



OPEN ACCESS

EDITED BY
Zhenguang Cai,
The Chinese University of Hong Kong, China

REVIEWED BY
Tao Zeng,
Hunan University, China
Yu-Yin Hsu,
Hong Kong Polytechnic University,
Hong Kong SAR, China

*CORRESPONDENCE

Fengming Liu
✉ fl7@iu.edu

RECEIVED 25 May 2024
ACCEPTED 30 July 2024
PUBLISHED 19 August 2024

CITATION

Liu F and Lin C-JC (2024) Relative clause attachment in Mandarin Chinese: insights from classifier-noun agreement.
Front. Lang. Sci. 3:1438323.
doi: 10.3389/flang.2024.1438323

COPYRIGHT

© 2024 Liu and Lin. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Relative clause attachment in Mandarin Chinese: insights from classifier-noun agreement

Fengming Liu^{1*} and Chien-Jer Charles Lin²

¹Department of Second Language Studies, Indiana University, Bloomington, IN, United States,

²Department of East Asian Languages and Cultures, Indiana University, Bloomington, IN, United States

Introduction: Previous studies have shown that relative clause (RC) attachment preferences vary across languages, often influenced by factors like morphosyntactic agreement (e.g., number and gender). Mandarin Chinese, with its limited inflectional morphemes compared to Indo-European languages, provides a distinct context for examining this. This study explores relative clause attachment ambiguity in Mandarin by manipulating classifier-noun agreement.

Method: This study conducted two self-paced reading experiments to investigate the influence of an initial classifier on comprehenders' anticipation of its associated noun and the impact of this prediction on RC attachment preferences in Mandarin Chinese.

Results: Experiment 1 revealed a significant effect of classifier-noun agreement in offline comprehension: there was an increase in selecting the high-attachment noun (NP_{high}) as the RC attachment site when the classifier agreed with NP_{high}, whereas there was a decrease in selecting NP_{high} when the classifier agreed with the low-attachment noun (NP_{low}). Online processing results supported this effect, showing that classifiers guide comprehenders' expectations by pre-activating semantic features of the upcoming noun, thus modulating RC attachment preferences. Experiment 2 introduced semantic compatibility between the RC and potential attachment nouns as an additional disambiguating cue, revealing a reliable prediction effect for the upcoming noun. Although the classifier's prediction effect was diminished, it remained influential in this condition.

Discussion: This study highlights the complexity of relative clause attachment in Mandarin, demonstrating the significant predictive roles of classifier-noun agreement and semantic compatibility.

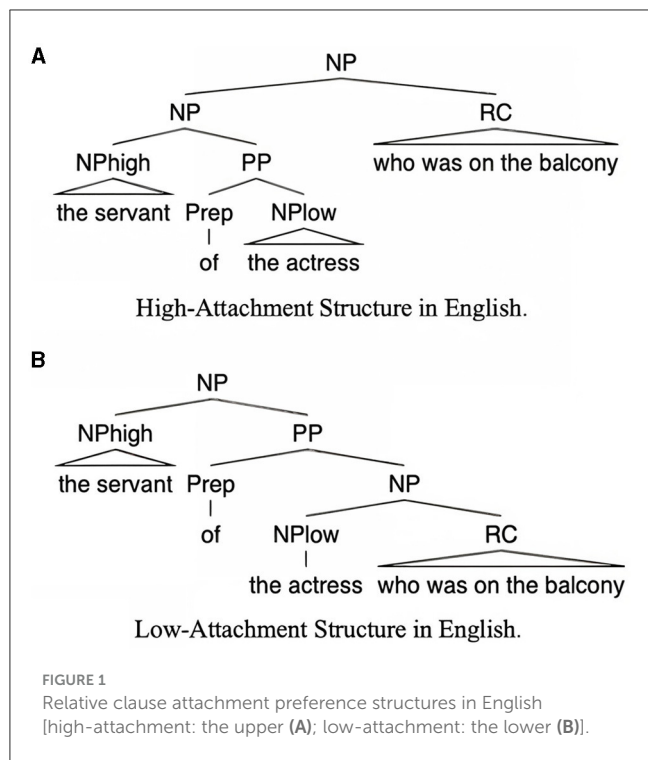
KEYWORDS

relative clauses, high/low-attachment, predictive processing, classifier-noun agreement, semantic compatibility, Mandarin

1 Introduction

The phenomenon of relative clause attachment ambiguity has attracted considerable attention among researchers. In English, resolving the attachment of relative clauses (RCs) is challenging when the head noun is inside a complex nominal phrase consisting of two noun phrases (NP_{high} of NP_{low}). This complexity introduces ambiguity, as the relative clause can modify either NP_{high} or NP_{low} (Cuetos and Mitchell, 1988).

Example (1) illustrates the attachment ambiguity with this type of sentence. When a relative clause *who was on the balcony* follows a complex noun phrase *the servant of the actress*, it can typically be attached to two potential noun sites. Thus, example (1) could imply that either *the servant* or *the actress* was the person *on the balcony*. To disambiguate the sentence, it becomes necessary to determine whether the relative clause attaches to the first potential host site (NP_{high}, *the servant*) or the second (NP_{low}, *the actress*). Resolutions of the first kind will be referred to as high attachment (HA) as NP_{high} occupies a higher



position in the syntactic structure. Correspondingly, instances of the second kind which favor NP_{low} will be labeled low attachment (LA) (see Figure 1 for illustrations).

- (1) Someone shot the servant of the actress who was on the balcony.

1.1 Cross-linguistic differences in relative clause attachments

Cross-linguistic studies have revealed variation in attachment preferences across languages. For instance, languages such as Spanish, Dutch, Italian, German, French, Japanese, and Korean tend to favor high attachment resolutions (Cuetos and Mitchell, 1988; de Vincenzi and Job, 1993; Hemforth et al., 1994; Brysbaert and Mitchell, 1996; Kamide and Mitchell, 1997; Zagar et al., 1997; Fernández, 2002; Lee, 2021). On the other hand, languages such as English and Mandarin tend to exhibit a preference for low attachment (Cuetos and Mitchell, 1988; Fernández, 2002; Kwon et al., 2019).

However, the distinctions between the two language groups are not as clear-cut as they seemed. Research findings often yield contradictory results due to methodological differences. For example, paper-based questionnaires typically tap into offline processing (later stages) while self-paced reading or eye-tracking techniques tap into online processing (earlier stages). Consequently, studies utilizing offline questionnaire tasks consistently report a high-attachment preference in Spanish (Cuetos and Mitchell, 1988; Fernández, 2002). In contrast, studies employing self-paced reading tasks have yielded inconsistent

results, finding either a low-attachment preference (Fernández, 2002) or a high-attachment preference (Mitchell and Cuetos, 1991). Moreover, the semantic and length properties of the test materials, such as using human versus non-human nouns or the length and complexity of the relative clauses, also contribute to these differences (Fernández, 2002; Hemforth et al., 2015; Kwon et al., 2019).

Mitchell and colleagues conducted a series of studies investigating relative clause attachment in several languages. One of the pioneering studies by Cuetos and Mitchell (1988) compared RC attachment preferences in Spanish and English. Despite sharing the SVO word order and head-initial RC structure, Spanish and English differ in the word order of adjectives and nouns. Unlike the prenominal adjective order in English, Spanish commonly follows a postnominal adjective order, where the adjective appears after the noun it modifies (Cuetos and Mitchell, 1988). Cuetos and Mitchell (1988) used equivalent materials, providing literal translations of sentences in both English and Spanish, in an offline questionnaire task (as shown in example (1) above). In this task, each sentence was followed by a question, such as “Who was on the balcony?” to determine the attachment preference. The results revealed a significant tendency among Spanish participants to attach the RC to the first noun (NP_{high}, i.e., servant), showing a strong high-attachment preference, whereas English participants preferred attachment to the second noun (NP_{low}, i.e., actress), indicating a low-attachment preference. This finding demonstrated a cross-linguistic difference in attachment preferences between Spanish and English. Subsequent studies, such as Fernández (2002), confirmed this cross-linguistic distinction, with questionnaire results replicating the findings in Cuetos and Mitchell (1988). In Fernández’s (2002) study, Spanish readers exhibited a higher likelihood (57%) of choosing the high-attachment site (NP_{high}) compared to English readers (43%).

Mandarin, on the other hand, follows an SVO pattern and a head-final RC structure. In addition, Mandarin employs the possessive particle *de* to express possession or genitive relationships instead of using a Norman genitive construction. This particle, *de*, is positioned after the possessor noun to indicate ownership. For instance, corresponding to *the servant of the actress*, *nyanyuan* “actress” *de* “DE” *puren* “servant” in Mandarin would mean *the actress’ servant*, as shown in example (2) below. Therefore, in Mandarin, the first NP (actress) serves as the low attachment site, while the second NP (servant) serves as the high attachment site. Kwon et al. (2019) found a prevailing preference toward low attachment in Mandarin in an offline sentence completion task.

- (2) Mandarin (adapted from Shen, 2006)

<i>Mouren</i>	<i>dasile</i>	<i>zhanzai</i>	<i>nyanyuan</i>	<i>puren.</i>
		<i>yangtai-</i>	<i>de</i>	
		<i>shang de</i>		
Someone-	shot	on the	actress-	servant.
Nom		balcony-	particle	
		Relativizer		
Subject_	Predicate_	RC	NP1-Low	NP2-
matrix	matrix			High
“Someone shot the servant of the actress who was on the balcony.”				

As we can see, relative clauses involving global ambiguity have been widely used in offline tasks to explore attachment preferences in a variety of languages. Cross-linguistic differences have been observed in languages that differ with respect to typological properties such as word order variability, alternative forms for expressing genitive meanings, the overt/covert relative pronouns, as well as left-branching/right branching RC structures.

1.2 Factors that modulate relative clause attachment preference

Most studies (e.g., Lee and Kweon, 2004; Kwon et al., 2019) discussed readers' attachment preferences within the context of offline tasks, tapping into the later phases of processing. To gain insights into readers' preferences during the earliest phases of processing, a number of studies have employed online tasks, such as self-paced reading and eye-tracking techniques, measuring differences in terms of processing difficulty associated with forcing attachment in either direction. For instance, an online task may compare reading times across three conditions: ambiguous, forced high-attachment, and forced low-attachment (Fernández, 2002). Readers' preferences in the early processing phases would manifest as increased reading times at critical regions when encountering materials that contain the dis-preferred forced attachment, as compared to reading times associated with either ambiguous materials or those containing the preferred forced attachment (Fernández, 2002). Studies on Indo-European languages often employ morphosyntactic agreement as a means to disambiguate relative clauses with complex head nouns. English, for instance, always utilizes number agreement to force high or low attachments, as shown in example (3) (Fernández, 2002, p. 195).

(3) Number agreement in English

- a.... the nephew of the teacher that was... (Ambiguous)
- b.... **the nephew** of the teachers that **was**... (Forced HA)
- c.... the nephews of **the teacher** that **was**... (Forced LA)

In terms of early attachment preferences, Fernández (2002) conducted a self-paced reading task comparing reading times in forced low and forced high conditions in both English and Spanish. Surprisingly, no differences were found between the two languages, as both exhibited faster reading times in the forced low attachment condition (3c) compared to the forced high attachment condition (3b). These results suggest that the initial attachment choice in relative clauses tends to be the low site. Fernández attributed the observed cross-language differences between Spanish and English in offline tasks to the influence of post-syntactic information. Similarly, a study by de Vincenzi and Job (1993) employed gender agreement in Italian to disambiguate relative clause attachment. Contrary to the high-attachment preference found in offline questionnaires, a self-paced reading task revealed a low attachment preference in the early phase, suggesting that the initial low-attachment choice is subsequently adjusted to high-attachment considering post-syntactic information. Kamide and Mitchell (1997) conducted a study in Japanese where they manipulated the pragmatic compatibility between the RC and the two potential attachment sites, aiming to force the RC to attach to

either NP. Using a self-paced reading task with seven segmented regions (RC/NP1-Low/Gen/NP2-High/Acc/Matrix subject/Final), the study revealed significantly longer reading times in the regions associated with NP1-Low and Gen in the forced high-attachment condition, compared to the forced low-attachment condition. Interestingly, the Final region exhibited significantly longer reading times in the forced low-attachment condition than in the forced high-attachment condition. These findings suggest that parsers initially exhibit a low-attachment when encountering the first potential attachment site (NP1-Low), but their preference switches to high attachment by the end of the sentence.

In contrast to the findings observed in Japanese, the languages Korean and Mandarin exhibit consistent attachment preferences in both offline and online tasks. Specifically, Korean consistently shows high attachment as the initial and eventual choice, while Mandarin consistently favors low attachment. Lee and Kweon (2004) employed a disambiguating device in the form of subject-verb agreement with an honorific marker *-si*. By manipulating the agreement between the verb and the potential agents using the honorific marker, the relative clauses are forced to attach to either site. Similarly, Shen (2006) manipulated the semantic compatibility of the RC with alternative host sites and conducted both a self-paced reading task and an offline questionnaire. The results from Shen (2006) demonstrated a low-attachment preference that persisted from the late online processing phase into the offline phase.

While online tasks have commonly utilized morphosyntactic agreement (e.g., number and gender) in relative clauses to disambiguate attachment preferences in languages like English, Spanish, Italian, and French, Mandarin Chinese differs from most Indo-European languages in having a relatively limited set of inflectional morphemes. However, Mandarin features classifier-noun agreement, which remains largely unexplored in relation to the RC attachment preferences. The present study addressed this issue by conducting two self-paced reading experiments using classifier-noun agreement as a disambiguating cue to resolve relative clause attachment preferences.

1.3 Predictions in sentences in relation to classifier-noun agreement

1.3.1 Prediction in language processing

Language comprehenders actively make predictions about upcoming linguistic units, whether they are syllables, words, or grammatical structures. The ability to incrementally form hypotheses about incoming linguistic material, known as predictive processing, is a fundamental property of sentence processing (Kuperberg and Jaeger, 2016; Pickering and Gambi, 2018). Comprehenders use context to make predictions, pre-activating linguistic representations before actually encountering them (Pickering and Garrod, 2013; Pickering and Gambi, 2018). This pre-activation allows the brain to prepare for expected linguistic input in advance, facilitating quicker and more efficient processing.

