Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-025-06086-y OPEN

Object type effects on the processing of Chinese aspectual verbs in complement coercion

Wenting Xue[1,2]✉, Jinmeng Dou[3] & Wenxin Qin[1,4]

This study investigates whether object types influence the processing of Chinese aspectual verbs in complement coercion. Previous research has claimed that Chinese aspectual verbs (e.g., *kāishǐ* "start") require an event-denoting object and typically combine with a verb phrase. If this claim holds true, then when these verbs take a noun phrase (NP), whether the NP denotes an event or an entity, the processing patterns are assumed to be similar, as they share the same syntactic structure (aspectual verb + NP) and lack a verb between the aspectual verb and its NP object. Using a self-paced reading experiment and two norming tests, this study measured reading times of Chinese speakers on NP objects (EntityNP vs. EventNP) preceded by aspectual verbs and neutral verbs as controls (selecting either an EventNP or EntityNP), as in *kāishǐ/zhǔnbèi zhè-fú zuòpǐn/zhè-cì tiáo-sè* "started/prepared this artwork/color-mixing. The results showed that aspectual verbs paired with EntityNPs (complement coercion instances) elicited longer reading times than those paired with EventNPs, and such a difference was not found between the two kinds of expressions with neutral verbs. The findings suggest that object types affect the processing of Chinese aspectual verbs. The detected processing difficulty of aspectual verbs with EntityNP objects was more likely due to the enriched composition of the expressions (where the EntityNP is coerced from an entity into an event sense to satisfy the verb's selectional constraints), which goes beyond the effect of any potential syntactic irregularities following the aspectual verb. This study underscores the importance of considering both semantic and syntactic factors when examining the processing of aspectual verbs in Chinese, and it provides a foundation for future research to build upon.

[1] School of Foreign Languages, Huazhong University of Science and Technology, Wuhan, China. [2] HUST-GYENNO CNS Intelligent Digital Medicine Technology Center, Wuhan, China. [3] School of Humanities, Shanghai Jiao Tong University, Shanghai, China. [4] Institute of Modern Languages and Linguistics, Fudan University, Shanghai, China. ✉email: xuewentingxwt@outlook.com

Introduction

Argument selection, a fundamental property of verbs, is essential for understanding sentence structure and interpretation. Verbs, such as *start* and *finish*, inherently select an argument that denotes an event (Jackendoff, 1997; Pustejovsky, 1991, 1995). These verbs can take the eventive argument in gerund or infinitive forms, as in (1a) and (1b), or combine with a noun phrase (NP), as in (1c) and (1d). While the eventive NP *the fight* in (1c) meets the verb's selectional constraints, the entity NP *the book* in (1d) does not, leading to a type mismatch. The mismatch is resolved by complement coercion—a mechanism that coerces the entity NP into an event meaning (e.g., an activity related to the book, such as "reading") to fit the verb's requirements (Pustejovsky, 1991, 1995).

1. (a) Mary started reading the book.
(b) Mary started to read the book.
(c) Mary started the fight.
(d) Mary started the book.

This study aims to explore how the type of complement/object (entity- vs. event-denoting) affects the comprehension of sentences with aspectual verbs in Mandarin Chinese, through two norming tests (i.e., acceptability judgment and cloze test) and a self-paced reading experiment.

Complement coercion and processing

Complement coercion reflects a syntax-semantic interface relation (Pustejovsky, 1995; Pylkkänen & McElree, 2006). The notion arises to address the mismatch between an event-selecting verb (EventV) and an entity-denoting noun phrase (EntityNP), as shown in example (1d). This mismatch is resolved by adding event information associated with the denotation of the object to create a compatible interpretation. Notably, the additional event information is not directly expressed by any single word or the syntactic form of the sentence, yet readers can grasp it during real-time comprehension.

Empirical research has shown that sentences with complement coercion are more demanding to interpret than those that do not require it (Baggio et al., 2010; Frisson & McElree, 2008; Husband et al., 2011; Kuperberg et al., 2010; McElree et al., 2006, 2001; Pickering et al., 2005; Pylkkänen & McElree, 2007; Traxler et al., 2005, 2002; Zarcone et al., 2017). This difficulty is reflected in increased reading times (RTs) and additional brain activities at the objects and/or subsequent words.

For instance, McElree et al. (2001) recorded participants' self-paced RTs of sentences with an EntityNP preceded by one of three types of verbs: complement coercion verbs, which inherently take an event-denoting object; preferred verbs, which express a preferred meaning for complement coercion sentences; and non-preferred verbs, which express a plausible but non-preferred meaning for complement coercion sentences, as exemplified in *The secretary began/typed/read the memo before the annual sales conference*. The authors found that readers spent more time on the coerced noun *memo* and the subsequent word *before* compared to other conditions. In their event-related potential (ERP) study, Baggio et al. (2010) observed a larger N400 effect (an ERP component rising from difficulties in processing the upcoming words from semantic memory; see Kutas and Federmeier 2011 for details) in the 300–550 ms time window with a centro-parietal distribution for coerced nouns than their non-coerced equivalents. In another functional magnetic resonance imaging (fMRI) study, Husband et al. (2011) reported increased neural activities in Broadman's area in the left inferior frontal gyrus for complement coercion sentences, suggesting enhanced cognitive efforts.

The processing difficulty in complement coercion is commonly attributed to the enriched composition that resolves the type conflicts between the EventV and EntityNP object (Jackendoff, 1997). This mismatch is repaired by changing the object to an event type to align with the verb's selectional requirements, leading to a reconstruction of the semantic representation and resulting in processing costs (Baggio et al., 2010; Frisson & McElree, 2008; Husband et al., 2011; Kuperberg et al., 2010; McElree et al., 2006, 2001; Pickering et al., 2005; Traxler et al., 2005, 2002).

Complement coercion verbs: aspectual verbs

While the processing costs relevant to complement coercion are well-documented, there's an ongoing debate about which verbs should be classified as complement coercion verbs. Upon reviewing the existing empirical research, we noticed that the studies have included verbs from various semantic categories, including aspectual verbs (e.g., *begin*, *start*, *continue*, *finish*, *complete, end*), psych verbs (e.g., *hate*, *savor*, *endure*, *prefer*, *enjoy*, *resist*), as well as other verbs that do not fit into clear semantic groups (e.g., *attempt*, *try*, *master*, *manage*) (Baggio et al., 2010; De Almeida, 2004; Delogu et al., 2017; Frisson & McElree, 2008; Husband et al., 2011; Kuperberg et al., 2010; Lapata et al., 2003; Lowder & Gordon, 2016; McElree et al., 2006, 2001; Pickering et al., 2005; Pylkkänen & McElree, 2007; Spalek & Tomaszewicz, 2018; Traxler et al., 2005, 2002). However, the heterogeneity of complement coercion verbs included in these studies has caught the attention of some scholars. They argued that aspectual verbs, which describe an event's initiation, termination, or continuation (Levin, 1993), may have distinct argument selection properties (Katsika et al., 2012; Piñango & Deo, 2016).

To test the hypothesis, scholars have conducted empirical research to examine whether aspectual verbs with an EntityNP behaved distinctly from other types of verbs (particularly psych verbs) (Katsika et al., 2012; Lai et al., 2017; Ma et al., 2023; Xue et al., 2024). In an eye-tracking experiment, Katsika et al. (2012) found longer RTs for aspectual verbs than control verbs, while psych verbs did not show the same effect. Similarly, Lai et al. (2017) measured participants' self-paced RTs for EntityNP objects preceded by aspectual verbs and *enjoy*-verbs, with *love*-verbs as controls. They observed that only aspectual verbs led to longer RTs relative to *love*-verbs. Although these studies suggest that psych verbs may not pose the same processing difficulties as aspectual verbs, the studies have their own limitations. Katsika et al. (2012) did not eliminate the possibility that the divergent processing patterns between aspectual and psych verbs might be due to differences in the predictability of their NP objects, as readers are known to be sensitive to predictability nuances (Smith & Levy, 2013). As for Lai et al. (2017), they did not conclusively show that *enjoy*-verbs were processed without incurring additional cognitive costs, as both *enjoy*-verbs and *love*-verbs, as part of the psych verb group (Pesetsky, 1995), may require similar, albeit less demanding, processing than aspectual verbs.

