Japanese Psychological Research
THE JAPANESE PSYCHOLOGICAL ASSOCIATION
Check for updates
Japanese Psychological Research
2026
doi: 10.1111/jpr.70035
Brief Report

Anticipating at a Distance: Classifier–Noun Agreement Processing in Japanese by Native Speakers and Chinese JFL Learners

WEIYU WANG *Shanghai University*

JINGYI ZHANG *Setsunan University*

YING SU and KATSUO TAMAOKA* *Shanghai University*

Abstract: This study examined how native Japanese speakers and Chinese learners of Japanese as a foreign language (JFL) process classifier–noun agreement involving short- and long-distance dependencies. Using a maze task, participants read sentences in which classifiers agree locally either with the subject noun or with a distant object noun. Native speakers showed faster reaction times for short-distance dependencies and a significant slowdown at the verb region under long-distance conditions, but no difference at the head noun, indicating anticipatory processing using verb cues. In contrast, Chinese JFL learners showed slower processing at both the classifier and following noun regions, but faster responses at the head noun in the long-distance condition. This suggests a bottom-up, non-predictive strategy, with resolution occurring only when the head noun is encountered. These findings highlight a key difference in real-time sentence processing: Native speakers engage in predictive integration, while JFL learners rely more on reactive parsing.

Key words: anticipatory processing, classifier–noun agreement, Japanese as a foreign language (JFL), L1 transfer, long-distance dependency.

Numeral classifiers are a grammatical feature shared by both Japanese and Mandarin Chinese (hereafter, Chinese; Allan, 1977). In both languages, classifiers appear between numerals and nouns, yielding structurally similar expressions, for example, *huta-ri-no ryôrinin* (二人の料理人, "two cooks") and *san-satsu-no hon* (三冊の本, "three books") in Japanese, and *liǎng wèi chúshī* (两位厨师, "two cooks") and *sān běn shū* (三本书, "three books") in Chinese. Apart from the Japanese possessive particle *no*, both languages share the [Number + Classifier + Noun] structure. This typological similarity suggests that Chinese learners of Japanese as a foreign language (JFL) may transfer L1-based syntactic strategies when processing Japanese classifier constructions. Such transfer may support accurate processing in simple, linear contexts. Yet, these long-distance dependencies may pose challenges for Chinese JFL learners to resolve them effectively. In such cases, learners may resort to

*Correspondence concerning this article should be sent to: Katsuo Tamaoka, School of Foreign Languages, Shanghai University, 99 Shangda Road, Baoshan District, Shanghai 200444, China. (E-mail: ktamaoka@gc4.so-net.ne.jp)

surface-level heuristics rather than native-like incremental parsing. Thus, this study examines how Chinese JFL learners process classifier–noun agreement, focusing on the role of syntactic distance and the influence of foreign language (FL) proficiency.

Short- and Long-Distance Classifier–Noun Agreement

The processing of long-distance syntactic dependencies by second-language (L2) learners has been a central topic in psycholinguistic research. Empirical studies have investigated whether and how L2 learners differ from native speakers in their ability to resolve long-distance structures, such as *wh*-movement (Hettiarachchi & Pires, 2022; Juffs, 2005; Pliatsikas et al., 2017; Slavkov, 2015), subject- and object-extracted relative clauses (Juffs & Harrington, 1995; Kanno, 2007), and island constraints (Boxell & Felser, 2017; Omaki & Schulz, 2011). A general finding across these studies is that L2 learners exhibit reduced sensitivity to syntactic information during real-time processing of long-distance dependencies, likely due to the increased cognitive demands associated with syntactic integration over distance. Importantly, this sensitivity is modulated by individual factors, such as L2 proficiency and the degree of syntactic similarity between the learner's first language (L1) and the target language.

Theoretical explanations have sought to explain these findings. Among them, the Shallow Structure Hypothesis (SSH; Clahsen & Felser, 2006) argues that L2 learners rely less on detailed syntactic representations and instead depend more on semantic and lexical cues, particularly when processing complex or nonlocal dependencies. Classifier constructions present an interesting case for examining such processing mechanisms. Classifiers encode semantic features of the noun they modify, including animacy, shape, size, and function (Allan, 1977; Croft, 1994; Huang & Schiller, 2021; Lakoff, 1986), and thus play a central role in guiding predictive processing during sentence comprehension. When a noun mismatches the semantic expectations set by a preceding classifier, processing difficulty arises. This phenomenon is known as the *classifier–noun agreement effect* (Frankowsky et al., 2022; Her et al., 2018; Huettig et al., 2010).

This effect has been observed in both native and L2 speakers. However, L2 learners often show reduced or delayed sensitivity to classifier cues, particularly when their L1 does not include a classifier system (e.g., English or Russian). Studies have shown that such learners rely less on classifiers to predict or integrate upcoming nouns (Grüter et al., 2020; Lau & Grüter, 2015; Mitsugi, 2020) or may use classifiers in qualitatively different ways (Zhu, 2021). In contrast, learners whose L1 includes classifiers, such as Chinese, may show more native-like sensitivity, though it remains unclear whether this extends to structurally complex contexts, such as long-distance dependencies.

In both Japanese and Chinese, classifier–noun agreement typically follows the linear order of [Number + Classifier + Noun], allowing the classifier to semantically constrain or agree with the noun it modifies. However, within object-extracted relative clause structures, this relationship can be disrupted by intervening constituents, giving rise to either short-distance or long-distance classifier–noun agreement configurations.