A number of studies have demonstrated that people can predict semantic information of upcoming words, particularly in highly

constraining contexts (Federmeier and Kutas, 1999; Kamide et al., 2003; Wang et al., 2018). However, less is known about predictions in less-constraining contexts, which may allow for the prediction of general information characterizing a group of words rather than specific words (Huang et al., 2023). Some studies have explored this issue by using morphosyntactic dependencies triggered by cues such as gender-marking (German: Hopp and Lemmerth, 2018), case marking (Turkish: Özge et al., 2019; German: Özge et al., 2022), and honorific marking (Lee and Yoo, 2023). These studies have found that people can use morphosyntax to generate predictions during the course of processing.

Hopp and Lemmerth (2018), for example, examined whether gender agreement between pre-nominal determiners/adjectives and nouns in German can aid comprehenders in predicting the upcoming nouns. Using a visual-world task, they found that gender marking on the determiner or adjective helped participants identify the object among competitors. Participants spent less time identifying the object when the gender marking (masculine or feminine) on the determiner or adjective provided disambiguating cues compared to when the gender marking (neuter) provided no disambiguating clues. These findings suggest that gender markings on determiners or adjectives pre-activate the gender information of the upcoming nouns, facilitating faster and more accurate identification of objects that share the gender features with the predicted nouns.

Lee and Yoo (2023) also demonstrated the pre-activation effect of honorific agreement between the subject and verb in Korean. In a self-paced reading task, they manipulated the honorific agreement by varying whether the subjects had honorific features. Results showed that subjects with honorific features elicited longer reading times when the verbs do not carry an honorific marker, compared to subjects without honorific features, supporting predictive processing, as participants pre-activate the honorific information when they encounter a subject with honorific features and expect an honorific marker *-si* on the following verb before encountering it. In addition, Özge et al. (2019, 2022) found that both Turkish

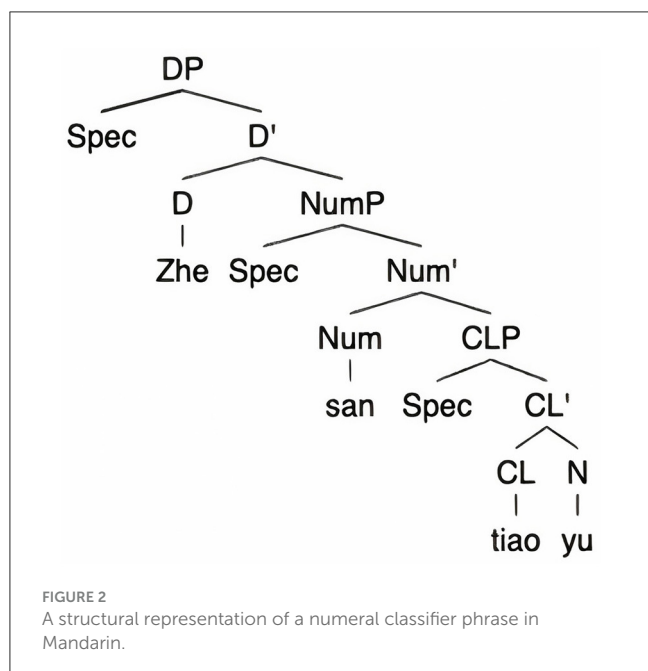
and German children as young as four years old were able to use sentence-initial case marking cues to predict the thematic role of the upcoming argument.

These studies have demonstrated the role of morphosyntactic agreement and dependencies in facilitating predictive processing in less-constraining contexts, suggesting that comprehenders use morphosyntactic cues to anticipate and prepare for upcoming linguistic information. However, gender, case, and honorific agreements are highly grammaticalized and correlate with overt morpho-syntactic markers. This raises the question of whether comprehenders can also rely on agreement or dependency relationships that are not explicitly marked to predict upcoming linguistic information during sentence comprehension. To explore this, the present study adopted the classifier-noun agreement in Mandarin Chinese to test whether people can utilize the semantic constraints of classifiers to make predictions for the upcoming nouns.

1.3.2 Classifier-noun agreement and its prediction effect

Classifiers are systems used to mark and categorize noun classes (Erbaugh, 2006; Ahrens and Huang, 2016). Mandarin is one of the languages that employs numeral classifiers, which appear after a number or determiner and before a noun (see example (4) and Figure 2 for illustrations). In Mandarin, classifiers exhibit selectivity for specific noun classes, establishing a classifier-noun agreement through selectional restrictions. For instance, the classifier *tiao* “CL(objects/animals)” typically pairs with objects or animals that are long, thin, cylindrical, and flexible, such as *yu* “fish” and *maojin* “towel”, while it cannot be used with other nouns like flowers or books (Ahrens and Huang, 2016, p. 177).

(4) <i>Zhe</i>	<i>san</i>	<i>tiao</i>	<i>yu</i> .
DET	NUM	CL	N
This	three	CL	fish
‘these	three	fishes’	



Several studies (Zhou et al., 2010; Chou et al., 2015; Chan, 2019) have investigated the dependencies between classifiers and nouns. Zhou et al. (2010) conducted an event-related potential (ERP) study to explore the temporal neural dynamics of semantic integration processes in a hierarchical structure: *subject noun + verb + numeral + classifier + object noun*. They manipulated the semantic (in)congruency in the classifier-noun, verb-noun, and verb-classifier at different levels of syntactic hierarchy during sentence processing. The results revealed significant N400 effects for nouns in all three mismatch conditions, indicating semantic processes at both lower- and higher-level syntactic structures. However, Chan (2019) challenged the interpretation of the N400 effect observed in the classifier-noun mismatch reported by Zhou et al. (2010). Chan argued that it might not be a typical N400, but rather a confounding result of LAN (left anterior negativity) and N400 components, “signaling the coexistence of morphosyntactic and semantic agreement in processing classifier phrases” (Chan, 2019, p. 14). This interaction between semantic and syntactic processes in classifier-noun agreement has also been observed in Chou et al. (2015).

Despite the controversial nature of classifier-noun agreement, researchers agree that classifiers semantically constrain their following nouns in features such as animacy, shape, or size, and that comprehenders can pre-activate semantic features of the upcoming nouns when they encounter classifiers (Kwon et al., 2017; Grüter et al., 2019; Huang et al., 2023). For example, Huang et al. (2023) investigated whether comprehenders use classifiers to predict the animacy of upcoming nouns. Native Chinese speakers were presented with animate-constraining (e.g., 位, *wei*) or inanimate-constraining classifiers (e.g., 本, *ben*) followed by congruent or incongruent nouns. ERP analyses revealed an N400 effect for incongruent conditions, indicating the difficulty of semantic integration with incompatible nouns, consistent with previous studies (Zhou et al., 2010; Chou et al., 2015; Chan, 2019). Additionally, representational similarity analysis showed that neural activity patterns were more similar following animate-constraining classifiers than inanimate-constraining classifiers before noun presentation, driven by perditions from classifiers about the animacy features of the upcoming nouns.

In addition, Kwon et al. (2017) explored the classifiers' prediction effect in a highly constraining discourse context using relative clauses (e.g., 张艺谋指导的... *zhangyimou zhidao de...*, "Zhang Yimou directed..."), which set an expectation for a specific head noun. The head nouns in these clauses were manipulated to be either predicted (电影, *dianying*, "movie") or non-predicted (e.g., 大厦, *dasha*, "building"), and paired with appropriate classifiers (部, *bu* for movie; 座, *zuo* for building), establishing classifier-noun agreement. The results demonstrated that participants pre-activated semantic features of the predicted head noun (电影, *dianying*, "movie") when they encountered semantically constraining relative clauses. This was evidenced by an N400 effect when a mismatched classifier (座, *zuo*) was read, reflecting challenges in processing a classifier incongruent with the pre-activated semantic information of the predicted noun.

What's intriguing about the classifier-noun agreement is that the classifier and its associated noun can be separated from each other, creating a long-distance dependency. In Mandarin, noun phrases strictly follow a head-final structure, requiring determiner-classifier sequences (such as demonstratives, numerals, and classifiers) and modifiers (adjectives and relative clauses) to precede the head noun. This structural characteristic enables the insertion of a modifying relative clause between the classifier and its associated noun, as demonstrated in (5).

(5) a. Intervening by a Subject-gap RC

Yi tiao	zhengzai	youyong	de	yu
Num-	<i>t_i</i> PROGRE-	V	RELATI-	head
CL(fish)	SSIVE		VIZER	noun _i
One-	<u>gap</u> be	swimming	DE	fish
CL(fish)		Subject RC		

'a fish that is swimming'

b. Intervening by an Object-gap RC

Yi tiao	linju	mai	de	yu
Num-	NP	V <i>t_i</i>	RELATI-	head
CL(fish)			VIZER	noun _i
One-	neighbor	bought <u>gap</u>	DE	fish
CL(fish)		Object RC		

'a fish that the neighbor bought'

In (5a), a subject-gap RC is inserted between the classifier *tiao* "CL(long-shaped objects/animals)" and its associated noun *yu* "fish", creating a long-distance dependency between them. Similarly, (5b) establishes a long-distance dependency between the classifier *tiao* and its associated noun, but also introduces a temporary local incongruity due to the additional noun *linjun* "neighbor" serving as the embedded subject in the object-gap RC. This issue of syntactic uncertainty and potential local semantic incongruity is also known as the garden path problem for processing head-final relative clauses (Lin and Bever, 2011; Lin, 2019).

Researchers (e.g., Hsu et al., 2014; Wu et al., 2018) have investigated whether the temporary mismatch between the classifier and noun as in (5b) can effectively serve as a cue for predicting an RC structure. For example, Hsu et al. (2014) investigated the impact of temporary classifier-noun incongruity on RC disambiguation by introducing an object-gap RC between a classifier and the head noun in a head-final RC structure. An ERP study revealed that the initial incongruity between the classifier and its adjacent noun did not guarantee an RC prediction. Wu et al. (2018) explored the same issue by manipulating the clause-initial classifier in a self-paced reading task. Their findings demonstrated that encountering temporary incongruity between a classifier *wei* "CL(people)" and its adjacent noun *shitou* "stone" in a sentence-initial position would lead the reader to postulate a phrasal boundary between them and anticipate a suitable noun (in this case, a human) that agrees with the classifier in the upcoming words. Moreover, they found that the temporary incongruity resulted in faster reading times in the region following the RC head noun, indicating that the initial classifier-noun mismatch facilitated the anticipation of an RC structure, albeit the effect was observed in the spill-over region.

Overall, the existing evidence indicates that classifiers can serve as cues for predicting the semantic features of upcoming nouns on a phrasal or sentential level (e.g., relative clauses). In other words, people can pre-activate semantic features such as animacy, humanness, size, or shape of the upcoming nouns when they encounter classifiers before the nouns appear, due to the semantic and syntactic dependencies between classifiers and their associated nouns in Mandarin Chinese. However, the extent to which the dependency between classifiers and the upcoming nouns can modulate RC attachment preferences remains underexplored. Therefore, this study investigated whether classifier-noun agreement can function as a morphosyntactic predictor of semantic dependencies and influence relative clause attachment preferences.

1.4 The present study

This study conducted two self-paced reading experiments to investigate the influence of an initial classifier on comprehenders' anticipation of its associated noun and the impact of this prediction on relative clause attachment preferences in Mandarin Chinese. This study involved the insertion of a subject-gap RC between a demonstrative-classifier phrase (DEM-CL) and a complex RC head noun (NP_{low} *de* NP_{high} , "NP_{high} of NP_{low}"), resulting in

the sequence “DEM-CL [*gap* VP *de*]_{RC} NP_{low} *de*₂ NP_{high}.” For instance, in sentence (6), a subject-gap RC (*who is drinking water*) was inserted between the demonstrative-classifier phrase (*zheming* “this-CL(people)”) and a complex nominal phrase (*yigong de xiaomao* “volunteer’s cat”).

(6)	<i>Zheming</i>	<i>zhengzai de</i>	<i>yigong</i>	<i>de</i>	<i>xiaomao</i>
		<i>heshui</i>			
	DEM-	RC	de1	NP _{low}	de2 NP _{high}
	CL(people)				
	This-	is	relati-	volunteer	poss. cat
	CL(people)	drinking	vizer		
		water			
					‘the cat of the volunteer who is drinking water.’

Note that the two occurrences of the functional morpheme DE realized as *de*₁ and *de*₂ are distinctive in their grammatical functions and interpretations, as *de*₁ is the relativizer of the RC while *de*₂ is a prenominal modifier indicating a possessive (genitive) relationship between NP_{low} and NP_{high}. Crucially, the DEM-CL phrase is placed before the relative clause to generate a long-distance classifier-noun agreement. In addition, since previous studies have shown that animacy interacts with RC types and can affect RC attachment (Kwon et al., 2019), in this experiment, we controlled for both animacy and RC type by adopting only animate nouns to serve the head nouns in subject-gap RCs. We also set up NP_{low} to be semantically [+human] and NP_{high} to be [+animal], and there exist the same possessive relationships between NP_{low} and NP_{high} in all experimental items.

Since attachment preferences are ultimately related to temporarily storing RCs and NPs in the working memory for interpretation, the present study also explores the potential role of working memory capacity (WMC) as an individual difference variable. The impact of WMC on RC attachment preferences has long been studied, but conflicting results still exist. Some offline studies have shown a link between low WMC and a preference for high attachment. For instance, Mendelsohn and Pearlmutter (1999) found that low-span subjects in English preferred high attachment, a finding supported by subsequent research (Swets et al., 2007; Cotter and Ferreira, 2014). However, other offline studies (German: Harding et al., 2019; English and Persian: Mahmoodi et al., 2022) do not find a correlation between WMC and RC attachment preferences. Similarly, in online tasks, some studies have shown that low-span subjects are inclined to attach RCs to the high-attachment site, compared to high-span subjects (English: Marefat et al., 2015). However, other studies have reported the opposite pattern (Traxler, 2007), and some have found no effect of WMC (Omaki, 2005; Traxler, 2009). Therefore, the present study incorporates WMC, measured by an adapted reading-span task for Mandarin sentence processing (Daneman and Carpenter, 1980) to investigate its impact in moderating RC attachment preferences in both offline and online stages.