Xue et al. (2024) addressed these concerns by controlling for NP object predictability and comparing self-paced RTs of Chinese speakers for EntityNPs preceded by aspectual and psych verbs, using non-psych verbs as controls. They found that EntityNPs following aspectual verbs were read more slowly than those following psych and control verbs. Ma et al. (2023) reported similar findings in another study conducted in Chinese.

The empirical evidence indicates that sentences with aspectual verbs and EntityNP objects exhibit unique processing patterns compared to their counterparts with non-aspectual verbs. This

suggests that the mechanism of complement coercion may be limited to aspectual verbs only.

Aspectual verbs in Mandarin Chinese

Aspectual verbs in Mandarin Chinese are thought to have different linguistic properties compared to their English counterparts (Lin & Liu, 2005). While these verbs can be followed by a NP object in English, some scholars argued that Chinese aspectual verbs should only take a verb phrase (VP) or clause as their object (Lu, 2020; Lu & Wen, 2018; Wen & Lu, 2017). Consequently, sentences like (2a) are deemed unacceptable, because the aspectual verb 开始 *kāishǐ* "begin" is followed by a NP这本书*zhè-běn shū* "the book". The sentence becomes acceptable only if a verb like 读*dú* "read" is inserted to specify an activity related to the object, as in (2b).

2. (a)*约翰开始这本书
yuēhàn kāishǐ zhè-běn shū
John begin this-classifier book
"John began the book."
(b) 约翰开始读这本书。
*yuēhàn kāishǐ **dú** zhè-běn shū*
John begin **read** this-classifier book
"John began **reading/to read** the book."

However, this claim has been challenged by studies following that Chinese aspectual verbs probably can co-occur with NPs (Hsu & Hsieh, 2013; Lin et al., 2009; Qin & Wang, 2012; Song, 2015; Xue et al., 2021b). Adopting a corpus-based approach, Lin et al. (2009) and Hsu and Hsieh (2013) analyzed instances of complement coercion from corpora, of which several instances pair aspectual verbs with EntityNP objects, as in (3).

3.开始一句话
kāishǐ yī-jù huà
begin one-classifier sentence
"begin a sentence"

Song (2014, 2015) discussed the differences in complement coercion between English and Chinese. They suggested that while complement coercion is a universal linguistic phenomenon, it is less common in Chinese due to different lexico-syntactic features. Particularly, Chinese aspectual verbs have weaker coercion potentials than their English equivalents, and the implicit event information of complement coercion tends to be expressed explicitly at the syntactic level of Chinese sentences. Similar viewpoints can also be found in Qin and Wang (2012).

Song (2014) further stated that although aspectual verbs can co-occur with eventive NPs, it is preferable to insert a verb specifying an action associated with the NP's denotation or a light verb (in Jespersen's 1965 sense, such as *have* and *take* in *have a rest* and *take a walk*) with little semantic content between the aspectual verbs and their NP objects. Take (4) as an example. Although the NP object教育改革*jiàoyù gǎigé* "educational reform" is an eventive nominal which satisfies the selectional constraints of the aspectual verb 开始 *kāishǐ* "begin", a light verb 进行*jìnxín* "do" is still preferred to appear after the matrix verb.

4.开始(进行)教育改革
kāishǐ (jìnxíng) jiàoyù gǎigé
begin (do) educational reform
"begin educational reforms"

The viewpoint that Chinese aspectual verbs tend to pair with a VP correlates with typological traits shared by English and Chinese (Song, 2014, 2015). Through a contrastive analysis of Chinese and English, Liu (2010, p. 17) proposed that "Chinese is a verby language whereas English is a nouny language." That means arguments that can be expressed by nouns in English can be freely represented by verbs in Chinese, and arguments that can be expressed either by verbs or nouns in English are essentially represented only by verbs in Chinese. With this in mind, the unstated event information in complement coercion with Chinese aspectual verbs tends to be expressed directly on the surface structure of sentences. This aligns with previous studies suggesting that complement coercion occurs less frequently in Chinese than in English (Qin & Wang, 2012; Song, 2014, 2015).

Puzzles in processing Chinese aspectual verbs with different object types

As previously discussed, some Chinese linguists held the point that Chinese aspectual verbs should combine with a VP as their object (Lin & Liu, 2005; Lu, 2020; Lu & Wen, 2018; Wen & Lu, 2017), and even when followed by an eventive NP, another verb is often inserted between the matrix verb and the object (Song, 2014, 2015). If this is the case, aspectual verbs paired with NP objects (either EventNP or EntityNP) might show similar processing patterns due to the absence of a verb after the matrix verb.

Empirical studies have indicated that Chinese aspectual verbs with EntityNP objects are more difficult to process than those with non-aspectual verbs (Xue et al., 2021a, 2024). For example, Xue et al. (2021a) measured self-paced RTs of Chinese speakers for EntityNPs preceded by aspectual verbs versus two types of non-aspectual verbs (verbs that express a preferred meaning of the aspectual verb sentences, and verbs that suggest a plausible but less preferred meaning), as in 医生开始/阅读/检查这份病例 *yīshēng kāishǐ / yuèdú / jiǎnchá zhè-fèn bìng-lì* "The doctor started/read/checked the medical record." The results showed that EntityNPs preceded by aspectual verbs had longer RTs than those preceded by the other two types of verbs.

Given the established processing difficulty with Chinese aspectual verbs and EntityNP objects, this study seeks to answer the research question: Will aspectual verbs followed by EventNP objects elicit similar processing patterns as those followed by EntityNP objects? Understanding these patterns could shed light on the processing of Chinese aspectual verbs across different object types.

Current study

This study examined the effect of object types on the processing of Chinese aspectual verbs using a 2 × 2 factorial design, following Traxler et al. (2002, Experiment 2). We explored the effects of object types (EntityNP vs. EventNP) and verb types (aspectual verbs vs. neutral verbs) across four conditions: (a) aspectual verb with EntityNP, (b) aspectual verb with EventNP, (c) neutral verb with EntityNP, and (d) neutral verb with EventNP. One example is shown in (5) below. Aspectual verbs, which typically select event-denoting objects and have been argued to combine with a VP, contrast with neutral verbs that can take either entity- or event-denoting objects and co-occur with NPs and VPs. EntityNPs express entities, while EventNPs express events. Given the syntactic features of Chinese aspectual verbs, conditions (a) and (b) lack a verb between the predicate verb and its object, with (a) also involving complement coercion; conditions (c) and (d) are well-formed compositions without a verb missed.

We anticipated three possible outcomes based on the design: If (a) > (b) > (c) = (d) in RTs, then both complement coercion and the absence of a verb between the aspectual verb and its NP object

may play a role. If (a) = (b) > (c) = (d), the absence of a verb may be crucial. However, if (a) > (b) = (c) = (d), complement coercion in (a) could be more influential.

(5) 画家开始 (a)这幅作品/ (b) 这次调色。(CL = classifier)
huàjiā kāishǐ zhè-fú zuòpǐn/zhè-cì tiáo-sè
artist start this-CL artwork/this-CL color-mixing
"The artist started this artwork/color-mixing."
画家准备 (c) 这幅作品/ (d)这次调色。(CL = classifier)
huàjiā zhǔnbèi zhè-fú zuòpǐn/zhè-cì tiáo-sè
artist prepare this-CL artwork/this-CL color-mixing
"The artist prepared this artwork/color-mixing."

The study includes two norming tests—an acceptability judgment and a cloze test—followed by a self-paced reading experiment. Details are presented as follows.

Acceptability judgment. The acceptability judgment task was undertaken to assess the acceptability of experimental sentences, ensuring they were acceptable to native Chinese speakers.

Participants. 320 native Mandarin Chinese speakers (191 women and 129 men; *age*: mean = 20, *range* = 18–27) rated the acceptability of 96 item quadruplets. Participants gave their informed consent and were compensated upon completion of the judgment task.