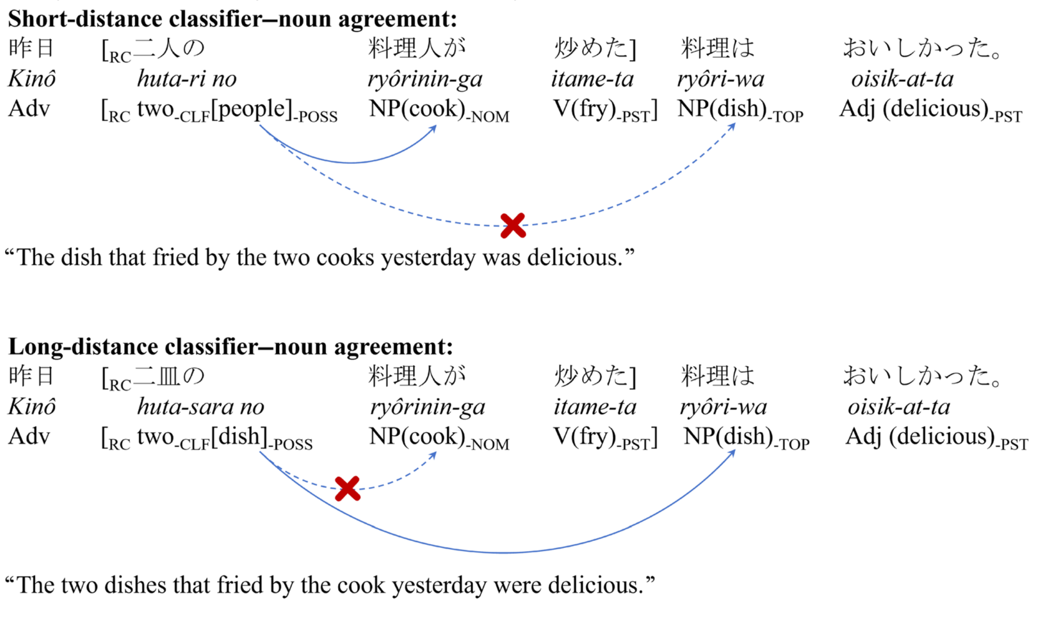
As illustrated in Figure 1, short-distance agreement occurs when the classifier directly precedes and modifies the subject noun within the relative clause. For example, in the phrase *huta-ri-no ryôrinin* (二人の料理人, "two cooks"), the classifier *huta-ri-no* ("two [people]") immediately modifies *ryôrinin* ("cook"), forming a local and structurally straightforward dependency. This proximity facilitates processing, as the classifier and its associated noun are both syntactically and linearly adjacent. In contrast, long-distance agreement arises when the classifier appears before the relative clause and semantically modifies the head noun of the clause (i.e., the object), which occurs later in the sentence. In the expression *huta-sara-no ryôrinin-ga itame-ta ryôri-wa* (二皿の料理人が炒めた料理, "the two dishes fried by the cook"), the classifier *huta-sara no* ("two [dish]") must be interpreted in relation to *ryôri* ("dish"), bypassing the subject noun *ryôrinin* ("cook") and the verb *itame-ta* ("fried"). Evidence from eye-tracking (F. Wu et al., 2014) and ERP studies (Hsu et al., 2014; Yasunaga & Sakamoto, 2007) suggest that long-distance dependencies impose greater demands on working memory and syntactic integration, particularly during real-time processing. These dependencies involve syntactic elements, such as classifiers and their associated nouns that are not adjacent in the linear structure of the sentence, often separated by intervening clauses or phrases. Such non-adjacent relationships pose significant challenges for L2 learners, who frequently struggle to maintain and integrate these distant elements during comprehension.

In sum, the key distinction lies in dependency length: Short-distance agreement involves local, easily processed relationships, whereas long-distance agreement requires the parser to resolve dependencies across multiple intervening elements. This structural contrast forms the basis for investigating how native Japanese speakers and Chinese JFL learners process classifier–noun agreement under differing syntactic conditions.

The Present Study

As discussed in previous sections, classifier–noun agreement in Japanese involves both short- and long-distance dependencies, which vary in processing complexity. While structural similarity between Chinese and Japanese may support Chinese JFL learners in processing simple classifier constructions, it remains unclear whether this advantage extends to more complex, long-distance configurations. One possibility is that learners rely on local heuristic cues, rather than engaging in full syntactic integration. Alternatively, with increased proficiency, they may adopt more native-like, incremental parsing strategies. To explore this, this study compares native Japanese speakers and Chinese JFL learners in their processing of classifier–noun agreement structures. By examining performance across short- and long-distance

Figure 1
Examples of short- and long-distance classifier–noun agreement.



agreement conditions, the study aimed to clarify whether Chinese learners rely on L1-based strategies or whether their processing reflects proficiency-driven development.

Methods

Participants

A total of 36 Japanese native speakers and 45 Chinese JFL learners participated in the experiment. One Japanese native speaker and one Chinese JFL learner were excluded from analysis due to low overall accuracy (below 70%). The final sample included 35 Japanese native speakers (Japanese native speaker group; M_{age} = 18 years and 10 months, SD_{age} = 6 months, $Range_{age}$ = 18 years–19 years and 10 months, 24 females) and 44 Chinese JFL learners (Chinese JFL learner group; M_{age} = 21 years and 11 months, SD_{age} = 27 months, $Range_{age}$ = 19 years and 7 months–25 years and 11 months, 28 females). Japanese proficiency was assessed using a 36-item multiple-choice grammar test (Max_{score} = 36). Test items were randomly selected from the grammar Simple Performance-Oriented Test (SPOT; Kobayashi et al., 1996) of the Tsukuba Test-Battery of Japanese. This experiment involving human participants was reviewed and approved by the Research Ethics Committee at Shanghai University. Before the experiment, all participants provided written informed consent and received compensation for their participation.

Materials

Thirty-six sentences containing classifier–noun agreement structures were constructed. Each sentence consisted of six constituent regions: a time phrase, a classifier, a first noun, a verb, a second noun, and a sentence-final phrase. The time phrase "yesterday" (*kinô*) is treated as a time adverb in Japanese, although it corresponds to a time noun in Chinese. Two experimental conditions were created by manipulating the classifier to yield either short- or long-distance classifier–noun agreement. In the short-distance condition (Example 1a), the classifier agrees with the subject noun within the relative clause. In the long-distance condition (Example 1b), the classifier instead agrees with the head noun of the relative clause (i.e., the object).

Example 1a. Short-distance classifier–noun agreement.

昨日 二人の 料理人が 炒めた 料理は おいしかった。

Kinô huta-ri-no ryôrinin-ga itame-ta ryôri-wa oisik-at-ta.

[$_{CP}$ Adv two$_{\text{-CLF-POSS}}$ NP(cook)$_{\text{-NOM}}$ V(fry)$_{\text{-PST}}$] NP(dish)$_{\text{-TOP}}$ Adj(delicious)$_{\text{-PST}}$

"The dish that the two cooks fried yesterday was delicious."

Example 1b. Long-distance classifier–noun agreement

昨日 二皿の 料理人が 炒めた 料理は おいしかった。

Kinô huta-sara-no ryôrinin-ga itame-ta ryôri-wa oisik-at-ta.

[$_{CP}$ Adv two$_{\text{-CLF-POSS}}$ NP(cook)$_{\text{-NOM}}$ V(fry)$_{\text{-PST}}$] NP(dish)$_{\text{-TOP}}$ Adj(delicious)$_{\text{-PST}}$

"The two dishes that the cook fried yesterday were delicious."