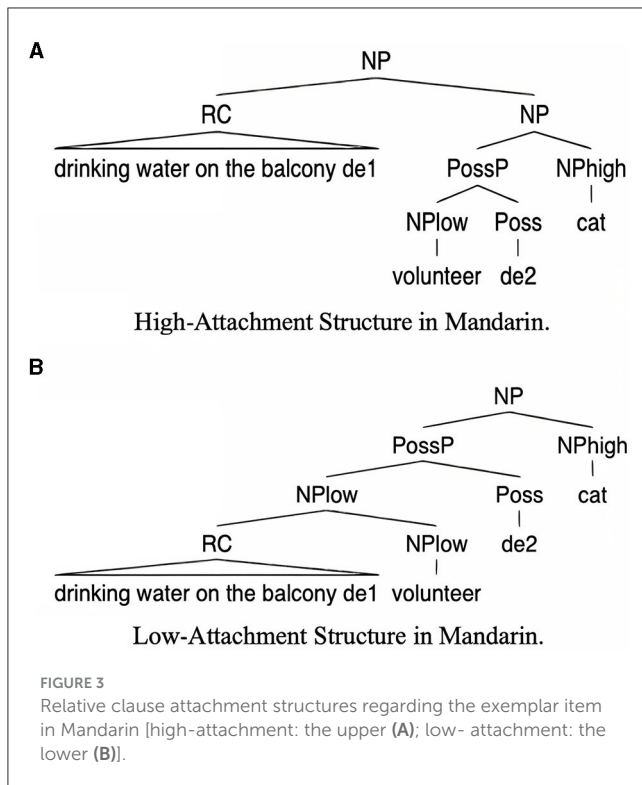
2 Experiment 1

Experiment 1 looked at comprehenders’ RC attachment preferences in Mandarin Chinese during both online and offline processing. We also studied how the presence of an initial classifier

TABLE 1 An experimental item in the four conditions.

CL-N agreement	Test item								
	Dem-CL	RC (Aspect/Loc/Prc)	de1	NP _{low}	de 2	NP _{high}	Adv _{Main}	VP _{Main}	NP _{Main}
CL-NP _{low}	这名/this-CL (people)	正在/阳台上/喝水/drinking water on the balcony	的/rel.	义工/volunteer	的/poss.	小猫/cat	旁边/nearby	有两把/there are two	椅子。 chairs.
CL-NP _{high}	这只/this-CL (animal)								
CL-NP _{low} /NP _{high}	这群/this-CL (group)								
No CL (control)									

English translation: There are two chairs nearby the cat of the volunteer that is drinking water on the balcony.



pre-activate semantic information of its associated nouns, thus impacting comprehenders' RC attachment preferences.

In this experiment, we manipulated the agreement between the classifier and the two potential attachment sites, NP_{low} and NP_{high}, and created four conditions: (i) CL-NP_{low}, where the classifier only agrees with NP_{low}, enforcing a low attachment; (ii) CL-NP_{high}, where the classifier only agrees with NP_{high}, enforcing a high attachment; (iii) CL-NP_{low}/NP_{high}, where the classifier could agree with either NP_{low} or NP_{high}; and (iv) No CL, representing a condition without a classifier. Both the CL-NP_{low}/NP_{high} and the No CL condition create global ambiguity in terms of RC attachment preferences. Table 1 demonstrates the four conditions with an example. See also Figure 3 for illustrations on syntactic structures regarding the exemplar item. In the CL-NP_{low} condition, the classifier *ming* "CL(people)" agrees only with the first noun NP_{low} *yigong* "volunteer" but not with NP_{high} *xiaomao* "cat". In the CL-NP_{high} condition, the classifier *zhi* "CL(animal)" agrees solely with the second noun NP_{high} *xiaomao* "cat" but not with NP_{low} *yigong* "volunteer". The CL-NP_{low}/NP_{high} condition introduced ambiguity by allowing the group classifier *qun* "CL(group)" to be congruent with either NP_{low} or NP_{high}. Finally, the No CL condition does not involve a classifier.

Drawing on the prediction effect in language processing (Kuperberg and Jaeger, 2016; Pickering and Gambi, 2018), and considering Mandarin's SVO word order and head-final RC structure, we predict that comprehenders pre-activate semantic features of the upcoming noun upon encountering the sentence-initial classifier. This anticipation persists until they come across the two potential attachment noun sites. We hypothesize that the classifier-noun agreement, similar to number agreement in English or gender agreement in Spanish, would serve as a disambiguating

factor influencing comprehenders' RC attachment preferences. When the classifier holds a semantic dependency with the lower noun NP_{low}, the RC may also preferably be attached to NP_{low}. However, when the classifier holds a semantic dependency with the higher noun, NP_{high}, the RC is more likely to be attached high to this non-local NP. By manipulating the agreement between the initial classifier and either of the two potential attachment sites (NP_{low} or NP_{high}), we may observe the semantic influence of classifiers on comprehenders' preferences for high or low RC attachments.

2.1 Methods

2.1.1 Participants

Eighty native Mandarin Chinese speakers (62 females and 18 males) participated in this experiment, recruited via Chinese social media and university classified ads in China. Participants' ages ranged from 18 to 25 years old. Informed consent was obtained for experimentation with participants. Compensation was provided for their participation. The entire study was approved by IRB (##2005886743).

2.1.2 Materials

Twenty-four experimental items were created,¹ all following the same syntactic structure and segmented into eleven regions (see Table 1 for an example). Six critical regions of interest are: de1, NP_{low}, de2, NP_{high}, ADV_{main} and VP_{main}. All the experimental items were counterbalanced using a Latin-square design. Each participant was presented with one version of each experimental item. Additionally, 56 unambiguous distractors were also included in this experiment. These included 14 subject-gap RCs with animate head nouns, 14 object-gap RCs with inanimate head nouns, 14 non-RC sentences featuring a passive *BEI* construction ("be-passive"), and 14 non-RC sentences with an existential *YOU* construction ("there exists..."). Each participant encountered a total of 80 trials, presented in a random order.

To enhance the natural flow of the stimuli and to prime participants' awareness of classifier-noun agreement prior to introducing each test item, a referential discourse context was provided. This context included equal mentions of classifier nominal phrases for both NP_{low} and NP_{high} twice. For instance, the context for the exemplar item in Table 1 would be "A CL(people) volunteer adopted a CL(animal) stray cat. One day, the CL(people) volunteer brought this CL(animal) cat to my house to play." The contexts remained consistent across all four conditions for each item, with the exception of the CL-NP_{low}/NP_{high} condition where the classifiers for people and animals were replaced with the classifier *qun* "CL(group)".

Before carrying out the experiment, two norming tests were conducted to validate the acceptability of classifier-noun agreement

¹ The numbers of test sentences in Experiment 1 ($k = 24$) and Experiment 2 ($k = 28$) were determined by referencing previous studies on RC attachment that employed similar methodologies (e.g., Kamide and Mitchell, 1999; Fernández, 2002; Shen, 2006; Kwon et al., 2019).

and the equitable use of relative clauses to modify NP_{low} or NP_{high}. Twenty-five native Mandarin Chinese speakers were paid to evaluate all the classifier-noun pairs used in both Experiment 1 and Experiment 2. Participants were required to rate the acceptability of classifier-noun pairs on a scale ranging from 1 to 7, with 1 signifying complete unacceptability and 7 indicating total acceptability. Only classifier-noun pairs rated above 6 were incorporated into the experiments: CL- NP_{low} (people) ($M = 6.71$; $SD = 0.19$); CL- NP_{high} (animal) ($M = 6.90$; $SD = 0.15$); Qun- NP_{low} (people) ($M = 6.65$; $SD = 0.17$); Qun- NP_{high} (animal) ($M = 6.73$; $SD = 0.16$).

In addition, two surveys were created using a Latin-square design to assess the modifying plausibility of relative clauses. Each survey contained 24 items from Experiment 1 and 56 items from Experiment 2 (will be discussed further in the Experiment 2 section). To maintain a balance between acceptable and unacceptable items, 40 fillers of comparable length and structure were included. These fillers were intentionally crafted to be semantically unacceptable, for example, “*the goldfish that is watching TV on the couch.*” A distinct group of 20 native speakers was randomly assigned to one of the surveys. Each participant assessed the acceptability of relative clauses modifying either [+human] nouns (e.g., “*the volunteer that is drinking water on the balcony*”) or [+animal] nouns (e.g., “*the cat that is drinking water on the balcony*”) on a 7-point Likert scale, where 1 indicated complete unacceptability and 7 represented full acceptability. The results demonstrated that relative clauses in Experiment 1 could modify both NP_{low} (people) ($M = 6.65$; $SD = 0.33$) and NP_{high} (animal) ($M = 6.58$; $SD = 0.40$) with equal acceptability. A two-tailed paired t-test revealed no significant differences between them ($t(23) = 0.16$, $p = 0.87$).

2.1.3 Procedure

In this experiment, a moving-window self-paced reading task was developed using jsPsych Version 6.2 (de Leeuw et al., 2023), and executed on the Cognition.run platform. Participants were presented with a context in the center of the screen. After reading the context, they could proceed by pressing the spacebar. The context would then be replaced by a fixation “+”. Pressing the spacebar again allowed participants to advance to the next page, where the fixation was replaced by a dashed line. Participants pressed the spacebar to move through the successive displays of the target sentence, with their reaction times recorded for each region.

Following the target sentence, participants advanced to the comprehension question phase, where they were presented with a forced-choice question regarding the RC attachment. For instance, “*___ is drinking water on the balcony? F: volunteer; J: cat.*” The relative order of the two options was balanced across test trials, with half of them having F taken by NP_{low} and J taken by NP_{high}, and the other half with the reverse order. Participants’ choices were recorded during this phase. To familiarize participants with the task procedure, three practice trials were administered before the actual experiment.

After completing the self-paced reading task, participants proceeded to a reading span task designed to assess their working memory capacity, adapted from Daneman and Carpenter (1980).

This task comprised five sets of sentences, with three trials in each set. The number of sentences in each set ranged from 2 to 6, representing different reading span sizes. Thus, a total of 60 sentences (2 sentences per trial * 3 trials + 3 sentences per trial * 3 trials + 4 sentences per trial * 3 trials + 5 sentences per trial * 3 trials + 6 sentences per trial * 3 trials) was included in the reading span task. Participants saw one sentence at a time on the screen and read it aloud. After reading all the sentences in a trial, they were required to recall the last word of each sentence in the correct order. If they couldn’t remember a word, they were instructed to type “do not remember.” Only words typed correctly in both form and order were considered accurate and recorded. The total number of words recalled correctly was calculated for each participant. Following the reading span task, participants were directed to complete a background questionnaire.

2.1.4 Data analysis

The data obtained from the self-paced reading task and the reading span task underwent preprocessing before any statistical analysis. Initially, the reading span scores were standardized using z-scores. Regarding comprehension question choices, NP_{low} (indicating low attachment) was coded as 0, while NP_{high} (indicating high attachment) was coded as 1. In the data analysis phase, a generalized linear mixed-effects model was applied to investigate participants’ comprehension question responses, with “condition” (reflecting classifier-noun agreement: CL-NP_{low}; CL-NP_{high}; CL-NP_{low}/NP_{high}; No CL) and “reading span z-score” (indicating working memory capacity) as fixed factors. Furthermore, both “subject” and “item” were considered as random factors. Since “condition” was a four-level categorical predictor, we converted it into three contrasts. The first contrast compares “No CL” with “CL-NP_{low}” (No CL = -0.5, CL-NP_{low} = 0.5, CL-NP_{high} = 0, CL-NP_{low}/NP_{high} = 0). The second contrast compares “No CL” with “CL-NP_{high}” (No CL = -0.5, CL-NP_{high} = 0.5, CL-NP_{low} = 0, CL-NP_{low}/NP_{high} = 0). The third contrast compares “No CL” with “CL-NP_{low}/NP_{high}” (No CL = -0.5, CL-NP_{low}/NP_{high} = 0.5, CL-NP_{low} = 0, CL-NP_{high} = 0). We first created the most complex model with the maximal random-effects structure, justified by the design (Barr et al., 2013), as shown in (7).

```
(7) Complex model = glmer [cq_choice ~ condition +
  rs_zscore + (1 + condition|subject) + (1 + condition +
  rs_zscore|item), data = data, family = "binomial"
  (link = "logit")]
```

Convergence failures or singularity warnings were dealt with by progressively simplifying the random effects structure until convergence was reached or singularity warning disappeared (Barr et al., 2013; Linck and Cummings, 2015). For model selection, we employed a stepwise backward elimination method using the Akaike Information Criterion (AIC, Akaike, 1998). We began with the most complex model and reduced the model complexity until no further improvement in model fit is observed (Matuschek et al., 2017). Considering the design of Experiment 1, the variables of interest were “condition” and “reading span z-score.” These variables were included in the models as fixed factors and were retained throughout the model selection procedure. In addition, *post hoc* tests were performed with Bonferroni correction for any

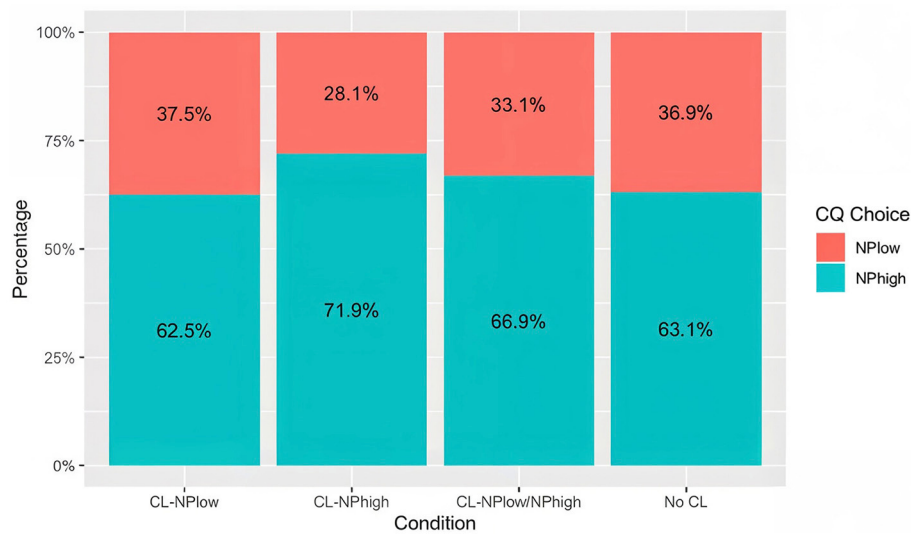


FIGURE 4
Percentage of comprehension question choices by conditions.

significant interactions that were identified. For details on the model selections and AIC values, please refer to the R script (find the link under the Data Availability section).