Materials. For the acceptability judgment task, researchers created 96 quadruplets with five aspectual verbs selected from Xue et al. (2024): 开始 *kaishǐ* "begin/start", 继续 *jìxù* "continue", 完成 *wánchéng* "finish", 结束 *jiéshù* "end", 停止 *tíngzhǐ* "stop". EntityNP and EventNP objects were constructed with a demonstrative determiner(这/那 *zhè/nà* "this/that"), followed by a classifier specific to their type—either an entity-type for concrete objects, or an event-type for activities—and a two-character noun. The nouns for EntityNPs denote physical objects; the nouns for EventNPs can be modified by a duration phrase (e.g., "2 h") or an event classifier (Han, 2016). For instance, an EntityNP might be 这幅作品*zhè-fú zuòpǐn* "this-CL artwork", while an EventNP could be这次调色*zhè-cì tiáo-sè* "this-CL color-mixing".

Procedures. The acceptability judgment task was carried out in two phases. Initially, 160 participants were invited to rate the sentences on a 6-point scale (1= "completely unacceptable", 6= "completely acceptable") based on their intuition. The 96-item sets were arranged into four lists using a Latin square design, ensuring that no single list contained all four conditions of any given item. To lessen the rating load for each participant, these lists were then broken down into eight sub-lists, each containing 48 test sentences. Additionally, 52 ungrammatical filler sentences were added. The 160 participants were randomly assigned to eight lists, with each list containing 20 participants. After data collection, 23 sets of items with ratings above 3.5 (the midpoint on the 6-point scale) were selected, and 22 sets with very low acceptability were discarded. The remaining 51 sets were revised and rated again in the second phase by the remaining 160 participants.

Results. Items that were rated below 3.5 on at least one of the four conditions were eliminated across the two phases, resulting in 44 sets for the following cloze norming and the self-paced reading experiment. Average acceptability ratings were as follows: 4.28 (*range* = 3.55–5.25) for aspectual verbs with EntityNPs, 4.50 (*range* = 3.5–5.65) for aspectual verbs with EventNPs, 4.40 (*range* = 3.5–5.15) for neutral verbs with EntityNPs, and 4.30 (*range* = 3.5–5.35) for neutral verbs with EventNPs. To eliminate potential scale biases, the acceptability ratings were z-transformed prior to statistical analysis. The analysis did not show a main effect of verb type or object type; instead, it showed an interaction between verb type and object type (F (1, 172) = 5.359, p = 0.022). Such an interaction was consistent with RT data collected from the subsequent self-paced reading experiment (see section "Self-paced reading"), and was further discussed in section "Discussion".

Cloze test. The cloze test was designed to measure the predictability of object NPs (the likelihood that participants would include the target words in their responses) following different verb types. Crucially, the cloze probability data would be incorporated into subsequent RT data analyses to eliminate any potential effect of NP predictability on the processing of the experimental sentences.

Participants. A new group of 160 Chinese speakers (92 women and 68 men; *age*: mean = 21 years; *range* = 18–33) were recruited for the cloze test. They gave their informed consent and received compensation upon completion of the cloze test.

Materials. The materials included 44 quadruplets of items identified through the acceptability judgment task. They were split into two lists, 22 quadruplets per list. Employing a Latin Square design, each main list was then broken down into four sub-lists using, ensuring that no single list contained all four conditions for any item. This process led to a total of eight lists.

Procedures. The 44 sets were presented as sentence fragments with a subject, verb, and demonstrative 这/那 *zhè/nà* "this/that", such as 演员完成/理解这___ *yǎnyuán wánchéng/ lǐjiě zhè*___ "The actor finished/understood the___." Participants completed the sentence with a NP (including a classifier and a two-character noun). The test began with two samples (e.g., 女孩撕碎这张白纸 *nǚhái sīsuì zhè-zhāng báizhǐ* "The girl tore the white paper"). The eight lists were randomly assigned to 160 participants, with each list completed by 20 participants.

Results. For the 44 sentence sets, participants' responses were evaluated against the target object NPs according to two specific criteria: First, the noun in the response had to match the target noun, and the responded classifier and the target classifier had to belong to the same semantic category (either event or entity type). For instance, a response like 这本小说*zhè-běn xiǎoshuō* "this novel" would be considered equivalent to the target NP 这部小说 *zhè-bù xiǎoshuō* "this novel", because the nouns (小说 *xiǎoshuō* "novel") are identical; the classifiers (本 *běn* vs. 部*bù*), though different, both can describe books, with 本 *běn* typically referring to individual books or volumes, and 部*bù* denoting larger collections. Similarly, 这场表演*zhè-chǎng biǎoyǎn* "this performance" would match 这次表演 *zhè-cì biǎoyǎn* "this performance", because their nouns (表演*biǎoyǎn* "performance") are the same; and the classifiers are of the event type, with 场 *chǎng* typically indicating events that have a defined space or duration, and 次*cì* denoting the order or sequence of events. However, a response like 这次报道*zhè-cì bàodào* "this instance of reporting" would not match the target NP 这篇报道*zhè-piān bàodào* "this article", because 次*cì* is an event-type classifier, whereas 篇*piān* "piece" is an entity-type classifier, quantifying written content in a way that acknowledges the completeness of a written work.

Second, the classifier in the response must match the target classifier, and the noun in the response must belong to the same

Table 1 An example set of the experimental stimuli.

Condition		Verb	Object NP	NP+1	NP+2	
AspV + EntityNP	画家	开始	这幅作品	之前	喝了	一杯咖啡。
	huàjiā	*kāishǐ*	*zhè-fú zuòpǐn*	*zhīqián*	*hēle*	yī-bēi kāfēi
	artist	start	this-CL artwork	before	drink	one-CL coffee
	"The artist drank a cup of coffee before starting this artwork."					
AspV + EventNP	画家	开始	这次调色	之前	喝了	一杯咖啡。
	huàjiā	*kāishǐ*	*zhè-cì tiáo-sè*	*zhīqián*	*hēle*	yī-bēi kāfēi
	artist	start	this-CL color-mixing	before	drink	one-CL coffee
	"The artist drank a cup of coffee before starting this color-mixing."					
NeuV + EntityNP	画家	准备*zhǔnbèi* prepare	这幅作品	之前	喝了	一杯咖啡。
	huàjiā		*zhè-fú zuòpǐn*	*zhīqián*	*hēle*	*yī-bēi kāfēi*
	artist		this-CL artwork	before	drink	one-CL coffee
	"The artist drank a cup of coffee before preparing this artwork."					
NeuV + EventNP	画家	准备	这次调色	之前	喝了	一杯咖啡。
	huàjiā	*zhǔnbèi* prepare	*zhè-cì tiáo-sè*	*zhīqián*	*hēle*	*yī-bēi kāfēi*
	artist		this-CL color-mixing	before	drink	one-CL coffee
	"The artist drank a cup of coffee before preparing this color-mixing."					

AspV aspectual verb, *NeuV* neutral verb, *EntityNP* entity-denoting noun phrase, *EventNP* event-denoting noun phrase, *CL* classifier.

semantic category as the target noun. For instance, a response like 这幅创作*zhè-fú chuàngzuò* "this piece of creation" was not considered equivalent to the target NP 这幅油画*zhè-fú yóuhuà* "this piece of oil painting", because 创作*chuàngzuò* "creation" is primarily an event-type noun, while 油画*yóuhuà* "oil painting" is an entity-type noun, even though the classifiers are the same (i.e., 幅*fú*).

The cloze probabilities for the object NPs following the two types of verbs were as follows: 0.06 (*range* = 0–0.55) for aspectual verb with EntityNP, 0.22 (*range* = 0–0.85) for aspectual verb with EventNP, 0.11 (*range* = 0–0.50) for neutral verb with EntityNP, and 0.12 (*range* = 0–0.60) for neutral verb with EventNP. The statistical analysis revealed a main effect of object type (F (1, 172) = 11.12, p = 0.001), and an interaction effect between object type and verb type (F (1, 172) = 8.02, p = 0.005). Since this study examined how object types influenced the comprehension of Chinese sentences with aspectual verbs, the cloze probabilities of the objects were essential for evaluating the processing of the stimuli. Thus, they were included in the final statistical models analyzing the RT data (see section "Data analysis" for details).

Self-paced reading. Building on the stimuli from the acceptability judgment task, we measured self-paced RTs of Chinese speakers on EntityNP and EventNP objects following aspectual or neutral verbs. The primary goal was to examine the effects of object types on the processing of Chinese sentences with aspectual verbs.