In all conditions, the numeral was fixed as *two*, appearing immediately before the classifier. As a confusing factor, animacy of nouns plays a rule in the processing. To avoid potential animacy effect (e.g., Mak et al., 2002; Traxler et al., 2005), the subject nouns were kept as people and the head nouns as inanimate entities. To prevent any single participant from encountering both the short- and long-distance versions of the same sentence, a Latin square design was employed to counterbalance the stimulus distribution across participants.

Procedure

The experiment was conducted using a maze task (Forster, 2010; Forster et al., 2009) to measure online processing. In the lexical maze task, participants are presented with two candidate continuations at each step and must select the acceptable option. There are generally two types of maze paradigms: grammatical maze tasks and lexical maze tasks. In grammatical maze tasks, participants choose between two grammatically plausible continuations, whereas in lexical maze tasks, they choose between a valid lexical item and a nonword. Because Japanese allows relatively flexible word order, it is difficult to construct reliable grammatical maze contrasts. For this reason, the lexical maze paradigm was applied. Although each choice constitutes a local lexical decision, the forced sequence of decisions incrementally builds a sentence. If a participant makes an incorrect choice at any step, the entire trial is immediately terminated and coded as incorrect. Consequently, successful completion of a trial necessarily reflects incremental sentence construction rather than isolated selection of visually plausible words.

Stimulus presentation and data collection were implemented using the E-Prime 3.0 platform. Responses were recorded using a Chronos Device. Each trial began with a slide for region 1: a fixation point ("++++"), followed by the simultaneous presentation of a time phrase, such as *kinô* ("Yesterday") on both the left and right sides of the screen. Participants were instructed to choose the time phrase side. After this, participants proceeded through five slides for each of the five regions, each presenting a pair of a word and a nonword with the same case marker (e.g., 料理人が and ね木こが) or the same suffix (e.g., おいしかった and そ花かった). In psycholinguistic research, such cumulative forced-choice processing is regarded as capturing relatively automatic, online sentence processing with minimal reliance on task-specific strategies (Forster, 2010; Forster et al., 2009; Tamaoka et al., 2024; J. Witzel & Witzel, 2016; N. Witzel et al., 2012). Because task performance is fully determined by the correctness of each incremental selection, no additional comprehension questions are required. To ensure that participants understood the procedure, each completed six practice trials before the main experiment.

Data Analysis

Accuracy and reaction time data from the maze task were analyzed using (generalized) linear mixed-effects (GLME/LME) models (Baayen et al., 2008), implemented in R (Version 4.4.2, R Core Team, 2024). For the native Japanese speaker group, fixed effects included condition (factor variable; short-distance vs. long-distance; long-distance as the reference level) and trial order (numerical variable). For the learner group, participants' Japanese proficiency (numeric variable) was also included, along with the Condition × Proficiency interaction. All numeric variables standardized (z-scores). Random intercepts and slopes were included for both participants and items.

The reaction time analysis was restricted to correctly answered trials and was conducted by region. Based on the Box–Cox power transformation technique (Box & Cox, 1964; Venables & Ripley, 2002), reaction times were transformed to reduce skewness using either a natural logarithm (logRT) or a negative reciprocal transformation (–1,000/RT). Degrees of freedom and p-values were estimated using Satterthwaite's approximations (Satterthwaite, 1946) with restricted maximum likelihood (Harville, 1977). To control the effect of temporal dependency (Baayen & Milin, 2010; S. Wu et al., 2023), the reaction times of previous region(s) were included as a co-variable. However, due to potential multicollinearity, a principal component analysis (PCA) was conducted on reaction times from Regions 3 to 6, and the first principal component (PC) was included in the models.

Model selection was based on the Akaike Information Criterion (Anderson et al., 2000) using the *buildmer* package (Version 2.11, Voeten, 2023). A maximal converging model (Barr et al., 2013) was first constructed. To balance Type I error and statistical power (Matuschek et al., 2017), non- or low-contributory predictors were iteratively removed, consequently getting the best-fitting models. In this selection process, the fixed effects of condition and Japanese proficiency, along with their interaction, were retained. After model selection, potentially influential outliers with absolute standardized residuals exceeding 2.5 standard deviations were removed based on the best-fitting models.

Results of Native Japanese Speaker Group

Results of GLME Analysis for Accuracy Rates

The best-fitting GLME model was specified as ACC ~ 1 + Condition + Trial order + (1 | Participant). The effect of condition did not reach significance level ($\beta = 0.91$, $SE = 0.56$, $z = 1.64$, $p = .101$), suggesting that processing accuracy of sentences with short-distance classifier–noun agreements ($M = 98.41\%$, $SD = 12.52\%$) was consistent with sentences with long-distance classifier–noun agreements ($M = 96.51\%$, $SD = 18.39\%$). Additionally, the trial order did not significantly affect accuracy rates ($\beta = 0.53$, $SE = 0.28$, $z = 1.90$, $p = .057$).

Results of LME Analysis for Reaction Times

LME analysis was applied to analyze reaction times for each sentence region. The mean reaction times, standard errors, and the statistical significance of condition effects across regions are summarized in Figure 2.

Significant effects of condition were observed in Regions 2, 3, and 4. In Region 2 (classifier phrases, e.g., 二人の vs. 二皿の), the processing speed for the short-distance condition (二人の, $M = 715$ ms, $SD = 160$ ms) was significantly faster ($\beta = -0.11$, $SE = 0.02$, $t(516.48) = -6.70$, $p < .001$) than that for the long-distance condition (二皿の, $M = 733$ ms, $SD = 198$ ms). In Region 3 (nouns, e.g., 料理人が), consistent with the findings from Region 2, nouns were processed significantly faster ($\beta = -0.17$, $SE = 0.03$, $t(32.75) = -6.05$, $p < .001$) under the short-distance condition ($M = 721$ ms, $SD = 142$ ms) than under the long-distance condition ($M = 887$ ms, $SD = 314$ ms). In Region 4 (verbs, e.g., 炒めた), verbs in the short-distance condition ($M = 747$ ms, $SD = 142$ ms) were again processed significantly faster ($\beta = -0.03$, $SE = 0.01$, $t(518.06) = -2.57$, $p < .05$) than those in the long-distance condition ($M = 887$ ms, $SD = 314$ ms). In contrast, no significant condition effects were found in Regions 1, 5, or 6 (all ps > .05).