In preprocessing the reading times, we first eliminated outliers using absolute cutoffs, excluding reading times <100 ms or $>3,000$ ms (Luce, 1986; Roberts and Felsler, 2011), constituting 6.13% of the total data. Subsequently, we log-transformed the reading times, a standard practice to stabilize variance and achieve approximately normal distribution of residuals (Box and Cox, 1964). Linear mixed-effects models were then conducted on the log-transformed reading times in each region, considering “condition” and “reading span z-score” as fixed factors and “subject” and “item” as random factors. The same contrast-coding for “condition” was applied here. We first created the most complex model with the maximal random-effects structure, justified by the design (Barr et al., 2013), as shown in (8).

$$(8) \text{ Complex model} = \text{lmer}(\log\text{RT} \sim \text{condition} + \text{rs_zscore} + (1 + \text{condition} | \text{subject}) + (1 + \text{condition} + \text{rs_zscore} | \text{item}), \text{data} = \text{data})$$

Convergence failures, singularity warnings, model selections, and post hoc tests were managed following the same principles as previously outlined. Refer to the R script for model selections and AIC values. We utilized the lme4 package version 1.1-34 and the lmerTest package version 3.1-3 for analyses. All statistical analyses were conducted in R version 4.3.1 (R Core Team, 2023).

2.2 Results and discussion

2.2.1 Offline comprehension questions

The selected model for comprehension question choices is shown in (9), with an AIC value of 1942.7. The by-subject random slope was eliminated due to convergence issues.

$$(9) \text{ Selected model} = \text{glmer}[\text{cq_choice} \sim \text{condition} + \text{rs_zscore} + (1 | \text{subject}) + (1 + \text{rs_zscore} | \text{item}), \text{data} = \text{data}, \text{family} = \text{“binomial”} (\text{link} = \text{“logit”})]$$

The participants’ comprehension question choices in each condition are shown in Figure 4, and statistical results are given in Table 2. Figure 4 illustrates a prevailing NP_{high} (High-Attachment) preference across all conditions: 62.5% in the CL-NP_{low} condition, 71.9% in the CL-NP_{high} condition, 66.9% in the CL-NP_{low}/NP_{high} condition, and 63.1% in the No CL condition. Table 2 further revealed that the selection of NP_{high} in the CL-NP_{low} condition was significantly lower than that in the No CL condition ($\beta = -0.49$, $SE = 0.21$, $z = -2.37$, $p < 0.05$, $OR^2 = 0.62$, 95% CI³ $[-0.89, -0.08]$). However, the selection of NP_{high} in the CL-NP_{high} condition was significantly higher than that in the No CL condition ($\beta = 0.85$, $SE = 0.21$, $z = 3.40$, $p < 0.001$, $OR = 2.34$, 95% CI $[0.43, 1.27]$). No significant differences were observed between the CL-NP_{low}/NP_{high} condition and the No CL condition ($\beta = 0.09$, $SE = 0.21$, $z = 0.41$, $p = 0.680$, $OR = 1.09$, 95% CI $[-0.32, 0.49]$). In addition, there was no significant effect of working memory capacity (i.e., rs_zscore) on the comprehension question responses ($\beta = 0.19$, $SE = 0.23$, $z = 0.81$, $p = 0.416$, $OR = 1.21$, 95% CI $[-0.26, 0.64]$).

To better understand the impact of “condition” on participants’ comprehension responses, the data was also fitted with a model using treatment coding for the categorical predictor “condition”. Pairwise comparisons conducted via the emmeans package (version 1.10.2) revealed a significant difference between the CL-NP_{low} and CL-NP_{high} conditions. The selection of NP_{high} was significantly higher in the CL-NP_{high} condition than in the CL-NP_{low} condition ($\beta = -0.67$, $SE = 0.17$, $z = -3.89$, $p < 0.001$, $OR = 0.51$, 95% CI $[-1.01, -0.33]$).

2 OR, Odd Ratios.

3 CI: 95% Confidence intervals for the coefficients.

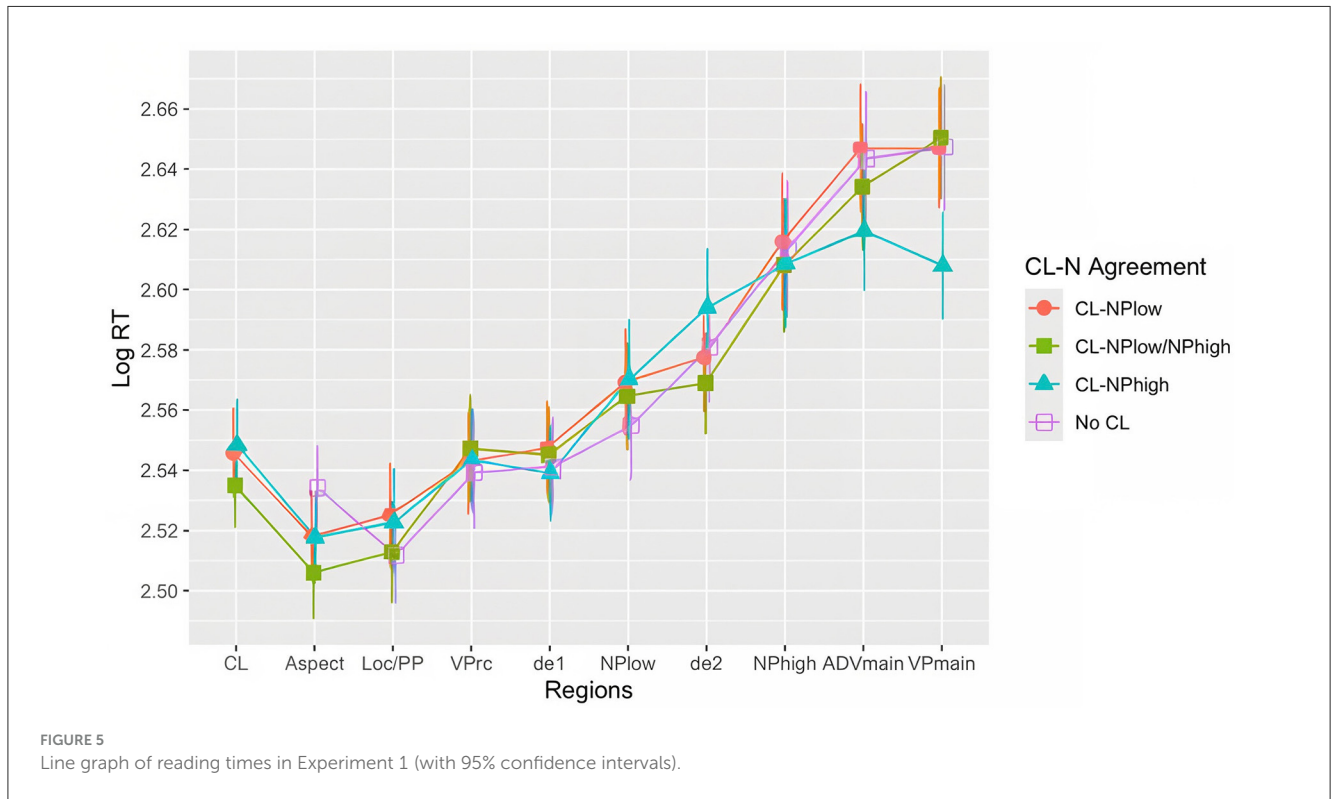
TABLE 2 Generalized linear mixed-effects model for comprehension question responses.

Fixed Effects	Estimate	SE	z-value	p-value	OR
Intercept	1.16	0.26	4.39	<0.001***	3.18
condition c1: No CL vs. CL-NP _{low}	-0.49	0.21	-2.37	0.018*	0.62
condition c2: No CL vs. CL-NP _{high}	0.85	0.21	3.40	<0.001***	2.34
condition c3: No CL vs. CL- NP _{low} /NP _{high}	0.09	0.21	0.41	0.680	1.09
rs_zscore	0.19	0.23	0.81	0.416	1.21

Selected model formula in R: glmer [cq_choice ~ condition+ rs_zscore + (1|subject) + (1 + rs_zscore| item), data = data, family = "binomial" (link = "logit")]. Random effects for subject ($\sigma^2 = 3.44$, SD = 1.86), item ($\sigma^2 = 0.49$, SD = 0.70), and rs_zscore by item ($\sigma^2 = 0.10$, SD = 0.32). SE, standard error; OR, odds ratios; Signif. codes: ***p < 0.001, *p < 0.05.

TABLE 3 The mean reading times and standard deviations in the six regions of interest (milliseconds).

Condition	de1	NP _{low}	de2	NP _{high}	ADV _{main}	VP _{main}
CL-NP _{low}	383.62 (183.79)	416.85 (249.54)	426.20 (261.51)	502.51 (410.65)	525.98 (392.29)	518.20 (395.60)
CL-NP _{high}	379.08 (195.28)	433.45 (320.86)	455.76 (328.73)	485.15 (389.93)	485.82 (352.79)	457.46 (298.05)
CL-NP _{low} /NP _{high}	384.39 (192.03)	413.59 (255.27)	411.01 (237.08)	490.85 (400.73)	511.29 (393.20)	521.89 (374.41)
No CL	382.57 (205.53)	408.56 (271.64)	437.06 (322.66)	500.58 (420.91)	531.72 (422.32)	521.56 (388.16)



2.2.2 Online reading times

Table 3 presents the means and standard deviations of raw reading times across four conditions in six regions of interest. Figure 5 displays a line graph depicting log-transformed reading times. Linear mixed-effects models were applied to the six regions of interest. The results showed no significant effects of “condition” or “reading span z-score” in the de1, NP_{low}, and NP_{high} regions. However, significant differences were observed between the “No CL” condition and the “CL-NP_{high}” condition in the spill-over

region of NP_{low} (i.e., de2), as well as in the two spill-over regions of NP_{high} (i.e., ADV_{main} and VP_{main}).

Table 4 presents the statistical results with the selected models using contrast-coding for each region. As Table 4 demonstrates, in the de2 region, reading times in the CL-NP_{high} condition were notably longer than those in the No CL condition ($\beta = 0.03$, $SE = 0.01$, $t = 2.20$, $p < 0.05$, $d = 0.174$,⁴ 95% CI [0.002,

⁴ d, Cohen's d.

TABLE 4 Linear mixed-effects models in de2, ADVmain, and VPmain regions.

Region	Fixed Effects	Estimate	SE	t-value	p-value
de2	Intercept	2.58	0.02	170.51	<0.001***
	condition c1: No CL vs. CL-NP _{low}	-0.01	0.01	-0.49	0.621
	condition c2: No CL vs. CL-NP _{high}	0.03	0.01	2.20	0.028*
	condition c3: No CL vs. CL- NP _{low} /NP _{high}	-0.02	0.01	-1.77	0.078
	rs_zscore	0.003	0.01	0.20	0.846
ADVmain	Intercept	2.64	0.01	180.64	<0.001***
	condition c1: No CL vs. CL-NP _{low}	0.02	0.02	1.40	0.162
	condition c2: No CL vs. CL-NP _{high}	-0.03	0.02	-2.08	0.037*
	condition c3: No CL vs. CL- NP _{low} /NP _{high}	-0.01	0.02	-0.34	0.734
	rs_zscore	-0.002	0.01	-0.13	0.901
VPmain	Intercept	2.64	0.01	202.90	<0.001***
	condition c1: No CL vs. CL-NP _{low}	0.02	0.01	1.18	0.239
	condition c2: No CL vs. CL-NP _{high}	-0.06	0.01	-4.09	<0.001***
	condition c3: No CL vs. CL- NP _{low} /NP _{high}	0.02	0.01	1.65	0.099
	rs_zscore	-0.01	0.01	-0.75	0.455

Selected model formula in R: (1) $de2_model = lmer(\log_de2 \sim condition + rs_zscore + (1|subject) + (1 + rs_zscore |item), data = data)$. Random effects for subject ($\sigma^2 = 0.02$, $SD = 0.13$), item ($\sigma^2 = 2.163e-04$, $SD = 0.01$), and rs_zscore by item ($\sigma^2 = 2.698e-05$, $SD = 0.01$). (2) $ADVmain_model = lmer(\log_adv_main \sim condition + rs_zscore + (1|subject), data = data)$. Random effects for subject ($\sigma^2 = 0.02$, $SD = 0.12$). (3) $VPmain_model = lmer(\log_vp_main \sim condition + rs_zscore + (1|subject) + (1|item), data = data)$. Random effects for subject ($\sigma^2 = 0.01$, $SD = 0.11$) and item ($\sigma^2 = 0.0002$, $SD = 0.01$). SE: standard error; Signif. codes: *** $p < 0.001$; * $p < 0.05$.

0.05)). In contrast, reading times in the CL-NP_{high} condition were significantly shorter than those in the No CL condition in the ADV_{main} region ($\beta = -0.03$, $SE = 0.02$, $t = -2.08$, $p < 0.05$, $d = -0.17$, 95% CI [-0.06, -0.002]), and the VP_{main} region ($\beta = -0.06$, $SE = 0.01$, $t = -4.09$, $p < 0.001$, $d = -0.33$, 95% CI [-0.09, -0.03]).

In addition, reading times in the six regions were fitted to models with treatment coding for the categorical predictor “condition” to explore pairwise comparisons. No significant differences emerged in the de1, NP_{low}, and NP_{high} regions. In the de2 region, reading times in the CL-NP_{high} condition were significantly longer than those in the CL-NP_{low}/NP_{high} condition ($\beta = 0.02$, $SE = 0.01$, $t = 2.42$, $p < 0.05$, $d = 0.16$, 95% CI [0.005, 0.04]). In the ADV_{main} region, reading times in the CL-NP_{high} were significantly shorter than in the CL-NP_{low} condition ($\beta = -0.03$, $SE = 0.01$, $t = -2.13$, $p < 0.05$, $d = -0.14$, 95% CI [-0.05, -0.002]). Similarly, in the VP_{main} region, reading times in the CL-NP_{high} were significantly shorter than in the CL-NP_{low} condition ($\beta = -0.04$, $SE = 0.01$, $t = -3.23$, $p < 0.01$, $d = -0.21$, 95% CI [-0.06, -0.02]) and the CL-NP_{low}/NP_{high} condition ($\beta = -0.04$, $SE = 0.01$, $t = -3.51$, $p < 0.01$, $d = -0.23$, 95% CI [-0.07, -0.02]). No other significant differences were found across levels in these regions. Additionally, no significant effect of “reading time z-score” was observed in any of the six regions.