Participants. 83 native Chinese speakers (49 women and 34 men; age: *mean* = 20 years; *range* = 18–27) participated in this experiment. However, seven participants were excluded due to their low comprehension question accuracy or a high number of outliers (for details, see section "Data analysis"). Participants had normal or corrected-to-normal vision, and none reported any language disorders. They gave informed consent and received financial compensation after the test was completed.

Materials. The materials were 44 quadruplets of items adapted from the acceptability judgment task, with each sentence containing 14 to 16 characters. A sample set is provided in Table 1. The five aspectual verbs, along with the number of items each verb appears in, are as follows: 开始 *kaishǐ* "begin/start" (6), 继续 *jìxù* "continue" (5), 完成 *wánchéng* "finish" (13), 结束 *jiéshù* "end" (8), and 停止 *tíngzhǐ* "stop" (12). We distributed these sets randomly across four lists using a Latin square design, ensuring that each list contained only one version of each item. To keep the research purpose from being apparent to participants, we added 47 fillers with diverse structures. Six practice trials were included at the start to help participants get familiar with the experimental procedures. Each list contained 97 sentences, followed by a comprehension question for each participant to maintain focus. None of these questions pertained to the predicate verbs or objects. For example, for the target sentence 画家开始这幅作品之前喝了一杯咖啡 *huàjiā kāishǐ zhè-fú zuòpǐn zhīqián hēle yī-bēi kāfēi* "The artist drank a cup of coffee before starting this artwork", the comprehension question translated in English is "Did the artist start this artwork after drinking a cup of coffee?"

Procedures. The experiment took place in a language lab, where participants were randomly assigned to one of the four lists. The experiment began with on-screen instructions, followed by six practice trials and 91 experimental trials. The experiment lasted approximately 20 minutes.

Using DMDX software (Forster & Forster, 2003), sentences were displayed one word at a time in a moving window format against a black background with white text. Trials began with a "+" sign and several sets of dashes. The first set of dashes were replaced by the first word when participants pressed the spacebar. Subsequent button presses revealed each following word, with dashes reappearing in place of the previous one. Participants read the sentences at their own pace. A comprehension question with "Yes" or "No" options was presented after each sentence. Participants provided their answers by pressing the "Yes" ("F" button) or "No" ("J" button) on the keyboard. The next trial began only after a response was provided. Except for the practice trials, all sentences were shown randomly, with the computer recording participants' RTs for each word and their answers to the comprehension questions.

Data analysis. Statistical analyses focused on RTs from four critical regions: the verb, the NP object, and the two regions right after the NP (that is, NP+1 and NP+2). The analysis of the verb region was done to identify any preliminary processing differences that might affect the upcoming NPs. The main effects of interest were anticipated in the object NP region or the subsequent spillover regions.

Prior to the statistical analyses, we did data cleaning following Xue et al. (2024). Participants with less than 75% accuracy on comprehension questions were excluded, resulting in the removal

Table 2 Comprehension question accuracy.

Condition	Mean accuracy
AspV + EntityNP	94.62%
AspV + EventNP	95.33%
NeuV + EntityNP	94.62%
NeuV + EventNP	94.86%

AspV aspectual verb, *NeuV* neutral verb, *EntityNP* entity-denoting noun phrase, *EventNP* event-denoting noun phrase.

Table 3 Mean reading times by condition and region.

Condition	Region			
	Verb	Object NP	NP+1	NP+2
AspV + EntityNP	352 [334, 369]	382 [357, 407]	407 [383, 432]	383 [364, 403]
AspV + EventNP	355 [332, 375]	369 [346, 388]	366 [347, 382]	347 [327, 366]
NeuV + EntityNP	355 [337, 374]	373 [348, 398]	374 [353, 391]	363 [343, 379]
NeuV + EventNP	351 [331, 367]	372 [346, 395]	370 [352, 386]	361 [342, 377]

The bracketed figures represent difference-adjusted 95% (percentile) mixed-effects model-based intervals (Politzer-Ahles, 2017) for each region.
AspV aspectual verb, *NeuV* neutral verb, *EntityNP* entity-denoting noun phrase, *EventNP* event-denoting noun phrase.

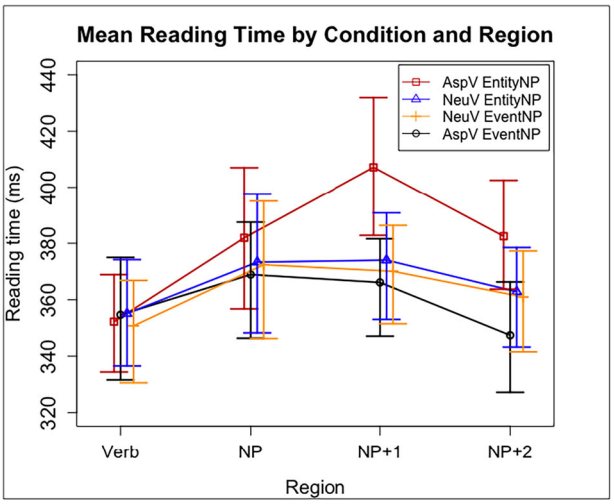

Fig. 1 Mean reading times and difference-adjusted 95% (percentile) mixed-effect-model-based intervals (Politzer-Ahles, 2017). The intervals roughly indicate significant differences between conditions: If the interval range of one condition does not overlap with the mean of another and vice versa, it suggests a statistically significant difference between those conditions within a mixed-effect model.

of three participants. Participants with more than 40% of RTs for the four regions over 2000 ms or under 100 ms were excluded, leading to four more exclusions. For the remaining 76 participants, the comprehension question accuracy for the four sentence types was presented in Table 2. RTs outside the range of 100 ms to 2000 ms were also excluded, totaling 275 data points (2.06%) lost.

The rest of the data was analyzed utilizing the R programming language (R Core Team, 2018). To get an overview of RTs for each condition across the four specified regions, mean RTs and the difference-adjusted 95% (percentile) mixed-effects model-based intervals (Politzer-Ahles, 2017) were first computed. The intervals serve as a gauge for roughly detecting significant differences between conditions: If the interval range of one condition does not overlap with the mean of another and vice versa, it suggests a statistically significant difference between those conditions within a mixed-effect model. Results are shown in Table 3, and visualized in Fig. 1.

In line with Xue et al. (2024), we performed separate linear mixed-effects models for each region's data with the *lme4* package (Bates et al., 2015). RTs, the dependent variable, were log-transformed to achieve a model with residuals approximating a normal distribution. A linear mixed model was applied to the four critical regions, encompassing the categorical fixed effects of Verb Type and Object Type, the continuous fixed effect of NP Probability, and the interaction between Verb Type and Object Type. NP Probability (see section "Cloze test") was factored into the model, whereas Sentence Acceptability (see section "Acceptability judgment") was not, as the former was deemed more pivotal to affect processing patterns of NP objects. The sentence acceptability was judged to guarantee that all sentences were acceptable to native speakers, thereby reducing any potential processing difficulties arising from the stimuli's unacceptability. For further validation, we did include sentence acceptability in linear mixed-effects models and conducted additional analyses (see section 1 of the Supplementary Materials online, where statistical results are shown in Table S1). However, the primary findings remained consistent with those reported in the manuscript.

We started our analysis with the full random effects structure supported by the present design, incorporating crossed random intercepts for both subjects and items, as well as random slope parameters for the main effects of Object Type, Verb Type and NP Probability, and the interaction between Verb Type and Object Type [RT ~ VerbType * ObjectType + NPProbability + (1 + VerbType * ObjectType + NPProbability | Subject)+(1 + VerbType * ObjectType + NPProbability | Item)]. The random effects structure was then simplified through model comparisons to identify the optimal model fit for each region [RT ~ VerbType * ObjectType + NPProbability + (1 + VerbType * ObjectType | Subject)+(1 + VerbType + ObjectType | Item)].