Results of Chinese JFL Learner Group

Results of GLME Analysis for Accuracy Rates

The final LME model was specified as ACC ~ 1 + Condition + Proficiency + Trial order + Condition:Proficiency + (1 | Participant) + (1 | Item). The effect of condition was significant ($\beta = 1.82$, $SE = 0.40$, $z = 4.52$, $p < .001$), indicating that sentences with short-distance classifier–noun agreement ($M = 96.46\%$, $SD = 18.49\%$) were processed more accurately than those with long-distance classifier–noun agreement ($M = 87.88\%$, $SD = 32.68\%$). Japanese proficiency also significantly contributed to accuracy ($\beta = 0.56$, $SE = 0.28$, $z = 1.98$, $p < .05$). In contrast, the interaction between condition and proficiency was not significant ($\beta = -0.29$, $SE = 0.37$, $z = -0.79$, $p = .432$).

Results of LME Analysis for Reaction Times

LME analysis was applied to analyze reaction times for each sentence region. The mean reaction times, standard errors, and the statistical significance of condition effects across regions are summarized in Figure 3. Significant condition effects were observed in Regions 2, 3, and 5.

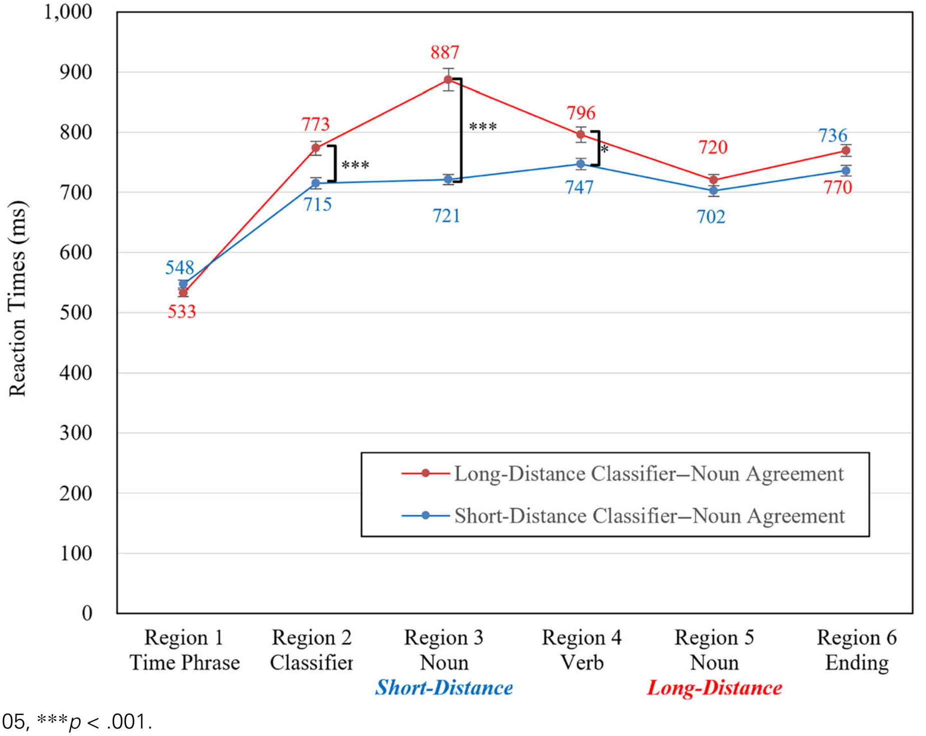
Table 1 summarizes the results of the LME analysis in Region 2. In Region 2 (classifiers, e.g., 二人の vs. 二皿の), the long-distance condition (e.g., 二皿の; $M = 1{,}277$ ms, $SD = 613$ ms) elicited significantly longer reaction times ($\beta = -0.17$, $SE = 0.02$, $t(623.25) = -8.96$, $p < .001$) than the short-distance condition (e.g., 二人の; $M = 1{,}075$ ms, $SD = 488$ ms). Notably, condition significantly interacted with

Japanese proficiency ($\beta = -0.04$, $SE = 0.02$, $t(622.65) = -2.21$, $p < .05$). To further illustrate this interaction, Figure 4 presents a Condition × Proficiency interaction plot showing the predicted values of negatively reciprocally transformed reaction times, along with 95% prediction intervals. As shown in Figure 4, increased Japanese proficiency facilitated faster processing, particularly in the long-distance condition. However, the main effect of Japanese proficiency was not significant ($\beta = -0.04$, $SE = 0.03$, $t(46.17) = -1.08$, $p = .286$). In addition, the effects of trial order ($\beta = -0.12$, $SE = 0.02$, $t(43.42) = -7.77$, $p < .001$) and reaction times of Region 1 ($\beta = 0.06$, $SE = 0.01$, $t(603.55) = 4.54$, $p < .001$) reached significance.

In Region 3 (nouns, e.g., 料理人が), the results of analysis are reported in Table 2. Reaction times were significantly longer ($\beta = -0.13$, $SE = 0.02$, $t(24.97) = -5.39$, $p < .001$) in the long-distance condition ($M = 1{,}686$ ms, $SD = 1{,}273$ ms) than in the short-distance condition ($M = 1{,}341$ ms, $SD = 783$ ms). A significant interaction between condition and Japanese proficiency was also observed ($\beta = -0.05$, $SE = 0.02$, $t(617.80) = -2.66$, $p < .01$). As shown in Figure 5, lower-proficiency Chinese JFL learners exhibited generally slower processing. In contrast, higher-proficiency learners processed short-distance agreement more efficiently than long-distance agreement. However, consistent with the findings in Region 2, the main effect of Japanese proficiency was not significant ($\beta = -0.04$, $SE = 0.02$, $t(28.07) = -1.54$, $p = .136$). Additionally, the effects of trial order ($\beta = -0.08$, $SE = 0.01$, $t(613.54) = -8.41$, $p < .001$) and PC ($\beta = 0.06$, $SE = 0.01$, $t(20.64) = 4.26$, $p < .001$) were both significant.

Region 5 (nouns, e.g., 料理は) exhibited a reversed condition effect. As shown in Table 3, the significant effect of condition ($\beta = 0.06$, $SE = 0.02$, $t(643.59) = 3.45$, $p < .001$) indicated that the long-distance condition ($M = 1{,}178$ ms,

Figure 2
Reaction times for each region in Japanese native speaker group.



*p < .05, ***p < .001.

SD = 531 ms) elicited shorter reaction times than the short-distance condition ($M = 1{,}301$ ms, $SD = 854$ ms). Additionally, a significant interaction between condition and proficiency was observed ($\beta = -0.03$, $SE = 0.02$, $t(635.77) = -2.03$, $p < .05$). As illustrated in the Condition × Proficiency interaction plot (Figure 6), the slope pattern closely resembled that observed in Region 3: Lower-proficiency Chinese JFL learners exhibited generally slower responses, particularly under the long-distance condition, whereas higher-proficiency learners processed short-distance classifier–noun agreement more efficiently.