In this experiment, we explored people’s preferences for RC attachment ambiguity, examining both online and offline processing phases. Additionally, we investigated whether an initial classifier pre-activates the semantic features of the upcoming noun, thereby influencing RC attachment preferences. To do this, we manipulated the agreement between the classifier and potential attachment sites (NP_{low} and NP_{high}), creating

four conditions: CL-NP_{low}, CL-NP_{high}, CL-NP_{low}/NP_{high}, and No CL.

The offline comprehension results indicated a prevalent preference for high attachment across all four conditions, revealing an intrinsic inclination toward high attachment during the offline processing phase. However, significant differences emerged: the selection of NP_{high} in the CL-NP_{high} condition was markedly higher than in the No CL condition, whereas the selection of NP_{high} in the CL-NP_{low} condition was significantly lower than in the No CL condition. These findings highlight the influence of classifier-noun agreement on people’s RC attachment choices. The CL-NP_{high} agreement increased the intrinsic high-attachment preference, while the CL-NP_{low} agreement decreased it. Consequently, a significant difference was observed between the CL-NP_{low} and CL-NP_{high} conditions in terms of comprehension question choices. Additionally, no significant difference was observed between the CL-NP_{low}/NP_{high} condition and the No CL condition, further confirming that the group classifier *qun* “CL(group)” in the CL-NP_{low}/NP_{high} condition was ambiguous, functioning similarly to the No CL condition.

These findings were further confirmed in the online processing phase. Before delving into the online data, it’s worth noting again that the word order of relative clauses in Mandarin is reversed compared to English. In Mandarin, the RC precedes the complex head noun, and the low-attachment site (NP_{low}) precedes the high-attachment site (NP_{high}). These features of Chinese RC structure suggest that when encountering NP_{low} first, the competitor (NP_{high}) has not yet appeared. Therefore, it would be inappropriate to describe low attachment as being “preferred” at this point in an online study of parsing ambiguity since there is no attachment conflict at this early stage. Therefore, we followed

Kamide and Mitchell (1997, 1999) in attributing this initial “low-attachment” to the incremental processing effect (Kamide and Mitchell, 1997, 1999). However, we do use the term “preference” for the high-attachment site if it turns out to attract attachment, as both attachment sites are now present, and readers can demonstrate their preferences either by maintaining the initial low-attachment or shifting to high-attachment.

Analyses of the self-paced reading times revealed that readers pre-activate semantic features of the upcoming nouns when they encounter the classifiers. These predictions are maintained until they encounter the two potential attachment nouns, affecting how people process these sites and modulating their RC attachment preferences.

Specifically, in the spill-over region of NP_{low} (i.e., de2), reading times in the CL-NP_{high} condition were significantly longer than in the No CL condition. In contrast, in the spill-over regions of NP_{high} (i.e., ADV_{main} and VP_{main}), reading times in the CL-NP_{high} condition were significantly shorter than in the No CL condition. These results indicate that when the sentence-initial classifier is CL(animal) (e.g., 只, *zhi*), it pre-activates the semantic features of an animal for the upcoming noun, and this prediction persists until comprehenders encounter the first potential noun, NP_{low}(people) (e.g., 义工, volunteer), which conflicts with the anticipation and thus creates processing difficulty. This is evidenced by the significantly increased reading times in the spill-over region compared with the control condition. The incongruence between the CL(animal) (e.g., 只, *zhi*) and NP_{low}(people) (e.g., 义工, volunteer) causes comprehenders’ expectation for an animal noun to persist until encountering the second potential noun, NP_{high}(animal) (e.g., 小猫, cat). The pre-activated semantic features of an animal match with NP_{high}(animal) (e.g., 小猫, cat), and the congruence between CL(animal) and NP_{high}(animal) facilitates processing, manifesting as significantly decreased reading times in the spill-over regions compared to the control condition.

On the other hand, when the classifier is CL(people) (e.g., 位, *wei*), it pre-activates the semantic features of people for the upcoming noun, persisting until the first potential noun NP_{low}(people) (e.g., 义工, volunteer), which matches well with the classifier’s anticipation. As a result, the NP_{low}(people) (e.g., 义工, volunteer) integrates smoothly with the preceding sentence, and no processing difficulty emerge in the NP_{low} and its spill-over de2 regions. After this integration, when comprehenders encounter an additional noun, NP_{high}(animal) (e.g., 小猫, cat), this additional noun creates integration difficulty, resulting in increased reading times. Therefore, the sentence-initial classifiers provide semantically constraining contexts for readers to make predictions about the upcoming noun, thus creating forced-low-attachment or forced-high-attachment conditions for RC interpretation.

In addition, the globally ambiguous conditions, namely the “No CL” and “CL-NP_{low}/NP_{high}” conditions, allow us to observe the locus of people’s intrinsic attachment preferences from the online phase to the offline phase. In these two conditions, self-paced reading times in the NP_{low} region indicate an initial low-attachment during early online processing due to “incremental pre-head attachment parsing” (Kamide and Mitchell, 1999, p. 635). Offline comprehension questions, however, show a

high-attachment preference. This inconsistency between the initial online low-attachment and the eventual offline high-attachment preference suggests that the initial low attachment is reconsidered during later online processing, influenced by discourse or post-syntactic information available later in the sentence, consistent with previous studies (de Vincenzi and Job, 1993; Kamide and Mitchell, 1997, 1999; Fernández, 2002).

This dynamic preference pattern has also been confirmed by the two disambiguated conditions, i.e., CL-NP_{low} and CL-NP_{high} conditions. According to Fernández (2002), readers’ attachment preferences in the online processing phases would manifest as increased reading times at critical regions when encountering materials that contain the dis-preferred forced attachment, as compared to reading times associated with either ambiguous materials or those containing the preferred forced attachment (Fernández, 2002). In this experiment, we observed that in the critical region, NP_{high} and its spill-over regions (ADV_{main} and VP_{main}), reading times in the forced-high-attachment condition (i.e., CL-NP_{high}) were significantly shorter than in the forced-low-attachment condition (CL-NP_{low}) and the ambiguous condition (No CL), indicating a high-attachment preference in the later online processing stage. Therefore, readers initially show low-attachment, which then switches to a high-attachment preference in the later stage of online processing due to discourse or post-syntactic information. This high-attachment preference remains until the offline phase.

3 Experiment 2

In Experiment 2, we further investigated the prediction effects of classifiers in resolving RC attachment ambiguity, particularly when an additional disambiguating cue, namely semantic compatibility, was introduced. This is because the RCs can be attached to both the low and the high NPs in Experiment 1, thus making it challenging to ensure how participants interpret these RCs in addition to relying on the semantic cues on the classifiers. Our aim was to discern the relative significance of each cue in predicting the upcoming noun and shaping attachment preferences.

Following the syntactic structure of Experiment 1, we introduced semantic compatibility as an additional variable alongside classifier-noun agreement in Experiment 2. Semantic compatibility was manipulated between the RC and its potential attachment sites, where each RC was exclusively compatible with either NP_{low} or NP_{high}. The variable of “semantic compatibility” therefore has two levels: that the RC is semantically compatible with NP_{low} (RC-NP_{low}) or that the RC is semantically compatible with NP_{high} (RC-NP_{high}). We also manipulated the classifier-noun agreement, including CL-NP_{low}, CL-NP_{high}, CL-NP_{low}/NP_{high}, and No CL levels. Thus, Experiment 2 had a 2*4 design incorporating two independent variables: semantic compatibility (2 levels) and classifier-noun agreement (4 levels), resulting in eight experimental conditions initially. However, due to the issue of crossing dependency, one of the conditions (RC-NP_{high} & CL-NP_{low}) was not possible and was eliminated. To maintain a factorial design, we therefore excluded the level of CL-NP_{low}. Consequently,

TABLE 5 An experimental item in the six conditions.

Semantic compatibility	CL-N agreement		Test item				
	Dem-CL	RC(Aspect/Loc/VPrc)	de1	NP _{low}	de2	NP _{high}	Adv _{main} /NP _{main} /NP _{main}
RC-NP _{high}	CL-NP _{high}	正在/院子里/舔爪子/licking its claws in the yard	的/rel.	义工/volunteer	的/poss.	小猫/cat	前面/有一排/房子。there is a row of houses in front of.
	CL-NP _{low} /NP _{high}	这群/This-CL (group)					
	No CL (control)	—					
RC-NP _{low}	CL-NP _{high}	正在/院子里/做手工/making handicrafts in the yard	的/rel.	义工/volunteer	的/poss.	小猫/cat	前面/有一排/房子。there is a row of houses in front of.
	CL-NP _{low} /NP _{high}	这只/This-CL (animal)					
	No CL (control)	—					

English translation for sentences above: there is a row of houses in front of the volunteer that is licking its claws in the yard. English translation for sentences below: There is a row of houses in front of the volunteer that is making handicrafts in the yard.

Experiment 2 adopted a 2*3 design with two independent variables: semantic compatibility (RC-NP_{low} and RC-NP_{high}) and classifier-noun agreement (CL-NP_{high}, CL-NP_{low}/NP_{high}, and No CL), establishing six experimental conditions. Table 5 illustrates an exemplar item in these six conditions.

We predict that when the RC aligns semantically with NP_{high} (i.e., RC-NP_{high}) and the classifier corresponds with NP_{high} (i.e., CL-NP_{high}), the coherence between semantic compatibility and classifier-noun agreement promotes a strong prediction for NP_{high}, resulting in a high-attachment preference. However, complications arise when the RC is semantically compatible with NP_{low} (i.e., RC-NP_{low}), but the classifier aligns with NP_{high} (i.e., CL-NP_{high}), generating a conflict in predictions of the upcoming noun. In such scenarios, the dominance of one cue prevails, establishing either a low-attachment preference or a high-attachment preference. Finally, in cases where the classifier is ambiguous (CL-NP_{low}/NP_{high}) or absent (No CL), the RC attachment preference is substantially predicted by the semantic compatibility between the RC and the potential noun phrases.

3.1 Methods

3.1.1 Participants

Sixty native Mandarin Chinese speakers aged between 18 and 30 participated in this experiment. Participants received compensation for their involvement, and informed consent was obtained from all participants.

3.1.2 Materials

Twenty-eight experimental items were created, all following the same syntactic structure and segmented into eleven regions, as illustrated in Table 6. Among these regions, six critical regions of interest were identified: de1, NP_{low}, de2, NP_{high}, ADV_{main}, and VP_{main}. All the experimental items were counterbalanced using a Latin-square design. Each participant was presented with one version of each experimental item. The experiment incorporated the same set of 56 fillers utilized in Experiment 1.

In line with the paradigm in experiment 1, each target sentence was introduced by a referential discourse context. To ensure the exclusive semantic compatibility of RCs with either NP_{low} or NP_{high}, the materials used in this experiment underwent validation using the same norming tests employed in Experiment 1. Participants utilized a 7-point Likert scale, ranging from 1 denoting complete unacceptability to 7 indicating full acceptability. The results of these acceptability judgments revealed a clear pattern: half of the RCs demonstrated semantic compatibility with NP_{high} ($M = 6.78$; $SD = 0.24$) but not NP_{low} ($M = 1.20$; $SD = 0.21$) [$t_{(27)} = 98.09$, $p < 0.001$]. Conversely, the other half of the RCs exhibited compatibility with NP_{low} ($M = 6.80$; $SD = 0.23$) but not NP_{high} ($M = 1.27$; $SD = 0.25$) [$t_{(27)} = 88.09$, $p < 0.001$].

3.1.3 Procedure and data analysis

In this experiment, we employed a moving-window self-paced reading task created using jsPsych Version 6.2 (de Leeuw et al., 2023) and executed on the Cognition.run platform. The

TABLE 6 The average reading times and standard deviations in the six regions of interest (milliseconds).

RC semantics	CL-N agreement	de1	NP _{low}	de2	NP _{high}	ADV _{main}	VP _{main}
RC-NP _{high}	CL-NP _{high}	381.61 (149.56)	443.13 (293.98)	476.99 (302.22)	468.37 (299.16)	484.00 (321.06)	460.36 (235.32)
	CL-NP _{low} /NP _{high}	396.93 (187.37)	446.97 (273.58)	442.51 (263.62)	479.30 (366.31)	500.27 (350.73)	470.17 (292.66)
	No CL	393.30 (188.51)	431.53 (274.85)	418.94 (231.93)	485.41 (361.01)	503.39 (381.16)	498.10 (325.66)
RC-NP _{low}	CL-NP _{high}	394.36 (205.18)	400.22 (223.33)	412.63 (270.04)	480.24 (369.92)	497.69 (363.32)	497.90 (342.77)
	CL-NP _{low} /NP _{high}	389.31 (189.23)	421.22 (268.00)	421.44 (265.66)	486.92 (386.73)	525.34 (367.33)	496.66 (295.11)
	No CL	389.05 (176.11)	411.07 (242.38)	416.67 (241.79)	507.84 (428.44)	552.19 (424.63)	533.82 (365.39)

experimental procedure followed that of Experiment 1. Prior to any statistical analysis, the data collected from both the self-paced reading task and the reading span task underwent preprocessing. Initially, reading span scores were standardized into z-scores. The accuracy rate for comprehension questions averaged 97.91% ($SD = 2.68$). Outliers were then removed using absolute cutoffs, excluding reading times <100 ms or $>3,000$ ms (Luce, 1986; Roberts and Felser, 2011), accounting for 2.61% of the total data. Subsequently, reading times were log-transformed. Linear mixed-effects models were applied to log-transformed reading times, incorporating “semantic compatibility” and “CL-N agreement” as interacting predictors, with “reading span z-score” as a co-predictor. Since “CL-N agreement” was a three-level categorical predictor, we converted it into two contrasts: the first compared “No CL” with “CL-NP_{high}” (No CL = -0.5 , CL-NP_{high} = 0.5 , CL-NP_{low}/NP_{high} = 0), and the second compared “No CL” with “CL-NP_{low}/NP_{high}” (No CL = -0.5 , CL-NP_{low}/NP_{high} = 0.5 , CL-NP_{high} = 0). In addition, “semantic compatibility” was a two-level categorical predictor, and it was contrast-coded (RC-NP_{low} = 0.5 , RC-NP_{high} = -0.5). We first created the most complex model with the maximal random-effects structure, justified by the design (Barr et al., 2013), as shown in (10).

$$(10) \text{ Complex model} = \text{lmer} [\log RT \sim \text{cl_n_agreement} * \text{rc_semantics} + \text{rs_zscore} + (1 + \text{cl_n_agreement} * \text{rc_semantics} | \text{subject}) + (1 + \text{cl_n_agreement} * \text{rc_semantics} + \text{rs_zscore} | \text{item}), \text{data} = \text{data}]$$

Convergence failures, model selections, and *post hoc* tests were managed following the same principles as in Experiment 1. Considering the design of Experiment 2, the variables of interest were “classifier-noun agreement”, “semantic compatibility”, and “reading span z-score.” These variables were included in the models as fixed factors and were retained throughout the model selection procedure.