The main effects of Object Type, Verb Type, and NP Probability, as well as the interaction between Object Type and Verb Type, were assessed via likelihood ratio tests (comparing models with and without the factor in question). The results are detailed in Table 4. As the interaction between Object Type and Verb Type was also detected at the two post-NP regions (see Table 4), further analyses were conducted to explore how Object Type interacted with Verb Type to influence the RTs of sentences. We examined in which Verb Type (either Aspectual Verb or Neutral Verb) the Object Type (EntityNP vs. EventNP) may affect the RTs, and in which Object Type (either EntityNP or EventNP) the Verb Type (Aspectual Verb vs. Neutral Verb) may affect the RTs. These were accomplished by summarizing the maximal fitting models when each factor was considered. For example, when the effect of Object Type was examined for aspectual and neutral verbs, respectively, we incorporated the fixed effect of Object Type, crossed random intercepts for both subjects and items, and random slope parameters for the main effect of Object Type [RT ~ ObjectType + (1 + ObjectType | Subject)+(1 + ObjectType | Item)]. The results are shown in Tables 5 and 6, respectively. When the effect of Verb Type was examined for EntityNP and EventNP objects, respectively, we incorporated the fixed effect of Verb Type, crossed random intercepts for both subjects and items, and random slope parameters for the main effect of Verb Type [RT ~ VerbType

Table 4 Results of likelihood ratio tests to examine the main effects of verb type, object type, and NP probability, and the interaction effect between object type and verb type.

Region	VT* OT		VT		OT		NPP	
	χ^2 (*df*)	*p*-value	χ^2 (*df*)	*p*-value	χ^2 (*df*)	*p*-value	χ^2 (*df*)	*p*-value
Verb	0.117 (1)	0.733	0.034 (1)	0.854	0.201 (1)	0.654	4.250 (1)	0.039
NP	2.118 (1)	0.146	1.057 (1)	0.304	1.661 (1)	0.197	0.044 (1)	0.833
NP+1	11.09 (1)	0.001***	2.676 (1)	0.102	5.467 (1)	0.019*	0.114 (1)	0.735
NP+2	10.835 (1)	0.001**	0.660 (1)	0.417	6.485 (1)	0.011*	1.079 (1)	0.299

VT verb type, *OT* object type, *NPP* NP probability.
$p \leq 0.001$ "***", $0.001 < p \leq 0.01$ "**", $0.01 < p \leq 0.05$ "*".

Table 5 Results of object type effects for aspectual verbs.

Region	Verb type: aspectual verb (*AspV* + *EntityNP vs. AspV* + *EventNP*)				
	Estimate	*Std.error*	*df*	*t*	*p*-value
Verb	-0.013	0.022	74.333	-0.588	0.558
NP	-0.056	0.033	50.692	-1.693	0.097
NP+1	-0.108	0.024	42.097	-4.583	0.000***
NP+2	-0.087	0.019	116.516	-4.56	0.000***

$p \leq 0.001$ "***".

Table 6 Results of object type effects for neutral verbs.

Region	Verb type: neutral verb (*NeuV* + *EntityNP vs. NeuV* + *EventNP*)				
	Estimate	*Std. error*	*df*	*t*	*p*-value
Verb	-0.002	0.023	36.579	-0.078	0.939
NP	-0.000	0.023	59.195	-0.008	0.994
NP+1	-0.008	0.021	43.765	-0.367	0.716
NP+2	-0.001	0.018	113.800	-0.054	0.957

Table 7 Results of verb type effects for EntityNP objects.

Region	Object type: entity NP (*AspV* + *EntityNP vs. NeuV* + *EntityNP*)				
	Estimate	*Std. error*	*df*	*t*	*p*-value
Verb	-0.002	0.021	54.455	-0.08	0.936
NP	-0.048	0.028	34.661	-1.675	0.103
NP+1	-0.090	0.025	36.292	-3.656	0.001***
NP+2	-0.060	0.020	94.175	-3.073	0.003**

$p \leq 0.001$ "***", $0.001 < p \leq 0.01$ "**".

Table 8 Results of verb type effects for EventNP objects.

Region	Object type: event NP (*AspV* + *EventNP vs. NeuV* + *EventNP*)				
	Estimate	*Std.error*	*df*	*t*	*p*-value
Verb	0.010	0.026	40.935	0.351	0.727
NP	0.006	0.023	119.300	0.254	0.800
NP+1	0.010	0.020	282.422	0.510	0.611
NP+2	0.026	0.019	115.345	1.364	0.175

+ (1 + VerbType | Subject)+(1 + VerbType | Item)]. The results are shown in Tables 7 and 8, respectively. The *df* and *p*-values presented in these tables were computed using the lmerTest package (Kuznetsova et al., 2017), which was initialized at the start of the analysis process.

Results

Comprehension question accuracy: The mean question accuracy rate was 94.86% across all four conditions. The accuracy rate for each condition was presented in Table 2. There were no significant conditional effects were observed, and no significant interaction effect was detected between the verb type and object type.

Reading times: Table 3 and Fig. 1 illustrate the mean RTs in milliseconds (ms) for all conditions at the four critical regions, along with the difference-adjusted 95% (percentile) mixed-effect-model-based intervals (Politzer-Ahles, 2017). As detailed in Table 4, the verb and the object regions did not reveal the main effects of Object Type, Verb Type, NP Probability, and the interaction effect between Object Type and Verb Type. The regions immediately following the object NP, i.e., NP+1 and NP+2, showed the main effect of Object Type and the interactions.

In the NP+1 region, the analyses revealed the main effects of Object Type ($\chi^2(1) = 5.467$, $p = 0.019$), and the interaction between Object Type and Verb Type ($\chi^2(1) = 11.09$, $p = 0.001$) (see Table 4). Considering the effect of Object Type, it was only found for aspectual verbs (*Estimate* = –0.108, *SE* = 0.024, $t = -4.583$, $p = 0.000$) (see Table 5), with EntityNP objects read slower than EventNP objects (407 ms vs. 366 ms) (see Table 3). This effect was not observed with neutral verbs (*Estimate* = –0.008, *SE* = 0.021, $t = -0.367$, $p = 0.716$) (see Table 6). As the analyses revealed an interaction effect between Object Type and Verb Type, we further looked at whether there was any effect of Verb Type exhibited, and found that the effect occurred only for EntityNP objects (*Estimate* = –0.090, *SE* = 0.025, $t = -3.656$, $p = 0.001$) (see Table 7), with aspectual verbs triggering longer RTs than neutral verbs (407 ms vs. 374 ms) (see Table 3). The Verb Type effect was not observed for EventNP objects (*Estimate* = 0.010, *SE* = 0.020, $t = 0.510$, $p = 0.611$) (see Table 8).

In the NP+2 region, the analyses demonstrated the main effect of Object Type ($\chi^2(1) = 6.485$, $p = 0.011$) and the interaction between Object Type and Verb Type ($\chi^2(1) = 10.835$, $p = 0.001$) (see Table 4). Considering the effect of Object Type, it was only found for aspectual verbs (*Estimate* = –0.087, *SE* = 0.019, $t = -4.56$, $p = 0.000$) (see Table 5), with EntityNP objects read slower than EventNP objects (383 ms vs. 347 ms) (see Table 3). Such an effect was not found for neutral verbs (*Estimate* = –0.001, *SE* = 0.018, $t = -0.054$, $p = 0.957$) (see Table 6). Considering the effect of Verb Type, it was only detected for EntityNP objects (*Estimate* = –0.060, *SE* = 0.020, $t = -3.073$, $p = 0.003$) (see Table 7), with aspectual verbs triggering longer

RTs than neutral verbs (383 ms vs. 363 ms) (see Table 3). The Verb Type effect was not observed for EventNP objects (*Estimate* = 0.026, *SE* = 0.019, *t* = 1.364, *p* = 0.175) (see Table 8).

Summing up, the interaction effect between Object Type and Verb Type reached the significance level at the two post-NP regions (i.e., NP +1 and NP +2), which resulted from the interplay of the effect of Object Type when the objects were preceded by aspectual verbs and the effect of Verb Type when the objects were EntityNPs.

Discussion

This study investigated the effect of NP object types (entity-denoting or event-denoting) on the processing of Chinese sentences with aspectual verbs. We contrasted the processing patterns of aspectual verbs followed by EntityNPs or EventNPs against neutral verbs as controls. The main finding was that aspectual verbs with EntityNPs were more difficult to process than those with EventNPs, and such a difference was not observed for neutral verbs; EntityNP objects exerted divergent processing patterns when preceded by aspectual verbs and neutral verbs. This study provides robust evidence that object types influence the processing of Chinese aspectual verb sentences; the detected processing cost results more likely from the enriched composition of the expressions (where the EntityNP is coerced from an entity into an event sense to satisfy the verb's selectional constraints), which goes beyond any effects of the absence of a verb/syntactic irregularity (if there is) between the Chinese aspectual verbs and the object NPs. Details are discussed below.