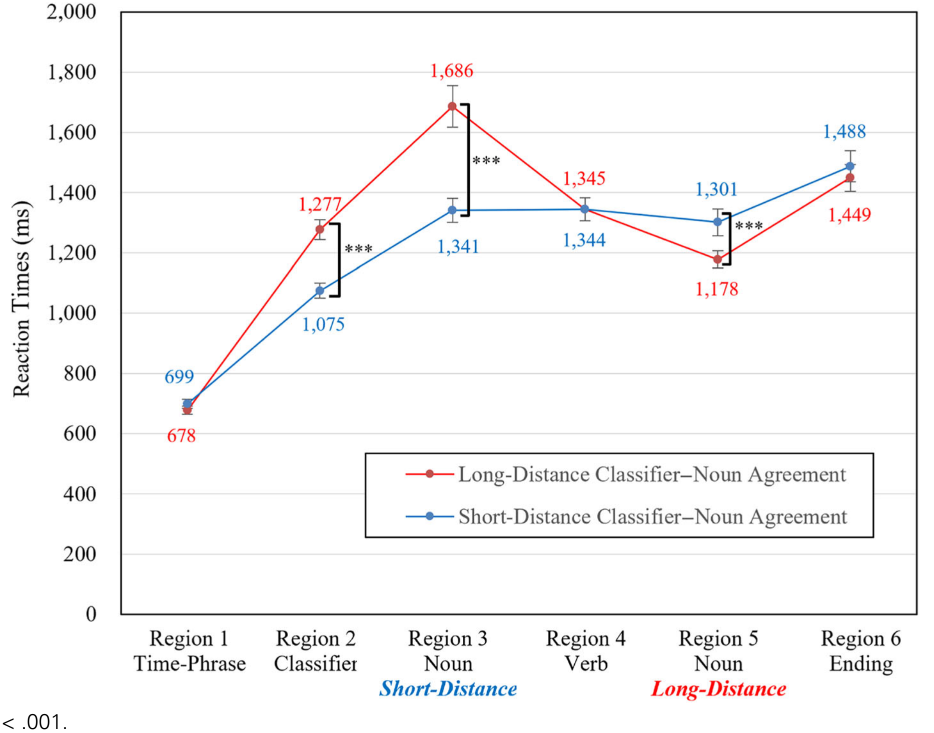
Discussion

The present study investigated how native Japanese speakers and Chinese JFL learners process short- and long-distance classifier–noun agreement structures, focusing on how syntactic dependency length influences real-time sentence comprehension.

The results from native Japanese speakers revealed no significant difference in accuracy between the short- and long-distance conditions, suggesting that both sentence types were grammatically well-formed and readily interpreted. Reaction time data showed significantly faster processing for short-distance dependencies in the classifier region (Region 2). Region 2 for short-distance dependencies was specialized as the human-denoting classifier (i.e., 二人の [+PEOPLE]) whereas varied in object-denoting classifiers for long-distance dependencies (e.g., 二皿の [+DISH], 二冊の [+ BOOK]), resulting in a repetition effect. Such effect led to the processing advantage for short-distance dependencies.

A clear processing advantage for short-distance dependencies was also observed in

Figure 3
Reaction times for each region in Chinese JFL learner group.



***$p < .001$.

Table 1
Results of reaction time analysis for Region 2

	Estimate	*SE*	*df*	*t* value	*p* value (>|*t*|)	*p*
(Intercept)	7.08	0.04	51.24	182.93	<.001	***
Condition	−0.17	0.02	623.25	−8.96	<.001	***
Proficiency	−0.04	0.03	46.17	−1.08	.286	*ns*
RT_1	0.06	0.01	603.55	4.54	<.001	***
Trial order	−0.12	0.02	43.42	−7.77	<.001	***
Condition × Proficiency	−0.04	0.02	622.65	−2.21	<.05	*

Note. Observations = 718, Participants = 44, Items = 36. Formula: logRT ∼1 + Condition + Proficiency + Condition:Proficiency + RT_1 + Trial order + (1 + Trial order | Participant) + (1 | Item) *ns* = not significant.
* $p < .05$, *** $p < .001$.

the subject noun (Region 3). This pattern aligns with previous findings on incremental sentence processing, where shorter dependencies reduce memory load and facilitate faster integration (Hsu et al., 2014; F. Wu et al., 2014; Yasunaga & Sakamoto, 2007). In the short-distance condition, the classifier (e.g., 二人の, "two people") immediately preceded and agreed with the subject noun (e.g., 料理人, "cook"), allowing for rapid integration. In contrast, in the long-distance condition (e.g., 二皿の料理人が炒めた料理, "the two dishes that the chef fried"), the classifier was semantically linked to the object noun (料理, "dish"), requiring the parser to maintain the classifier in working memory while processing intervening material. Moreover, semantic factors may have contributed to the observed differences. The inanimate classifiers used in the long-distance condition (e.g., 皿, 冊, 枚, 本) likely required more complex conceptual retrieval than the animate classifier 人, which was consistently used in the short-distance condition. In particular, the mismatch between the classifier and the

Figure 4
Predicting reaction times (logRT) by condition and Japanese proficiency in Region 2.

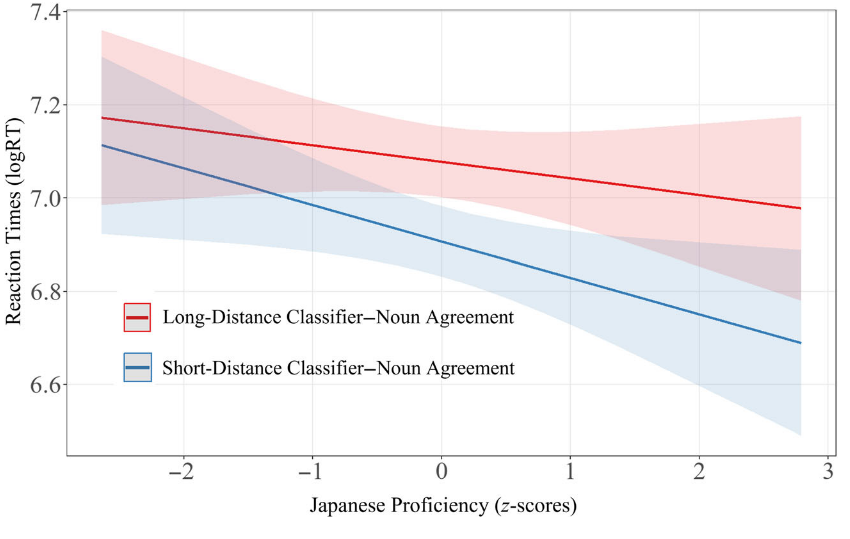


Table 2
Results of reaction time analysis for Region 3

	Estimate	*SE*	*df*	*t* value	*p* value (>|*t*|)	*p*
(Intercept)	−0.75	0.03	53.89	−23.58	<.001	***
Condition	−0.13	0.02	24.97	−5.39	<.001	***
Proficiency	−0.04	0.02	28.07	−1.54	.136	*ns*
PC	0.06	0.01	20.64	4.26	<.001	***
Trial order	−0.08	0.01	613.54	−8.41	<.001	***
Condition × Proficiency	−0.05	0.02	617.80	−2.66	<.01	**