3.2 Results and discussion

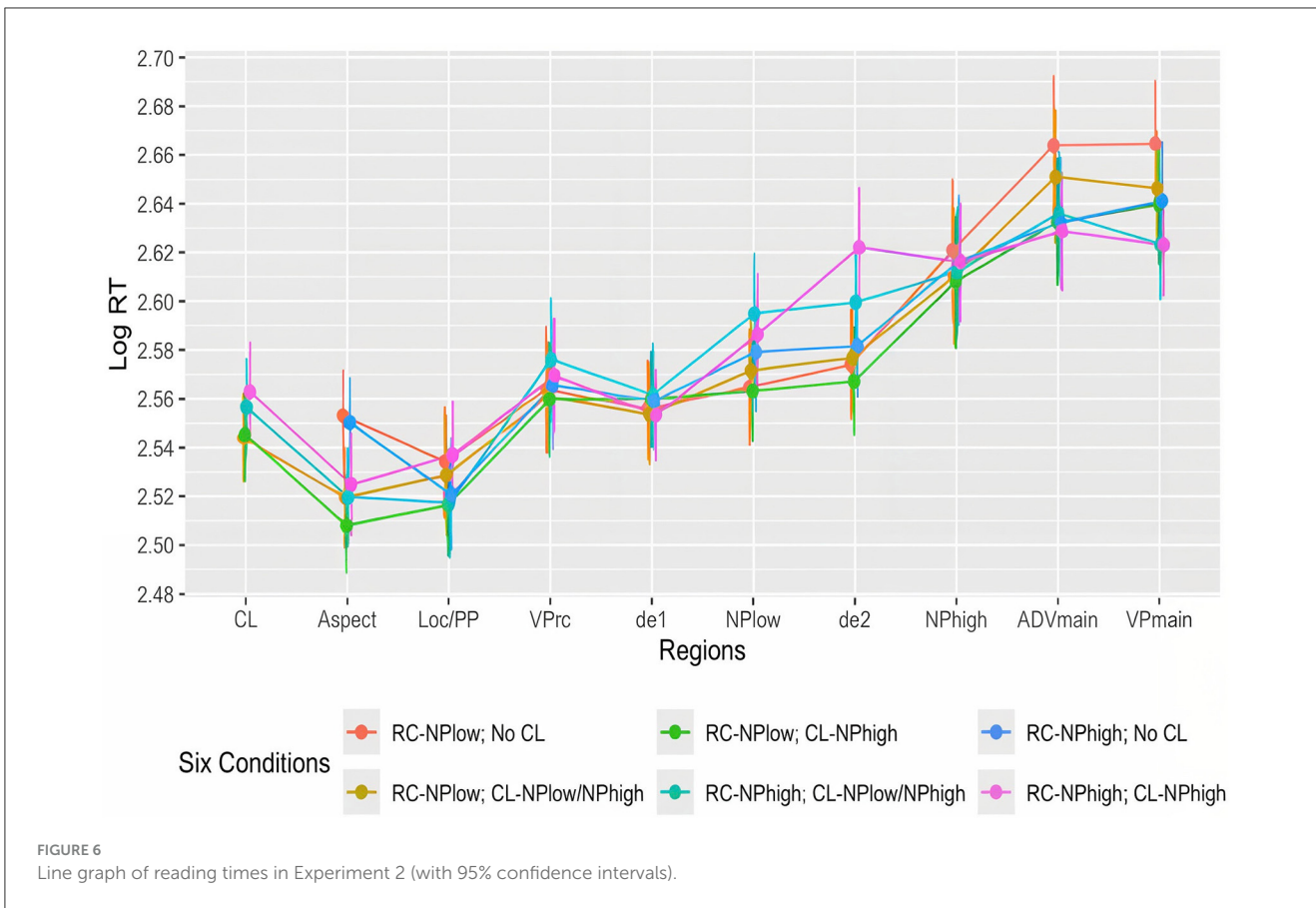
Table 6 displays the means and standard deviations of raw reading times across six conditions in six regions of interest. Meanwhile, Figure 6 presents a line graph illustrating log-transformed reading times. Linear mixed-effects models were fitted in the six regions of interest.

The results showed no significant effects of “classifier-noun agreement”, “semantic compatibility”, or “reading span z-score”, nor interaction effects in the de1 and NP_{high} regions ($ps > 0.05$). However, a significant effect of “semantic compatibility” was observed in the NP_{low} region, its spill-over de2 region, as well as in the spill-over region of NP_{high} (VP_{main}). Table 7 presents the statistical results from selected models using the contrast-coding approach (also see Figure 7 for illustration).

In the NP_{low} region, reading times in the RC-NP_{high} condition were significantly longer than those in the RC-NP_{low} condition ($\beta = -0.02$, $SE = 0.01$, $t = -2.81$, $p < 0.01$, $d = -0.15$, 95% CI [-0.04 , -0.01]). No significant effects of “classifier-noun agreement” or “reading span z-score” were observed, and there was no interaction effect in this region.

Moving to the spill-over de2 region, reading times in the RC-NP_{high} condition were significantly longer than those in the RC-NP_{low} condition ($\beta = -0.03$, $SE = 0.01$, $t = -3.60$, $p < 0.001$, $d = -0.20$, 95% CI [-0.05 , -0.01]). In addition, an interaction effect between “classifier-noun agreement” and “semantic compatibility” was observed in the de2 region. To further explore the interaction effect, the data was fitted with a model using treatment coding for the two categorical predictors, and pairwise comparisons between conditions were conducted. Results showed that reading times in the “RC-NP_{high} CL-NP_{high}” condition were significantly longer than those in the “RC-NP_{low} No CL” condition ($\beta = -0.05$, $SE = 0.01$, $t = -3.72$, $p < 0.01$, $d = -0.33$, 95% CI [-0.08 , -0.02]), the “RC-NP_{low} CL-NP_{low}/NP_{high}” condition ($\beta = -0.05$, $SE = 0.01$, $t = -3.45$, $p < 0.01$, $d = -0.31$, 95% CI [-0.07 , -0.02]), the “RC-NP_{low} CL-NP_{high}” condition ($\beta = -0.06$, $SE = 0.01$, $t = -4.29$, $p < 0.001$, $d = -0.38$, 95% CI [-0.08 , -0.03]), and the “RC-NP_{high} No CL” condition ($\beta = 0.04$, $SE = 0.01$, $t = 3.16$, $p < 0.05$, $d = 0.27$, 95% CI [0.02 , 0.06]). Moreover, a marginally significant effect of “reading span z-score” was observed in the de2 region ($\beta = -0.03$, $SE = 0.01$, $t = -1.88$, $p = 0.065$, $d = -0.18$, 95% CI [-0.05 , 0.001]).

In the first spill-over region of NP_{high}, the ADV_{main} region showed a marginally significant effect of “semantic compatibility.” Reading times in the RC-NP_{high} condition were shorter than those in the RC-NP_{low} condition ($\beta = 0.02$, $SE = 0.01$, $t = 1.71$, $p = 0.088$, $d = 0.09$, 95% CI [-0.002 , 0.03]). This trend became more pronounced in the second spill-over region of NP_{high}, the VP_{main} region, where a significant main effect of “semantic compatibility” was observed. Reading times in the RC-NP_{high} condition were significantly shorter than those in the RC-NP_{low} condition ($\beta = 0.02$, $SE = 0.01$, $t = 2.23$, $p < 0.05$, $d = 0.12$, 95% CI [0.002 , 0.04]). No other effects were identified in this region.



In this experiment, we examined the impact of two disambiguating cues, namely classifier-noun agreement and the semantic compatibility between the RC and the head noun, on influencing RC attachment preferences. We manipulated the relationship between the prenominal classifier and two potential head nouns, as well as the relationship between RC semantics and the potential head nouns.

Our findings revealed a prediction effect of RC semantics for upcoming nouns, affecting how people process the two potential attachment sites and thus modulating RC attachment preferences. Specifically, a significant main effect of semantic compatibility was observed in the NP_{low} region, the spill-over region of NP_{low} (i.e., de2), and the spill-over regions of NP_{high} (i.e., ADV_{main} and VP_{main}). In the NP_{low} and de2 regions, reading times in the RC-NP_{high} condition were significantly longer than in the RC-NP_{low} condition. Conversely, in the ADV_{main} and VP_{main} regions, reading times in the RC-NP_{high} condition were significantly shorter than in the RC-NP_{low} condition.

These results indicate that the semantics embedded in relative clauses provides a constraining context that predicts a specific upcoming noun. Despite the classifier-noun agreement, when the RC semantics is compatible with NP_{high}, which is an animal noun, it pre-activates the semantic features of an animal for the upcoming noun. This prediction clashes with the semantics of the first potential noun, NP_{low} (people), creating processing difficulty and resulting in slower reading times in the NP_{low} and de2 regions. This incongruence maintains the anticipation for an animal noun

until the second potential noun, NP_{high} (animal) appears. The pre-activated semantic expectation of an animal then matches with the NP_{high} (animal), facilitating its integration into the preceding sentence and manifesting as fast reading times in the spill-over regions of NP_{high}.

On the other hand, despite the classifier-noun agreement, when the RC semantics is compatible with NP_{low}, it leads to the anticipation of a human noun, and this prediction matches with the first potential noun NP_{low} (people), facilitating the integration of NP_{low} (people) with the preceding RC, which is evidenced by fast reading times in the NP_{low} and de2 regions. When readers encounter an additional noun, NP_{high} (animal), it creates integration difficulty, resulting in slow reading times in the spill-over regions of NP_{high}.

Furthermore, we observed an interaction effect between classifier-noun agreement and semantic compatibility in the de2 region. Reading times in the condition where both RC semantics and the classifier agreed with NP_{high} (i.e., “RC-NP_{high} CL-NP_{high}”) were significantly longer than those in the condition where RC semantics were compatible with NP_{high} but there was no sentence-initial classifier (i.e., “RC-NP_{high} No CL”). This suggests that both RC semantics and sentence-initial classifiers contribute to predictions for the upcoming noun. The prediction effect is notably stronger when both cues contradict with the semantics of the closest NP, namely NP_{low}, compared to when only RC semantics makes the prediction, underscoring the importance

TABLE 7 Linear mixed-effects models in NP_{low}, de2, and VP_{main} regions.

Region	Fixed Effects	Estimate	SE	t-value	p-value
NP _{low}	Intercept	2.58	0.02	158.35	<0.001***
	rc_semantics	-0.02	0.01	-2.81	<0.01**
	cl_n_agreement c1: No CL vs. CL-NP _{high}	-0.005	0.01	-0.45	0.652
	cl_n_agreement c2: No CL vs. CL-NP _{low} /NP _{high}	0.01	0.01	1.47	0.143
	rs_zscore	-0.03	0.02	-1.64	0.107
	rc_semantics: cl_n_agreement c1	-0.01	0.02	-0.32	0.748
	rc_semantics: cl_n_agreement c2	-0.004	0.02	-0.19	0.853
de2	Intercept	2.59	0.01	184.58	<0.001***
	rc_semantics	-0.03	0.01	-3.60	<0.001***
	cl_n_agreement c1: No CL vs. CL-NP _{high}	0.01	0.01	1.41	0.160
	cl_n_agreement c2: No CL vs. CL-NP _{low} /NP _{high}	0.003	0.01	0.32	0.753
	rs_zscore	-0.03	0.01	-1.88	0.065
	rc_semantics: cl_n_agreement c1	-0.05	0.02	-2.62	<0.01**
	rc_semantics: cl_n_agreement c2	0.01	0.02	0.66	0.509
VP _{main}	Intercept	2.64	0.01	216.39	<0.001***
	rc_semantics	0.02	0.01	2.23	0.029*
	cl_n_agreement c1: No CL vs. CL-NP _{high}	-0.02	0.01	-1.44	0.151
	cl_n_agreement c2: No CL vs. CL-NP _{low} /NP _{high}	-0.01	0.01	-0.72	0.470
	rs_zscore	-0.02	0.01	-1.41	0.163
	rc_semantics: cl_n_agreement c1	-0.01	0.02	-0.31	0.760
	rc_semantics: cl_n_agreement c2	0.004	0.02	0.16	0.874

The best-fitting model formula in R: NP_{low}: lmer(log_np1 ~ cl_n_agreement * rc_semantics + rs_zscore + (1 + rc_semantics | subject) + (1 | item), data = data). Random effects for subject ($\sigma^2 = 0.02$, SD = 0.13), rc_semantics by subject ($\sigma^2 = 6.533e-04$, SD = 0.03), and item ($\sigma^2 = 1.886e-05$, SD = 0.004). de2: lmer(log_de2 ~ cl_n_agreement * rc_semantics + rs_zscore + (1 + rc_semantics | subject), data = data). Random effects for subject ($\sigma^2 = 0.01$, SD = 0.11) and rc_semantics by subject ($\sigma^2 = 0.0009$, SD = 0.03). VP_{main}: lmer(log_vp_main ~ cl_n_agreement * rc_semantics + rs_zscore + (1 + rc_semantics | subject), data = data). Random effects for subject ($\sigma^2 = 0.01$, SD = 0.10) and rc_semantics by subject ($\sigma^2 = 0.001$, SD = 0.03). Signif. codes: ***p < 0.001; **p < 0.01; *p < 0.05.

of classifiers' prediction effects. Moreover, reading times in the condition where both RC semantics and the classifier agreed with NP_{high} (i.e., "RC-NP_{high} CL-NP_{high}") were significantly longer than in the condition where RC semantics were compatible with NP_{low} but the classifier agreed with NP_{high} (i.e., "RC-NP_{low} CL-NP_{high}"), suggesting the significance of RC semantics prediction effects.