The observation—EntityNP objects preceded by aspectual verbs (complement coercion) were processed slower than those preceded by neutral verbs—suggests that the former is more taxing to interpret than the latter. The result aligns with the literature on complement coercion in Indo-European languages (such as English, Polish, and German), both theoretically (Jackendoff, 1997; Pustejovsky, 1991, 1995; Pustejovsky & Bouillon, 1995) and empirically (Baggio et al., 2010; De Almeida et al., 2016; Frisson & McElree, 2008; Husband et al., 2011; Katsika et al., 2012; Kuperberg et al., 2010; Lai et al., 2017; Lowder & Gordon, 2016; McElree et al., 2006; McElree et al., 2001; Pickering et al., 2005; Pylkkänen & McElree, 2007; Spalek & Tomaszewicz, 2018; Traxler et al., 2005, 2002; Zarcone et al., 2017). Theoretically, aspectual verbs, which describe an event's initiation, termination, or continuation (Levin, 1993), semantically require an event-denoting object (e.g., *begin reading/to read the book*). When they co-occur with an entity-denoting object (e.g., *began the book*), an event sense associated with the object's denotation is supposed to be added (e.g., an activity of reading or writing) (Pustejovsky, 1991, 1995; Pustejovsky & Bouillon, 1995). Such an operation results in an enriched composition of the expressions (Jackendoff, 1997), making their semantic representations more complex. Note that the operation could also involve "reconfigurations of relevant syntactic representations" (Traxler et al., 2005, p. 4). Empirically, this type of sentences has been reported to elicit additional RTs and brain activities compared to those with non-aspectual verbs (Katsika et al., 2012; Lai et al., 2017; Zarcone et al., 2017).

The abovementioned observation is also consistent with studies on languages outside the Indo-European language family, such as Mandarin Chinese. Several Chinese empirical studies have found similar processing difficulty with aspectual verb sentences containing EntityNP objects (Ma et al., 2023; Xue et al., 2021a, 2021b, 2024). For instance, Xue et al. (2021a) compared RTs of Chinese speakers for EntityNPs (e.g., 这份病例*zhè-fèn bìnglì* "the medical record") when preceded by aspectual verbs (e.g., 开始 *kāishǐ* "start") and two other non-aspectual verbs (e.g., 阅读/检查 *yuèdú/ jiǎnchá* "read/checked", in which the former denoted a more preferred interpretation than the latter for the aspectual verb expression 医生开始这份病例已经有一会儿了*yīshēng kāishǐ zhè-fèn bìnglì yǐjīng yǒu yīhuìér le* "The doctor started the medical record for a while"). It was found that EntityNPs preceded by aspectual verbs had significantly longer RTs than those preceded by the other verbs. In another study, Xue et al. (2024) compared EntityNPs (e.g., 这本小说*zhè-běn xiǎoshuō* "the novel") preceded by aspectual verbs (e.g., 开始*kāishǐ* "start") and subject-experiencer psych verbs (e.g., 享受*xiǎngshòu* "enjoy") (which have been under controversy whether they have similar selectional restrictions to aspectual verbs i.e., inherently taking an event-denoting object), in relative to control verbs (e.g., 撰写 *zhuànxiě* "write"), as in 作家开始/享受/撰写这本小说*zuòjiā kāishǐ/xiǎngshòu/zhuànxiě zhè-běn xiǎoshuō* "The author started/enjoyed/wrote the novel." The results showed that only aspectual verbs led to increased RTs on EntityNPs. These findings and those of the present study suggest that Chinese aspectual verbs paired with EntityNP objects pose greater processing difficulty during on-line language processing.

Note that this study included five Chinese aspectual verbs as event-selecting verbs, namely, 开始 *kaishǐ* "begin", 继续*jìxù* "continue", 完成 *wánchéng* "finish", 停止 *tíngzhǐ* "stop", and结束 *jiéshù* "end". One may concern that these verbs may introduce heterogeneity due to their different syntactic behaviors. Specifically, aspectual verbs开始 *kaishǐ* "begin", 继续*jìxù* "continue", 停止 *tíngzhǐ* "stop", and结束 *jiéshù* "end" are unaccusative verbs and describe actions or events occurring to the subject. The verb 完成*wánchéng* "finish", on the other hand, is a transitive verb that describes the temporal characteristics of an event, and requires an active agent. To address this concern, two additional statistical analyses were performed on the RTs: One analysis examined whether 完成*wánchéng* "finish" significantly affected the results of the study by removing data from the items with完成 *wánchéng* "finish". The other analysis compared the processing patterns of完成*wánchéng* "finish" with those of the other four aspectual verbs. The results indicated that including the verb 完成*wánchéng* "finish" did not significantly alter the study's findings, nor were there any notable processing differences between完成*wánchéng* "finish" and the other four aspectual verbs (see section 2 of the Supplementary Materials online, where statistical results are shown in Tables S2–S11).

Although Chinese aspectual verbs are acknowledged to semantically require an event-denoting object, some Chinese linguists argued that the event-denoting object may be only represented with a VP (Lin & Liu, 2005; Lu & Wen, 2018; Wen & Lu, 2017). Song (2014, 2015) further pointed out that when this group of verbs combines with an eventive nominal that, though, meets the verb's selectional constraints, an additional verb is still preferably inserted after the matrix verb, either a verb explicitly stating how the event denoted by the nominal is undertaken or a light verb having little semantic content. In this sense, it is possible that the previous reported processing costs of Chinese aspectual verb sentences (Ma et al., 2023; Xue et al., 2021a, 2021b, 2024) may stem from the syntactic anomaly of the absence of a verb between the matrix verb and its object, rather than the complex semantic compositions that have been highlighted previously (Baggio et al., 2010; Frisson & McElree, 2008; Husband et al., 2011; Jackendoff, 1997; Kuperberg et al., 2010; McElree et al., 2006, 2001; Pickering et al., 2005; Pustejovsky, 1991, 1995; Traxler et al., 2005, 2002).

If the assumptions mentioned above were true, we would expect aspectual verbs paired with either EntityNP or EventNP objects to show similar processing patterns. This is because both structures lack a verb between the aspectual verbs and their NP objects, and thus have the same basic sentence structure. We

would also expect that aspectual verbs paired with EventNPs would require more processing effort than their neutral verb equivalents, given the syntactic irregularity involved. However, our findings did not align with these expectations. Instead, we observed that aspectual verbs elicited much longer RTs when paired with EntityNP objects than EventNP objects, with the latter showing processing patterns similar to their neutral verb equivalents.

Our results challenge the previous claim that Chinese aspectual verbs only combine with VPs (Lin & Liu, 2005; Lu, 2020; Lu & Wen, 2018; Wen & Lu, 2017). Rather, our results suggest that these verbs may be able to combine with both VPs and NPs (either an EventNP or EntityNP) while still maintaining their semantic requirements for event interpretation. When the aspectual verbs combine with an EventNP (e.g., 开始这次调色 *kāishǐ zhè-cì tiáo-sè* "start this color-mixing"), the event-denoting object satisfies the verb's selectional requirements, allowing for interpretations that are as straightforward as those neutral verb equivalents (e.g., 准备这次调色*zhǔnbèi zhè-cì tiáo-sè* "prepare this color-mixing"). When the aspectual verbs combine with an EntityNP (e.g., 开始这幅作品*kāishǐ zhè-fú zuòpǐn* "start this artwork"), the coercion operation is needed to shift the object to an event type. This shift resolves the type incompatibility between the verb and the object, but also adds a layer of processing complexity (Baggio et al., 2010; Frisson & McElree, 2008; Husband et al., 2011; Katsika et al., 2012; Kuperberg et al., 2010; Lai et al., 2017; Lowder & Gordon, 2016; McElree et al., 2006, 2001; Pickering et al., 2005; Pylkkänen & McElree, 2007; Spalek & Tomaszewicz, 2018; Traxler et al., 2005, 2002; Zarcone et al., 2017). Thus, selectional constraints are a fundamental property of predicate verbs (Spalek & Tomaszewicz, 2018), guiding how we construct and interpret expressions.

