonese: A comprehensive grammar*. Routledge. https://doi.org/10.4324/9780203835012

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*, 381–401. https://doi.org/10.1016/j.jml.2006.06.006

Puebla, C., & Felser, C. (2024). Discourse-based pronoun resolution in non-native sentence processing. *Bilingualism: Language and Cognition*, *27*, 557–571. https://doi.org/10.1017/S1366728923000676

R Core Team. (2025). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. Available at https://www.R-project.org/

Stone, K., Lago, S., & Schad, D. J. (2021). Divergence point analyses of visual world data: Applications to bilingual research. *Bilingualism: Language and Cognition*, *24*, 833–841. https://doi.org/10.1017/S1366728920000607

Xiao, R., McEnery, T., & Qian, Y. (2006). Passive constructions in English and Chinese: A corpus-based contrastive study. *Languages in Contrast*, *6*, 109–149. https://doi-org.easyaccess2.lib.cuhk.edu.hk/10.1075/lic.6.1.05xia

Cite this article: Hwang, H., & Han, G. (2026). Processing of passives using morphological information within the verb in a second language: Slower, but as robust as in a first language. *Studies in Second Language Acquisition,* 48: 478–490. https://doi.org/10.1017/S0272263125101472

https://doi.org/10.1017/S0272263125101472 Published online by Cambridge University Press