Note. Observations = 725, Participants = 44, Items = 36. Formula: −1,000/RT ~ 1 + Condition + Proficiency + Condition:Proficiency + PC + Trial order + (1 + PC | Participant) + (1 + Condition | Item). *ns* not significant. ** $p < .01$, *** $p < .001$.

immediately following noun (e.g., 二皿の料理人, where an inanimate classifier precedes a human noun) may have triggered processing difficulty known as the classifier–noun agreement effect (Frankowsky et al., 2022; Her et al., 2018; Huettig et al., 2010). This semantic incongruity likely disrupted the flow of incremental interpretation, further slowing down reading in the long-distance condition.

Among native Japanese speakers, a significant difference between the two conditions was observed in the verb region (Region 4), with long-distance dependencies eliciting longer processing times than short-distance dependencies. This pattern suggests that native Japanese speakers actively exploit intervening material (i.e., the verb) to project upcoming structural expectations. Converging evidence is provided by Hsu et al. (2014), who investigated classifier–noun agreement violations in native Chinese speakers using ERP measures. In the region preceding the object noun (the relative

Figure 5
Predicting reaction times (−1,000/RT) by condition and Japanese proficiency in Region 3.

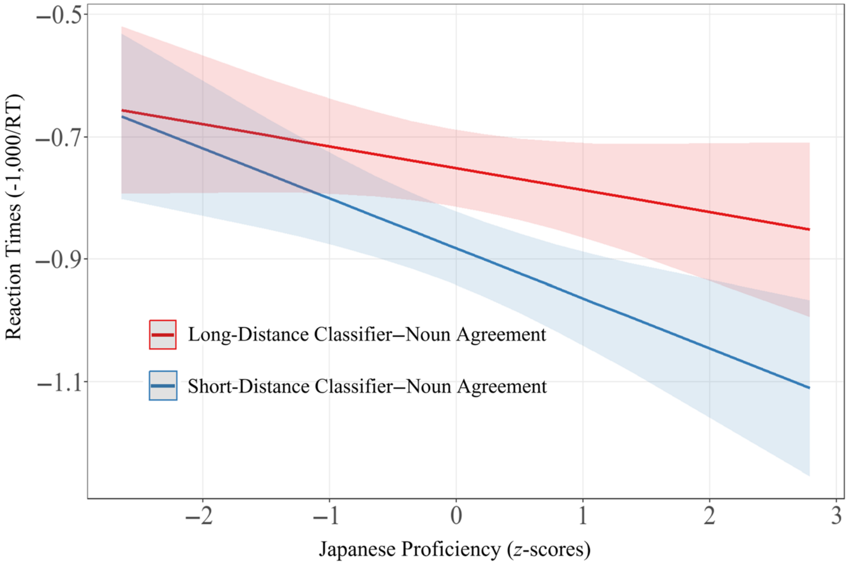


Table 3
Results of reaction time analysis for Region 5

	β	*SE*	*df*	*t* value	*p* value (>|*t*|)	*p*
(Intercept)	−0.98	0.04	58.45	−27.23	<.001	***
Condition	0.06	0.02	643.59	3.45	<.001	***
Proficiency	−0.03	0.02	51.98	−1.38	.174	*ns*
PC	0.07	0.01	565.74	8.53	<.001	***
Condition × Proficiency	−0.03	0.02	635.77	−2.03	<.05	*

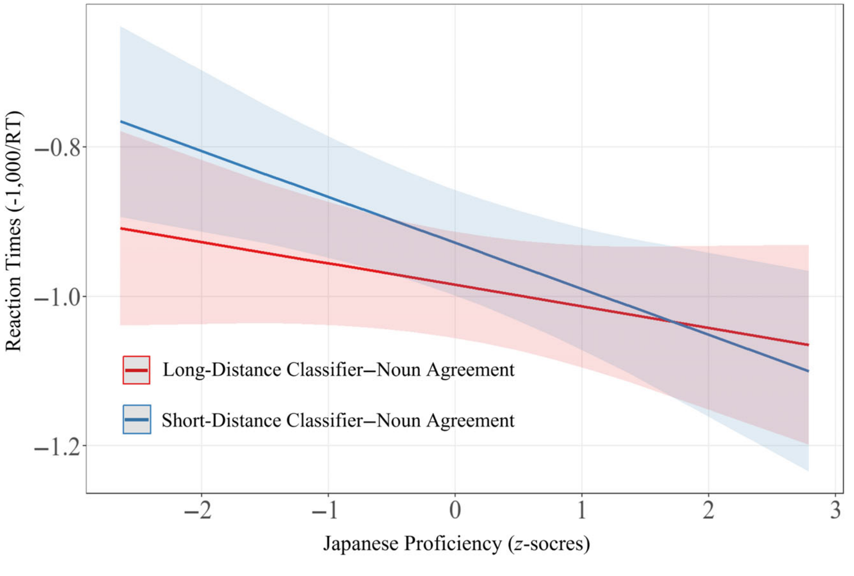
Note. Observations = 718, Participants = 44, Items = 36. Formula: −1,000/RT ~ 1 + Condition + Proficiency + Condition:Proficiency + PC + (1 | Participant) + (1 | Item). *ns* = not significant.
* $p < .05$, *** $p < .001$.

clause marker *de* in Chinese), they observed a significant N400 effect, such that long-distance dependencies elicited larger negativities than short-distance dependencies. The authors interpreted this effect as reflecting predictive processing.