The interpretation is further supported by our acceptability ratings, which showed that both types of aspectual verb sentences were judged as fully acceptable, with ratings comparable to the neutral verb controls (see section "Results"). These results collectively demonstrate that while aspectual verbs semantically select an event, they can syntactically combine with NPs, with processing costs arising specifically when complement coercion is required.

Although Chinese aspectual verbs can be paired with either a VP or a NP, similar to their English counterparts, according to Song (2015), there are constraints to the lexical-semantic properties of the EntityNPs that can follow these verbs in Chinese. Take the Chinese aspectual verb完成*wánchéng* "finish" as an example. For the construction "完成 + EntityNP", the object (e.g., *book*, *essay*) refers to man-made physical objects that imply a clear agentive role (factors involved in the bringing about of the physical object, e.g., writing; see Pustejovsky 1991 for details). The implicit event in this construction is limited to the agentive role of the object's denotation (Song, 2015). In contrast, the English construction "finish + EntityNP" allows for both the agentive role and telic role (purpose of the physical object, e.g., reading; see Pustejovsky 1991 for details) to contribute to the understanding of the unstated event sense (De Almeida, 2004; Pustejovsky, 1991). Therefore, a Chinese expression like 完成这本小说*wánchéng zhè-běn xiǎoshuō* "finish the novel" is more likely to mean "finish writing the novel" rather than "finish reading the novel", whereas both interpretations are available in English.

One may concern whether the longer RTs for the EntityNPs after the aspectual verbs may be simply due to the infrequent co-occurrence between an aspectual verb and an individual classifier. As the EntityNP was structured as [*zhe* $_{\text{demonstrative}}$ + classifier $_{\text{individual}}$ + Noun], the incompatibility between the aspectual verb and the individual classifier could have been detected by the readers when the classifier was presented, before they even read the noun that followed. However, from our perspective, although frequency effects could theoretically contribute to processing differences, our cloze test and mixed-effects models (controlling for the predictability of the NP objects) suggested that the frequency may be not a confounding variable for the present study.

The cloze test (see section "Cloze test") explicitly measured the cloze probability of NP objects following aspectual and neutral verbs. In this test, the sentences were presented with a subject, verb, and demonstrative 这 *zhè* "this", and we asked participants to complete the experimental sentences with a NP (including a classifier and a two-character noun). The cloze probability served as an indicator of the predictability and frequency of the classifiers and the following two-character nouns after aspectual verbs and neutral verbs. Importantly, cloze probability was incorporated as a covariate in our mixed-effects models (see section "Data analysis"). The results did not show a significant effect of cloze probability on RTs in the critical regions (see section "Results"). This indicates that the processing difficulty of aspectual verbs with EntityNPs cannot be attributed simply to low predictability or infrequent co-occurrence between aspectual verbs and the classifiers.

Additionally, the present study was conducted via a self-paced reading paradigm. Via the paradigm, the NP objects appeared as a single word on the screen but not appeared separately with the classifiers first occurring and then the following nouns. Thus, it is more meaningful to consider the co-occurrence frequency between aspectual verbs and the whole NP compositions rather than the frequency between the verbs and the classifiers only. Furthermore, our findings are consistent with prior work on complement coercion, where processing cost was attributed to semantic enrichment of the EventV paired with EntityNP, rather than lexical frequency of the EntityNPs (e.g., Katsika et al. 2012, Ma et al. 2023, McElree et al. 2001, Traxler et al. 2002, Xue and Liu 2021c, Xue et al. 2021a, 2021b, 2024).

This argument can be supported by two key observations in this study. First, neutral verbs showed no significant RT differences between EntityNP ([*zhe* $_{\text{demonstrative}}$ + classifier $_{\text{individual}}$ + Noun]) and EventNP ([*zhe* $_{\text{demonstrative}}$ + classifier $_{\text{event}}$+ Noun] objects (see section "Results"). Second, Spillover effects (NP +1 and NP +2) are classically associated with compositional semantics rather than lexical access (Traxler et al., 2002; Husband et al., 2011)

One may further question that the observation—no apparent RT differences were found between the aspectual verb sentences with EventNPs and the neutral verb equivalents—may not necessarily mean that the former are well-formed sentences. A few Chinese studies suggested that syntactic anomalies may not block access to the semantics of expressions; readers can often interpret meanings based on context (Yang et al., 2015; Zhang et al., 2013, 2010). However, if there are syntactic anomalies involved in the aspectual verb sentences, then when they combine with an EventNP, the corresponding expressions would be generally considered ungrammatical. Yet, our acceptability data (see section "Results") showed that either an EventNP or an EntityNP followed the aspectual verbs, the combinations were well-accepted.

Taking together the acceptability ratings and RT data, we may conclude that these sentences, despite their potential syntactic irregularities, were deemed acceptable by our participants. Future studies may employ more sophisticated methods, such as EEG and fMRI, to further validate these findings, and explore the neural mechanism underlying the processing of such sentences.

Our findings also shed light on the processing costs associated with complement coercion—involving aspectual verbs plus EntityNP objects—which seems to arise not from the verb or object per se but from their interactions (Traxler et al., 2002). The

lack of significant RT differences between the two types of neutral verb sentences suggested that the substantial RT differences between the two types of aspectual verb sentences were not due to the EntityNP object itself, relative to EventNP objects. Similarly, no significant RT differences were observed between EventNP objects preceded by aspectual versus neutral verbs, indicating that the notable RT differences between the two types of EntityNP sentences were not due to the verb itself. This interactive effect in complement coercion aligns with our acceptability ratings (see section "Results").

In sum, this study provides empirical evidence showing that Chinese aspectual verbs exhibited distinct processing patterns when combined with different types of NP objects, with aspectual verbs paired with EntityNP objects requiring more cognitive efforts than those paired with EventNP objects. This research advances understanding Chinese aspectual verbs in at least three aspects. Firstly, it aligns with existing literature by showing that sentences with aspectual verbs and complement coercion—like those with EntityNP objects—are more taxing to process than those without complement coercion. Secondly, between the NP objects following aspectual verbs, only the EntityNP objects, which do not match the selectional restrictions of the verbs, incur processing costs. The costs are likely due to the semantic enrichment needed to reconcile the type mismatch between the verb and the object, which exceeds the effect of any syntactic irregularities if there are. Thirdly, the study also demonstrates the interplay between verb and object types in sentence interpretation, reinforcing that interpreting complement coercion relies on neither the verb nor the object itself but their interaction.

While our behavioral data challenge the idea that Chinese aspectual verbs only pair with VPs, they do not, of course, exclude the possibility that the expressions, including aspectual verbs and NP objects, may engage different brain regions or require additional brain activities compared to their counterparts including non-aspectual verbs. Future studies may employ more robust experimental methods like EEG and fMRI to validate the findings. This study makes a pioneering effort to empirically examine the subcategories of Chinese aspectual verbs, offering fresh insights for Chinese linguists to delve deeper into the linguistic characteristics of these verbs and to confirm the results presented here.

Conclusion

This study has demonstrated that processing Chinese aspectual verbs is significantly influenced by the type of NP objects. By comparing self-paced reading times of native Chinese speakers on EntityNP objects and EventNP objects following aspectual verbs as opposed to neutral verbs, respectively, our main findings reveal that aspectual verbs with EntityNP objects are more challenging to process than those with EventNP objects, and such a difference is not observed with neutral verbs; moreover, aspectual verbs with EntityNP objects are also more challenging to process than those with neutral verbs. The results suggest that the processing difficulty of aspectual verbs with EntityNPs (complement coercion) stems from the enriched semantic composition required to reconcile the type mismatch between the Chinese aspectual verb and the EntityNP object. Such an effect extends beyond the effect of syntactic irregularities (if there are) between the matrix verb and its object. Therefore, selectional restrictions are an essential characteristic of a predicate verb, which determines the compatibility of a verb with its argument and how we construct and interpret sentences. Our findings underscore the importance of considering both semantic and syntactic factors when examining the processing of aspectual verbs in Chinese, and they pave the way for future studies to build upon this groundwork.