These results indicate that when both cues were present, RC semantic compatibility took precedence over classifier-noun agreement in resolving RC attachment ambiguities, as the classifier effect was only observed in the de2 region, whereas the RC semantic effect was observed in NP_{low}, NP_{low}'s spill-over region de2, and NP_{high}'s spill-over regions, ADV_{main} and VP_{main}. Essentially, RC semantic compatibility emerged as the predominant predicting factor in disambiguating RC attachment resolutions. However, this does not necessarily mean that the classifiers' prediction effect observed in Experiment 1 disappeared. An interaction effect between semantic compatibility and classifier-noun agreement was observed in the de2 region, suggesting that the classifiers' prediction effect did not disappear but instead diminished.

4 General discussion

4.1 Classifier-noun agreement and its prediction effect

The first research question investigates whether classifier-noun agreement can function as a disambiguating cue in resolving relative clauses that contain a head noun embedded within a complex nominal phrase consisting of two noun phrases.

The study found that classifier-noun agreement in Mandarin can disambiguate relative clauses to create forced-high/low-attachment conditions (Fernández, 2002). This finding aligns with previous studies that utilized morphosyntactic dependencies to resolve relative clause attachment ambiguities. For example, Fernández (2002) used number agreement in English, de Vincenzi and Job (1993) used gender agreement in Italian, and Lee and Kweon (2004) used honorific agreement in Korean.

We found that readers use classifier-noun agreement to resolve ambiguous RC attachments. The predictive role of classifiers influences how people anticipate and integrate upcoming information. Specifically, when a prenominal classifier

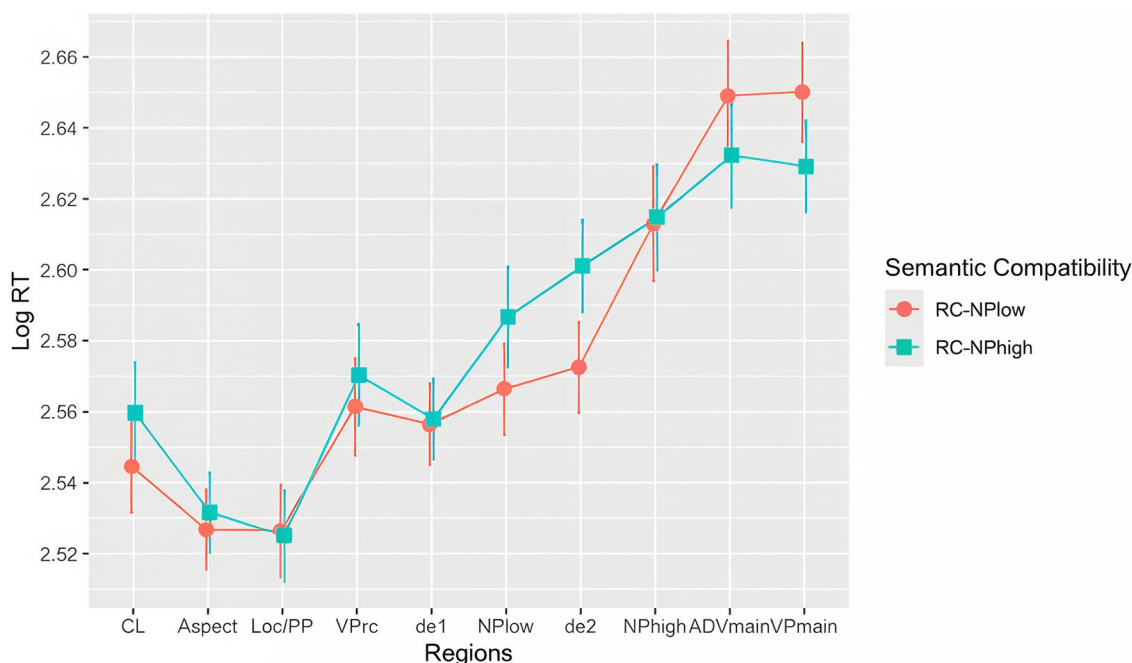


FIGURE 7

Line graph of reading times with 95% confidence intervals illustrating the semantic compatibility effect.

is encountered, it pre-activates specific semantic features of the upcoming noun. For example, if the classifier is CL(animal) (e.g., 只, *zhi*), it primes the comprehender to expect an animal noun. This pre-activation persists until the expected noun is encountered, guiding the comprehender's expectations and influencing processing efficiency. When the pre-activated semantic features triggered by the classifier and the subsequent noun are congruent (e.g., a classifier predicting an animal followed by a noun denoting an animal), processing becomes faster. This congruence reduces cognitive load, as evidenced by shorter reading times and easier integration into the preceding sentence.

Conversely, when the pre-activated semantic features and the subsequent noun are incongruent (e.g., a classifier predicting animals followed by a noun denoting a person), this mismatch creates processing difficulty. Readers experience increased reading times and cognitive load as they work to reconcile the unexpected noun with their initial predictions. Furthermore, the expectation triggered by the classifier can persist beyond the immediate noun. If the first encountered noun does not match the classifier, the anticipation continues until a matching noun is found.

The observed classifier prediction effect in this study aligns with predictive processing in sentence comprehension, wherein the linguistic features of upcoming words are pre-activated before the bottom-up input of the actual word (Kuperberg and Jaeger, 2016; Pickering and Gambi, 2018). Comprehenders utilize sentence-initial classifiers to provide contextual information, making predictions by pre-activating the semantic features of the upcoming noun before it is encountered (Kwon et al., 2017; Grüter et al., 2019; Chow and Chen, 2020; Huang et al., 2023). This pre-activation enables the brain to anticipate forthcoming linguistic input, promoting faster and more efficient

processing (Kuperberg and Jaeger, 2016; Pickering and Gambi, 2018).

In addition, the prediction triggered by classifiers persists until the expected noun is encountered, even if the expected noun is distant from the classifier due to an inserted RC. Upon encountering the expected noun, the classifier's prediction effect facilitates its integration, though the facilitative effect is delayed and observed in the spill-over region of the critical noun. This phenomenon was also noted in Wu et al. (2018).

However, the classifier-noun agreement effect diminished when it co-occurred with a strong disambiguating cue, RC semantics. People primarily rely on semantic information embedded in relative clauses to anticipate and ease the integration of upcoming nouns. In Experiment 2, both the sentence-initial classifier and the relative clause provided disambiguating information, with both preceding the two potential noun sites. This scenario led to both the classifier and the RC making predictions about the upcoming noun. Results indicated that RC semantics played a dominant role in predicting the upcoming noun, as it provided a highly-constraining context that enabled people to make fine-grained predictions, consistent with previous findings (Federmeier and Kutas, 1999; Kamide et al., 2003; Wang et al., 2018).

While RC semantics plays a dominant role in resolving RC attachment ambiguities, the classifier's prediction effect, though diminished, remains influential, as the study observed an interaction effect between RC semantics and classifier-noun agreement. When both the classifier and the RC semantics predict the same type of noun, the prediction effect is stronger, leading to more efficient processing. However, when predictions from different cues conflict, processing becomes more challenging. This finding is consistent with Kwon et al. (2017), which explored

classifiers' prediction effects in a highly semantically constraining context provided by a relative clause. Their study suggests that readers pre-activate semantic features of the upcoming noun when they encounter the relative clause and the classifier, comparing these pre-activated features from both sources. When there is incongruence between them, an N400 was elicited, indicating that even in highly constraining contexts, the classifier's prediction effect is also activated.

On the other hand, classifier-noun agreement also influences comprehension question decisions in the offline processing phase. Experiment 1 demonstrated a significant increase in the selection of NP_{high} as the RC attachment site when there was agreement between CL (animals) and NP_{high} (animals). Conversely, there was a significant decrease in the selection of NP_{high} as the attachment site when there was agreement between CL (people) and NP_{low} (people), indicating a robust classifier-noun agreement effect in manipulating RC attachment preferences in offline phases.

Therefore, classifier-noun agreement can be used to resolve RC attachment ambiguities in both online and offline processing phases, influencing readers to generate either high-attachment or low-attachment preferences. With this in mind, we propose that classifier-noun agreement can function similarly to morphosyntactic agreement (e.g., gender or number) in Indo-European languages in disambiguating RC attachment, allowing us to observe attachment preferences during both the online and offline phases (Fernández, 2002). This will be discussed further in the next section.

4.2 Relative clause attachment preferences

To address the research question concerning the preference for relative clause attachment in Mandarin Chinese, our study found a low attachment during the early stage of online processing. This early low-attachment was prominent in the critical region of NP_{low} and its immediate spill-over region, de2. This initial low attachment can be explained by the effect of incremental processing, as observed in Japanese (Kamide and Mitchell, 1997, 1999).

Furthermore, our study suggests a shift in attachment preference during the later stage of online processing due to discourse or post-syntactic information. Specifically, a high-attachment preference surfaced in the spill-over regions of NP_{high}, namely ADV_{main} and VP_{main}. This shift indicated a transition from the initial low-attachment to a late-stage high-attachment preference, a pattern also observed in other languages such as Spanish (Fernández, 2002), Italian (de Vincenzi and Job, 1993), and Japanese (Kamide and Mitchell, 1997). A similar pattern was evident in the study conducted by Kamide and Mitchell (1997) in Japanese, where they manipulated pragmatic compatibility between the RC and potential attachment sites. Their findings, based on a self-paced reading task segmented into seven regions (RC/NP1-Low/Gen/NP2-High/Acc/Matrix subject/Final), demonstrated significantly shorter reading times in the regions associated with NP1-Low and Gen in the forced low-attachment condition compared to the forced high-attachment condition. Notably, the Final region in their study exhibited substantially

shorter reading times in the forced high-attachment condition than in the forced low-attachment condition. These results suggested an initial low attachment when encountering the first potential attachment site (NP1-Low), followed by a shift to high attachment toward the end of the sentence, similar to the findings in our present study.

However, our findings contrast with those of Shen (2006), who observed a low-attachment preference in both the late online processing phase and the offline phase. Shen (2006) employed a self-paced reading task and an offline questionnaire to investigate RC attachment preferences in Mandarin. She manipulated the semantic compatibility between the RC and the two possible attachment sites, resulting in three conditions: ambiguous, forced-high, and forced-low conditions. The self-paced reading results indicated no significant differences between the three conditions regarding reading times at the low attachment site (NP_{low}), suggesting the absence of attachment preferences in the early online processing phase. However, at the high-attachment site (NP_{high}), reading times in the forced-high condition were significantly longer than those in the forced-low and ambiguous conditions, indicating a low-attachment preference in the late online processing phase. Furthermore, this low-attachment preference persisted into the offline phase, as demonstrated by the offline questionnaire responses (71.88% for low-attachment response).

In the following, we discuss three potential discourse/post-syntactic factors that may contribute to the observed shift to high-attachment preference in our study. First, we consider the semantic effect. In our materials, NP_{low} consistently referred to people while NP_{high} invariably denoted animals. One possibility for the observed high attachment can be attributed to people preferring animals over human beings for RC attachment, prompting a tendency to attach RCs to the distant NP_{high} during late online processing and offline phases, thus resulting in high attachment preferences. However, this possibility is unlikely since human nouns are more prominent in the animacy hierarchy and should be a stronger attractor for modification and attention (Hawkins, 1994). Furthermore, in constructing our stimuli, we made sure the actions described in the RCs, such as "drinking water," "sleeping," and "taking a rest," are equally compatible with both animals and human beings, as was evidenced by our norming test results.

Another potential factor is the frequency effect. In both experiments, human nouns (NP_{low}) had higher frequencies than animal nouns (NP_{high}). Experiment 1 showed a marginally significant difference between NP_{low} ($M = 0.07$; $SD = 0.12$) and NP_{high} ($M = 0.03$; $SD = 0.04$) ($t = 1.94$, $p = 0.07$). Experiment 2 found a significant difference between NP_{low} ($M = 0.07$; $SD = 0.11$) and NP_{high} ($M = 0.02$; $SD = 0.03$) ($t = 2.31$, $p < 0.05$). However, since we are comparing the reading times at the NP_{low} region or the NP_{high} region across different conditions, rather than directly comparing the reading times between the NP_{low} and NP_{high} regions, frequency should not be a concern for interpreting the results. In fact, our findings show a tendency of high attachment, which is to attach the RC to a word of lower frequency. This is the opposite of what a typical frequency effect may predict (i.e., presumably words of higher frequencies may be stronger attractors).

Lastly, we turn to the potential length effect, a phenomenon posited by the Implicit Prosody Hypothesis (Fodor, 2002; Jun, 2003), which suggests that the length of RCs influences attachment preferences. Long RCs in head-final nominal structures are argued to favor high attachment, forming a single chunk that modifies the complex nominal phrase as a whole (e.g., long RC de1//NP_{low} de2 NP_{high}), while short RCs tend to group with their adjacent noun phrase, generating a preference for low attachment (e.g., short RC de1 NP_{low}//de2 NP_{high}) (Fodor, 2002; Jun, 2003; Hemforth et al., 2015). In the current study, the RCs are longer than those in two other studies on Chinese RC attachment preferences (Shen, 2006; Kwon et al., 2019)⁵ as they ranged from 7 to 12 characters (Experiment 1: $M = 9.63$, $SD = 1.24$; Experiment 2: $M = 9.46$, $SD = 1.14$). Additionally, Mandarin RCs are head-final; language users, therefore, encounter the embedded relative clause before the complex head noun is reached (Lin, 2019). Therefore, our study revealed a tendency to high-attachment driven by the long RC effect around the NP_{high} region, which affects the offline questionnaire stage.

However, our findings appear to contradict the production results in Kwon et al. (2019), who conducted a sentence completion task and concluded that Mandarin exhibits a preference for low attachment in the offline stage. It is noteworthy that Kwon et al. (2019) acknowledged that “almost all responses were descriptive relative clauses (78.4%)” (p. 5). In essence, most RCs generated by participants were simple attributive adjectival phrases of short length, such as *qinlao de nongfu de nongchang* “the farm of the farmer who is diligent,” where the RC length is three syllables. Therefore, their discovery of low-attachment preference aligns with the Implicit Prosody Hypothesis, possibly due to the short length of the RCs in their study (Fodor, 2002; Jun, 2003).