In contrast, no comparable effect was observed in the learner group, suggesting that Chinese JFL learners may rely on a *good-enough* processing strategy (Ferreira et al., 2002; Swets et al., 2008; Tan & Foltz, 2020). Under this account, comprehenders construct partial or underspecified representations when task demands do not require fully detailed analysis. For learners, an early detection of a potential mismatch may lead to the suspension of active prediction, prompting them to wait for a compatible noun and resolve the dependency as soon as sufficient information becomes available rather than continuously updating predictive expectations.

The results in the following region (Region 5, the object noun) further differentiate native speakers' predictive processing from the more

Figure 6
Predicting reaction times (−1,000/RT) by condition and Japanese proficiency in Region 5.



reactive, good-enough processing observed in Chinese JFL learners. In this region, Chinese JFL learners showed a processing advantage for the long-distance condition, whereas native Japanese speakers showed no corresponding reaction-time difference. One possible explanation is that native speakers continuously maintain and update their predictions throughout sentence processing, such that facilitation from anticipation and costs associated with integration or reanalysis offset each other at the object noun. As a result, predictive processing may not manifest as a clear behavioral advantage. This interpretation is consistent with the ERP findings of Hsu et al. (2014), who reported a biphasic N400–P600 pattern at the object noun, reflecting the combined effects of anticipatory semantic activation and subsequent integration or reanalysis. Such overlapping facilitatory and inhibitory processes may therefore cancel out in reaction-time measures, yielding no observable difference despite ongoing predictive processing. For Chinese JFL learners, such cancellation is unlikely because predictive processing is weak and not continuously sustained. Instead, in the maze task, processing tends to terminate once a noun that satisfies the numeric classifier is encountered, resulting in a clear processing advantage when dependency resolution occurs later in the sentence.

In addition, among Chinese JFL learners, proficiency modulated reaction times of the two conditions in Regions 2, 3, and 5. For Region 2 (the classifier; Figure 4), as Chinese JFL learners' proficiency increased, processing speed for object-denoting classifiers improved more slowly than for subject-denoting classifiers, indicating relatively greater difficulty in lexical access for object-denoting classifiers. For Region 3 (subject NP; Figure 5), Japanese proficiency increased Chinese JFL learners' sensitivity to the classifier–noun mismatch. Higher-proficiency JFL learners showed a clearer divergence between the short- and long-distance conditions than lower-proficiency JFL learners, indicating improved ability to detect and integrate classifier–noun agreement. By contrast, lower-proficiency JFL learners showed less distinction between conditions, suggesting difficulty resolving the mismatch as it first emerged. As for Region 5 (head noun; Figure 6), proficiency reduced the processing advantage for the long-distance condition. Lower-proficiency JFL learners showed a bigger processing advantage, suggesting reliance on quickly resolving classifier–noun relations once the correct noun appeared. With increasing Japanese proficiency, this advantage diminished, indicating a shift toward more native-like parsing. Taken together, these patterns show that proficiency modulates how Chinese JFL learners differentiate between short- and long-distance dependencies, revealing a developmental shift from good-enough locally driven toward more structurally informed processing.

In sum, the present findings show that Japanese classifier–noun agreement processing by Chinese JFL learners is shaped by both dependency distance and language proficiency. Although Chinese JFL learners benefit from structural similarity between Chinese and Japanese, this similarity does not guarantee native-like online processing. Native speakers use verb information to project and revise expectations across long-distance dependencies, whereas learners rely more on locally available semantic cues and delay integration until the head noun is encountered. Proficiency modulates this pattern, revealing a gradual shift toward more structurally guided processing, but anticipatory parsing remains difficult even when L1 transfer is possible.

Conclusion

This study examined how native Japanese speakers and Chinese JFL learners process classifier–noun structures involving short- and long-distance syntactic dependencies. While both groups showed high accuracy in processing short-distance constructions, only native speakers demonstrated efficient, anticipatory processing using verb information to predict and integrate upcoming nouns before encountering the head noun. In contrast, Chinese JFL learners exhibited delayed integration, relying more heavily on locally available semantic cues, and showing bottom-up processing with resolution occurring only when the head noun was encountered. These findings suggest that, despite structural similarities between the Japanese and Chinese classifier systems, the ability to engage in predictive parsing does not automatically transfer from the first language to the foreign language. Rather, such anticipatory mechanisms appear to develop gradually with increased proficiency and exposure. The results highlight the complex interplay between dependency distance, L1–FL (or L1–L2) similarity, and learner proficiency in shaping real-time sentence comprehension.

Conflict of Interest

The authors declare no conflicts of interest associated with this manuscript.

References

Allan, K. (1977). Classifiers. *Language*, *53*(2), 285–311. https://doi.org/10.1353/lan.1977.0043

Anderson, D. R., Burnham, K. P., & Thompson, W. L. (2000). Null hypothesis testing: Problems, prevalence, and an alternative. *Journal of Wildlife Management*, *64*(4), 912–923. https://doi.org/10.2307/3803199

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

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28. https://doi.org/10.21500/20112084.807

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*(2), 211–243. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Boxell, O., & Felser, C. (2017). Sensitivity to parasitic gaps inside subject islands in native and non-native sentence processing. *Bilingualism: Language and Cognition*, *20*(3), 494–511. https://doi.org/10.1017/S1366728915000942

Clahsen, H., & Felser, C. (2006). Continuity and shallow structures in language processing. *Applied PsychoLinguistics*, *27*(1), 107–126. https://doi.org/10.1017/S0142716406060206

Croft, W. (1994). Semantic universals in classifier systems. *Word*, *45*(2), 145–171. https://doi.org/10.1080/00437956.1994.11435922

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*(1), 11–15. https://doi.org/10.1111/1467-8721.00158

Forster, K. I. (2010). Using a maze task to track lexical and sentence processing. *Mental Lexicon*, *5*(3), 347–357. https://doi.org/10.1075/ml.5.3.05for

Forster, K. I., Guerrera, C., & Elliot, L. (2009). The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods*, *41*(1), 163–171. https://doi.org/10.3758/BRM.41.1.163

Frankowsky, M., Ke, D., Zwitserlood, P., Michel, R., & Bölte, J. (2022). The interplay between classifier choice and animacy in Mandarin-Chinese noun phrases: An ERP study. *Language, Cognition and Neuroscience*, *37*(7), 866–882. https://doi.org/10.1080/23273798.2022.2026420

Grüter, T., Lau, E., & Ling, W. (2020). How classifiers facilitate predictive processing in L1 and L2 Chinese: The role of semantic and grammatical cues. *Language, Cognition and Neuroscience*, *35*(2), 221–234. https://doi.org/10.1080/23273798.2019.1648840

Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, *72*(358), 320–338. https://doi.org/10.1080/01621459.1977.10480998

Her, O. S., Chen, Y. C., & Yen, N. S. (2018). Neural correlates of quantity processing of Chinese numeral classifiers. *Brain and Language*, *176*, 11–18. https://doi.org/10.1016/j.bandl.2017.10.007

Hettiarachchi, S., & Pires, A. (2022). Second language acquisition of constraints on WH-movement by L2 English speakers: Evidence for full-access to syntactic features. *Language*, *7*(2), 134. https://doi.org/10.3390/languages7020134

Hsu, C. C., Tsai, S. H., Yang, C. L., & Chen, J. Y. (2014). Processing classifier-noun agreement in a long-distance: An ERP study on Mandarin Chinese. *Brain and Language*, *137*, 14–28. https://doi.org/10.1016/j.bandl.2014.07.002

Huang, S., & Schiller, N. O. (2021). Classifiers in Mandarin Chinese: Behavioral and electrophysiological evidence regarding their representation and processing. *Brain and Language*, *214*, 104889. https://doi.org/10.1016/j.bandl.2020.104889

Huettig, F., Chen, J., Bowerman, M., & Majid, A. (2010). Do language-specific categories shape conceptual processing? Mandarin classifier distinctions influence eye gaze behavior, but only during linguistic processing. *Journal of Cognition and Culture*, *10*, 39–58. https://doi.org/10.1163/156853710X497167

Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research*, *21*(2), 121–151. https://doi.org/10.1191/0267658305sr255oa

Juffs, A., & Harrington, M. (1995). Parsing effects in second language sentence processing: Subject and object asymmetries in wh-extraction. *Studies in Second Language Acquisition*, *17*(4), 483–516. https://doi.org/10.1017/S027226310001442X

Kanno, K. (2007). Factors affecting the processing of Japanese relative clauses by L2 learners. *Studies in Second Language Acquisition*, *29*(2), 197–218. https://doi.org/10.1017/S0272263107070118

Kobayashi, N., Ford-Niwa, J., & Yamamoto, H. (1996). A new test method for measuring Japanese-language proficiency: SPOT. *Sekai No Nihongo Kyōiku*, *86*, 201–218. (In Japanese.)

Lakoff, G. (1986). Classifiers as a reflection of mind. In C. G. Craig (Ed.), *Noun classes and categorization: Proceedings of a symposium on categorization and noun classification, Eugene, Oregon, October 1983* (pp. 13–52). John Benjamins. https://doi.org/10.1075/tsl.7.04lak

Lau, E., & Grüter, T. (2015). Real-time processing of classifier information by L2 speakers of Chinese. In *Proceedings of the 39th Annual Boston University Conference on Language Development* (pp. 311–323). Cascadilla Press.

Mak, W. M., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language*, *47*(1), 50–68. https://doi.org/10.1006/jmla.2001.2837

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Mitsugi, S. (2020). Generating predictions based on semantic categories in a second language: A case of numeral classifiers in Japanese. *International Review of Applied Linguistics in Language Teaching*, *58*(3), 323–349. https://doi.org/10.1515/iral-2017-0118

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

Pliatsikas, C., Johnstone, T., & Marinis, T. (2017). An fMRI study on the processing of long-distance wh-movement in a second language. *Glossa: A Journal of General Linguistics*, *2*(1), 101. https://doi.org/10.5334/gjgl.95

R Core Team. (2024). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, *2*, 110–114. https://doi.org/10.2307/3002019

Slavkov, N. (2015). Long-distance wh-movement and long-distance wh-movement avoidance in L2 English: Evidence from French and Bulgarian speakers. *Second Language Research*, *31*(2), 179–210. https://doi.org/10.1177/0267658314554939

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Under specification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, *36*(1), 201–216. https://doi.org/10.3758/MC.36.1.201

Tamaoka, K., Yu, S., Zhang, J., Otsuka, Y., Lim, H., Koizumi, M., & Verdonschot, R. G. (2024). Syntactic structures in motion: Investigating word order variations in verb-final (Korean) and verb-initial (Tongan) languages. *Frontiers in Psychology*, *15*, 1360191. https://doi.org/10.3389/fpsyg.2024.1360191

Tan, M., & Foltz, A. (2020). Task sensitivity in L2 English speakers' syntactic processing: Evidence for good-enough processing in self-paced reading. *Frontiers in Psychology*, *11*, 575847. https://doi.org/10.3389/fpsyg.2020.575847

Traxler, M. J., Williams, R. S., Blozis, S. A., & Morris, R. K. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Language*, *53*(2), 204–224. https://doi.org/10.1016/j.jml.2005.02.010

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. Springer.

Voeten, C. C. (2023). Buildmer: Stepwise elimination and term reordering for mixed-effects regression. R Package Version 2.11. Retrieved from https://CRAN.R-project.org/package=buildmer

Witzel, J., & Witzel, N. (2016). Incremental sentence processing in Japanese: A maze investigation into scrambled and control sentences. *Journal of Psycholinguistic Research*, *45*(3), 475–505. https://doi.org/10.1007/s10936-015-9356-4

Witzel, N., Witzel, J., & Forster, K. (2012). Comparisons of online reading paradigms: Eye tracking, moving-window, and maze. *Journal of Psycholinguistic Research*, *41*(2), 105–128. https://doi.org/10.1007/s10936-011-9179-x

Wu, F., Luo, Y., & Zhou, X. (2014). Building Chinese relative clause structures with lexical and syntactic cues: Evidence from visual world eye-tracking and reading times. *Language Cognition and Neuroscience*, *29*(10), 1205–1226. https://doi.org/10.1080/01690965.2013.841969

Wu, S., Liu, D., & Huang, S. (2023). The effects of over- and under-specified linguistic input on L2 online processing of referring expressions. *Journal of Psycholinguistic Research*, *52*(1), 283–305. https://doi.org/10.1007/s10936-022-09879-3

Yasunaga, D., & Sakamoto, T. (2007). On-line processing of floating quantifier constructions in Japanese: Using event-related brain potentials. *Journal of Japanese Linguistics*, *23*, 21–34. https://doi.org/10.1515/jjl-2007-0103

Zhu, S. (2021). An empirical study of the classifier effect in advanced L2 users of Chinese. *International Journal of Bilingualism*, *25*(1), 151–167. https://doi.org/10.1177/1367006920943301

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site: http://onlinelibrary.wiley.com/doi/10.1111/jpr.70035/suppinfo.

Data S1. Supplementary material.

(Received August 7, 2025; accepted March 6, 2026)