In conclusion, we explored the complexity of relative clause attachments in Mandarin Chinese, examining the influence of linguistic cues on comprehenders’ attachment preferences through two self-paced reading experiments. Our results demonstrate that classifier-noun agreement in Mandarin effectively resolves RC attachment ambiguities, influencing both online and offline processing phases. This study aligns with previous research using morphosyntactic dependencies to disambiguate relative clause attachment. The findings show that classifier-noun agreement, by pre-activating semantic features of the upcoming noun, guides comprehenders’ expectations and enhances processing efficiency. While the classifier’s prediction effect is significant, RC semantics can overshadow it in contexts with strong disambiguating cues. However, the classifier’s prediction effect remains influential, especially when congruent with RC semantics, leading to more efficient processing. This suggests that classifier-noun agreement functions similarly to morphosyntactic agreement in Indo-European languages, offering insights into attachment preferences during both online and offline phases. In addition, our study reveals a dynamic pattern of RC attachment preferences in Mandarin

Chinese, characterized by an early low-attachment during early online processing, transitioning to a high-attachment preference in later online and offline stages, particularly favoring NP_{high} as the attachment site.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: OSF platform “Frontiers_RC Attachment” https://osf.io/dtgjc/?view_only=b574c2aa1c5748d0828fee586babf680.

Ethics statement

The studies involving humans were approved by Indiana University Human Subjects and Institutional Review Boards. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their informed consent to participate in this study.

Author contributions

FL: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft. C-JL: Conceptualization, Formal analysis, Methodology, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by C-JL’s research fund and the Open Access Article Publishing Fund from the IU Libraries and the Office of the Vice Provost for Research of Indiana University Bloomington.

Acknowledgments

We are grateful to all our participants and members of the Language and Cognition Laboratory for their valuable feedback and comments. ChatGPT version 4o, a language model developed by OpenAI, has been used in the editing process to improve the readability and language of the manuscript.

Conflict of interest

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

⁵ Both Shen (2006) and Kwon et al. (2019) did not control RC lengths in their studies. Based on the three examples provided in the article, RC lengths in the second rating task are 4, 7, and 7 characters in Kwon et al. (2019). In Shen (2006), the RC lengths ranged from 5 to 14 characters ($M = 7.64$; $SD = 1.89$), with 87.5% of RCs having fewer than 10 characters.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Ahrens, K., and Huang, C. (2016). "Classifiers," in *A Reference Grammar of Chinese Reference Grammars*, eds. C. Huang and Shi, D. (Cambridge: Cambridge University Press), 169–198. doi: 10.1017/CBO9781139028462.008
- Akaike, H. (1998). "Information theory and an extension of the maximum likelihood principle," in *Selected Papers of Hirotugu Akaike*, eds. E. Parzen, K. Tanabe, and Kitagawa, G. (New York: Springer), 199–213. doi: 10.1007/978-1-4612-1694-0_15
- Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Memory Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001
- Box, G. E. P., and Cox, D. R. (1964). An analysis of transformations. *J. Royal Stat. Soc. Series B (Methodological)* 26, 211–252. doi: 10.1111/j.2517-6161.1964.tb00553.x
- Brybaert, M., and Mitchell, D. C. (1996). Modifier attachment in sentence processing: Evidence from Dutch. *Quart. J. Exp. Psychol.* 49A, 664–695. doi: 10.1080/0713755636
- Chan, S. H. (2019). An elephant needs a head but a horse does not: an ERP study of classifier-noun agreement in Mandarin. *J. Neuroling.* 52:100852. doi: 10.1016/j.jneuroling.2019.100852
- Chou, C. J., Chang, C. T., Tsai, J. L., and Lee, C. Y. (2015). "Interplay between semantic and syntactic information in Chinese Classifier-noun agreement: An ERP comparison," in *Poster at the Seventh Annual Neurobiology of Language Conference (SNL 2015)* (Chicago, IL).
- Chow, W.-Y., and Chen, D. (2020). Predicting (in)correctly: listeners rapidly use unexpected information to revise their predictions. *Lang. Cognit. Neurosci.* 35, 1149–1161. doi: 10.1080/23273798.2020.1733627
- Cotter, B. T., and Ferreira, F. (2014). The relationship between working memory capacity, bilingualism, and ambiguous relative clause attachment. *Mem. Cognit.* 1–18. doi: 10.3758/s13421-024-01561-4
- Cuetos, F., and Mitchell, D. C. (1988). Cross-linguistic differences in parsing: restrictions on the use of the Late Closure strategy in Spanish. *Cognition* 30, 73–105. doi: 10.1016/0010-0277(88)90004-2
- Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learning Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6
- de Leeuw, J. R., Gilbert, R. A., and Luchterhandt, B. (2023). jsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *J. Open Source Softw.* 8:5351. doi: 10.21105/joss.05351
- de Vincenzi, M., and Job, R. (1993). Some observations on the universality of the late-closure strategy. *J. Psycholinguist. Res.* 22, 189–206. doi: 10.1007/BF01067830
- Erbaugh, M. (2006). "Chinese classifiers: Their use and acquisition," in *The handbook of East Asian Psycholinguistics*, eds. P. Li, H. T. Li, E. Bates, and J. L. Tzeng (Cambridge: Cambridge University Press), 39–51. doi: 10.1017/CBO9780511550751.005
- Federmeier, K. D., and Kutas, M. (1999). A rose by any other name: long-term memory structure and sentence processing. *J. Mem. Lang.* 41, 469–495. doi: 10.1006/jmla.1999.2660
- Fernández, E. M. (2002). "Relative clause attachment in bilinguals and monolinguals," in *Bilingual Sentence Processing*, eds. R. R. Heredia and J. Altarriba (Amsterdam, NL: North Holland Elsevier Science Publishers), 187–216.
- Fodor, J. D. (2002). "Prosodic disambiguation in silent reading," in *Proceedings of the North East Linguistic Society (NELS)*, ed. M. Hirotani (Amherst, MA: GLSA Publications), 112–132.
- Grüter, T., Lau, E., and Ling, W. (2019). How classifiers facilitate predictive processing in L1 and L2 Chinese: the role of semantic and grammatical cues. *Lang. Cognit. Neurosci.* 35, 221–234. doi: 10.1080/23273798.2019.1648840
- Harding, E. E., Sammler, D., and Kotz, S. A. (2019). Attachment preference in auditory German sentences: Individual differences and pragmatic strategy. *J. Front. Psychol.* 10, 1–10. doi: 10.3389/fpsyg.2019.01357
- Hawkins, J. A. (1994). *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
- Hemforth, B., Fernández, S., Clifton, C., Frazier, L., Konieczny, L., and Walter, M. (2015). Relative clause attachment in German, English, Spanish and French: effects of position and length. *Lingua* 166, 43–64. doi: 10.1016/j.lingua.2015.08.010
- Hemforth, B., Konieczny, L., and Scheepers, C. (1994). "Principle based and probabilistic approaches to human parsing: how universal is the human language processor?" in *Paper presented at the Verarbeitung natürlicher Sprache Conference (KONVENS'94)*.
- Hopp, H., and Lemmerth, N. (2018). Lexical and syntactic congruency in L2 predictive gender processing. *Stud. Second Lang. Acquisit.* 40, 171–199. doi: 10.1017/S0272263116000437
- Hsu, C. N., Tsai, S., Yang, C. L., and Chen, J. (2014). Processing classifier-noun agreement in a long distance: an ERP study on Mandarin Chinese. *Brain Lang.* 137, 14–28. doi: 10.1016/j.bandl.2014.07.002
- Huang, Z., Feng, C., and Qu, Q. (2023). Predicting coarse-grained semantic features in language comprehension: evidence from ERP representational similarity analysis and Chinese classifier. *Cereb. Cortex* 33, 8312–8320. doi: 10.1093/cercor/bhad116
- Jun, S. A. (2003). Prosodic phrasing and attachment preferences. *J. Psycholinguist. Res.* 32, 219–249. doi: 10.1023/A:1022452408944
- Kamide, Y., Altmann, G. T. M., and Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: evidence from anticipatory eye movements. *J. Mem. Lang.* 49, 133–156. doi: 10.1016/S0749-596X(03)00023-8
- Kamide, Y., and Mitchell, D. C. (1997). Relative clause attachment: non-determinism in Japanese parsing. *J. Psycholinguist. Res.* 26, 247–254. doi: 10.1023/A:1025017817290
- Kamide, Y., and Mitchell, D. C. (1999). Incremental pre-head attachment in Japanese parsing. *Lang. Cognit. Proc.* 14, 631–662. doi: 10.1080/016909699386211
- Kuperberg, G. R., and Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Lang. Cognit. Neurosci.* 31, 32–59. doi: 10.1080/23273798.2015.1102299
- Kwon, N., Ong, D., Chen, H., and Zhang, A. (2019). The role of animacy and structural information in Relative Clause Attachment: Evidence from Chinese. *Front. Psychol.* 10:1576. doi: 10.3389/fpsyg.2019.01576
- Kwon, N., Sturt, P., and Liu, P. (2017). Predicting semantic features in Chinese: evidence from ERPs. *Cognition* 166, 433–446. doi: 10.1016/j.cognition.2017.06.010
- Lee, D., and Kweon, S. (2004). A sentence processing study of relative clause in Korean with two attachment sites. *Discourse Cogn.* 11, 125–140.
- Lee, S. Y. (2021). The Effect of Honorific Affix on Processing of an Attachment Ambiguity. *Japanese/Korean Linguist.* 28, 1–10.
- Lee, S. Y., and Yoo, M. Y. (2023). Honorific agreement in Korean: retrieval processing and predictive processing. *Linguist. Res.* 40, 387–407. doi: 10.17250/KHISLI.40.3.202312.002
- Lin, C. J. C. (2019). "Chinese psycholinguistics: A typological overview," in *The Routledge Handbook of Chinese Applied Linguistics*, eds. R. C. R. Huang, J. S. Zhuo, and B. Meisterernst (London: Routledge), 773–785. doi: 10.4324/9781315625157-48
- Lin, C. J. C., and Bever, T. G. (2011). "Garden path and the comprehension of head-final relative clauses," in *Processing and Producing Head-Final Structures*, eds. H. Yamashita, Y. Hirose, and J. L. Packard (New York, NY: Springer), 277–297. doi: 10.1007/978-90-481-9213-7_13
- Linck, J., and Cunnings, I. (2015). The utility and application of mixed effects models in second language research. *Lang. Learn.* 65, 185–207. doi: 10.1111/lang.12117
- Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. New York: Oxford University Press.
- Mahmoodi, M. H., Sheykhmololuki, H., Zoghipaydar, M. R., and Shahsavari, S. (2022). Working memory capacity and relative clause attachment preference of Persian EFL learners: does segmentation play any role? *J. Psycholinguist. Res.* 51, 683–706. doi: 10.1007/s10936-021-09825-9
- Marefat, H., Samadi, E., and Yaseri, M. (2015). Semantic priming effect on relative clause attachment ambiguity resolution in L2. *Appl. Res. Engl. Lang.* 4, 78–95. doi: 10.22108/are.2015.15504

- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., and Bates, D. (2017). Balancing type I error and power in linear mixed models. *J. Mem. Lang.* 94, 305–315. doi: 10.1016/j.jml.2017.01.001
- Mendelsohn, A., and Pearlmuter, N. (1999). "Individual differences in relative clause attachment preferences," in *Poster Presented at the CUNY Conference on Human Sentence Processing* (New York, NY).
- Mitchell, D. C., and Cuetos, F. (1991). "The origins of parsing strategies," in *Current Issues in Natural Language Processing*, ed. C. Smith (Austin: Center for Cognitive Science, University of Texas) 1–12.
- Omaki, A. (2005). *Working Memory and Relative Clause Attachment in First and Second Language Processing* [Doctoral dissertation]. Honolulu: University of Hawaii at Manoa.
- Özge, D., Kornfilt, J., Maquate, K., Küntay, A. C., and Snedeker, J. (2022). German-speaking children use sentence-initial case marking for predictive language processing at age four. *Cognition* 221:104988. doi: 10.1016/j.cognition.2021.104988
- Özge, D., Küntay, A., and Snedeker, J. (2019). Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension. *Cognition* 183, 152–180. doi: 10.1016/j.cognition.2018.10.026
- Pickering, M. J., and Gambi, C. (2018). Predicting while comprehending language: a theory and review. *Psychol. Bull.* 144, 1002–1044. doi: 10.1037/bul0000158
- Pickering, M. J., and Garrod, S. (2013). An integrated theory of language production and comprehension. *Behav. Brain Sci.* 36, 329–347. doi: 10.1017/S0140525X12001495
- R Core Team (2023). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/> (accessed December 20, 2023).
- Roberts, L., and Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Appl. Psycholinguist.* 32, 299–331. doi: 10.1017/S0142716410000421
- Shen, X. (2006). *Late Assignment of Syntax Theory: Evidence from Chinese and English*. [Doctoral dissertation]. Exeter: University of Exeter.
- Swets, B., Desmet, T., Hambrick, D. Z., and Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: a psychometric approach. *J. Exp. Psychol.* 136, 64–81. doi: 10.1037/0096-3445.136.1.64
- Traxler, J. M. (2007). Working memory contributions to relative clause attachment processing: a hierarchical linear modeling analysis. *Mem. Cognit.* 35, 1107–1121. doi: 10.3758/BF03193482
- Traxler, J. M. (2009). A hierarchical linear modeling analysis of working memory and implicit prosody in the resolution of adjunct attachment ambiguity. *J. Psycholinguist. Res.* 38, 491–509. doi: 10.1007/s10936-009-9102-x
- Wang, L., Kuperberg, G., and Jensen, O. (2018). Specific lexico-semantic predictions are associated with unique spatial and temporal patterns of neural activity. *Elife* 7:e39061. doi: 10.7554/eLife.39061.019
- Wu, F., Kaiser, E., and Vasishth, S. (2018). Effects of early cues on the processing of Chinese relative clauses: evidence for experience-based theories. *Cogn. Sci.* 42, 1101–1133. doi: 10.1111/cogs.12551
- Zagar, D., Pynte, J., and Rativeau, S. (1997). Evidence for early-closure attachment on first-pass reading times in French. *Q. J. Exp. Psychol.* 50, 421–438. doi: 10.1080/713755715
- Zhou, X., Jiang, X., Ye, Z., Zhang, Y., Lou, K., and Zhan, W. (2010). Semantic integration processes at different levels of syntactic hierarchy during sentence comprehension: an ERP study. *Neuropsychologia* 48, 1551–1562. doi: 10.1016/j.neuropsychologia.2010.02.001