Data availability

All data and R-codes for statistical analyses can be found at https://osf.io/fzc4g/?view_only=4acc8ec21d6145ba8a14789cf6ef7953.

Received: 15 November 2024; Accepted: 9 October 2025;
Published online: 20 December 2025

References

Baggio G, Choma T, Van Lambalgen M, Hagoort P (2010) Coercion and compositionality. J Cogn Neurosci 22(9):2131–2140
Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67(1):1–48
De Almeida RG (2004) The effect of context on the processing of type-shifting verbs. Brain Lang 90(1-3):249–261
De Almeida RG, Riven L, Manouilidou C, Lungu O, Dwivedi VD, Jarema G, Gillon B (2016) The neuronal correlates of indeterminate sentence comprehension: an fMRI Study. Front Hum Neurosci 10:614
Delogu F, Crocker MW, Drenhaus H (2017) Teasing apart coercion and surprisal: Evidence from eye-movements and ERPs. Cognition 161:46–59
Forster KI, Forster JC (2003) DMDX: A Windows display program with millisecond accuracy. Behav Res Methods Instrum Comput 35(1):116–124
Frisson S, McElree B (2008) Complement coercion is not modulated by competition: Evidence from eye movements. J Exp Psychol Learn Mem Cogn 34(1):1–11
Han L (2016) The definition and system construction of event nouns in Mandarin Chinese J. East China Norm Univ (Humanities and Social Sciences) 48(5):161–175
Hsu C-C, Hsieh S-K (2013) To coerce or not to coerce: A corpus-based exploration of some complement coercion verbs in Chinese. In: Pustejovsky, J (ed) Proceedings of the 6th International Conference on Generative Approaches to the Lexicon (GL2013) pp 13–20
Husband EM, Kelly LA, Zhu DC (2011) Using complement coercion to understand the neural basis of semantic composition: evidence from an fMRI study. J Cogn Neurosci 23(11):3254–3266
Jackendoff R (1997) The architecture of the language faculty. Cambridge, MA: MIT Press
Jespersen O (1965) A modern English grammar on historical principles, part VI, morphology. London: George Allen and Unwin Ltd
Katsika A, Braze D, Deo A, Piñango MM (2012) Complement coercion: Distinguishing between type-shifting and pragmatic inferencing. Ment Lex 7(1):58–76
Kuperberg GR, Choi A, Cohn N, Paczynski M, Jackendoff R (2010) Electrophysiological correlates of complement coercion. J Cogn Neurosci 22(12):2685–2701
Kutas M, Federmeier KD (2011) Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu Rev Psychol 62(1):621–647
Kuznetsova A, Brockhoff PB, Christensen R (2017) lmerTest package: Tests in linear mixed effects models. J Stat Softw 82(13):1–26
Lai Y-Y, Lacadie C, Constable T, Deo A, Piñango MM (2017) Complement coercion as the processing of aspectual verbs: evidence from self-paced reading and fMRI. In: Hampton JA, Winter Y (eds) Compositionality and concepts in linguistics and psychology pp 191–222. Cham: Springer
Lapata M, Keller F, Scheepers C (2003) Intra-sentential context effects on the interpretation of logical metonymy. Cogn Sci 27(4):649–668
Levin B (1993) English verb classes and alternation. Chicago: The University of Chicago Press
Lin S-Y, Hsieh S-K, Huang Y-J (2009) Exploring Chinese type coercion: A web-as-corpus study. Paper presented at the 5th International Conference on Generative Approaches to the Lexicon, Pisa, Italy
Lin T-HJ, Liu C-YC (2005) Coercion, event structure, and syntax. Nanzan Linguist 2:9–31
Liu D (2010) Chinese as a verby language: On typological differences between verby language and nouny languages. Chin Teach World 1:3–17
Lowder MW, Gordon PC (2016) Eye-tracking and corpus-based analyses of syntax-semantics interactions in complement coercion. Lang Cogn Neurosci 31(7):921–939
Lu Z (2020) Hanyu tidongci de shijian zhicheng texing [Temporal reference properties of Chinese aspectual verbs]. J Beijing Int Stud Univ 42(6):106–123
Lu Z, Wen B (2018) Jiyu dongci yuyixue de hanyu tidongci fenxi [A semantic anlaysis of aspectual verbs in Chinese based on verb semantics]. J PLA Univ Foreign Lang 41(2):80–103
Ma Y, Buccola B, Wang Z, Cousins S, Godfroid A, Beretta A (2023) Expressions with aspectual verbs elicit slower reading times than those with psychological verbs: An eye-tracking study in Mandarin Chinese. J Psycholinguist Res 52(1):179–215

McElree B, Frisson S, Pickering MJ (2006) Deferred interpretations: Why starting Dickens is taxing but reading Dickens isn't. Cogn Sci 30(1):181–192
McElree B, Pylkkänen L, Pickering MJ, Traxler MJ (2006) A time course analysis of enriched composition. Psychon Bull Rev 13(1):53–59
McElree B, Traxler MJ, Pickering MJ, Seely RE, Jackendoff R (2001) Reading time evidence for enriched composition. Cognition 78(1):B17–B25
Pesetsky DM (1995) Zero syntax: Experiencers and cascades: Cambridge, MA: MIT press
Pickering MJ, McElree B, Traxler MJ (2005) The difficulty of coercion: A response to de Almeida. Brain Lang 93(1):1–9
Piñango MM, Deo A (2016) Reanalyzing the complement coercion effect through a generalized lexical semantics for aspectual verbs. J Semant 33(2):359–408
Politzer-Ahles S (2017) An extension of within-subject confidence intervals to models with crossed random effects. https://www.tqmp.org/RegularArticles/vol13-1/p075/
Pustejovsky J (1991) The generative lexicon. Comput Linguist 17(4):409–441
Pustejovsky J (1995) The generative lexicon. Cambridge, MA: MIT press
Pustejovsky J, Bouillon P (1995) Aspectual coercion and logical polysemy. J Semant 12(2):133–162
Pylkkänen L, McElree B (2006) The syntax-semantics interface: On-line composition of sentence meaning. In: Traxler M, Gernsbacher MA (eds) Handbook of psycholinguistics pp 539–579. New York: Elsevier
Pylkkänen L, McElree B (2007) An MEG study of silent meaning. J Cogn Neurosci 19(11):1905–1921
Qin H, Wang K (2012) A usage-based analysis of logical metonymies with aspectual verbs. J Foreign Lang 35(5):29–37
R Core Team (2018) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing
Smith N, Levy R (2013) The effect of word predictability on reading time is logarithmic. Cognition, 128(3):302-319
Song Z (2014) A comparative study of event coercion between Chinese and English. Lang Linguist 15(2):199
Song Z (2015) Sheng-chéng-cí-kù lǐ-lùn yǔ hàn-yǔ shì-jiàn qiǎng-pò xiàn-xiàng yán-jiū [Generative Lexicon and Event Coercion in Chinese]: Peking University Press
Spalek AA, Tomaszewicz BM (2018) Complement coercion in Polish and the role of selectional restrictions revealed in a self-paced reading study. In: Truswell R, Cummins C, Heycock C, Rabern B, Rohde H (eds) Proceedings of Sinn und Bedeutung 21. pp 1141–1158
Traxler MJ, McElree B, Williams RS, Pickering MJ (2005) Context effects in coercion: Evidence from eye movements. J Mem Lang 53(1):1–25
Traxler MJ, Pickering MJ, McElree B (2002) Coercion in sentence processing: Evidence from eye-movements and self-paced reading. J Mem Lang 47(4):530–547
Wen B, Lu Z (2017) Hanyu tidongci de jufa yuyi fenxi [The syntactic properties of Chinese aspectual verbs and their syntax-semantic analysis] J Guangdong Univ Foreign 28(3):29–35
Xue W, Liu M (2021c) Complement coercion in Mandarin Chinese: evidence from a self-paced reading study. J Psycholinguist Res 50(4):757–776
Xue W, Liu M, Politzer-Ahles S (2021a) Processing of complement coercion with aspectual verbs in Mandarin Chinese: Evidence from a self-paced reading study. Front Psychol 12:1973
Xue W, Liu M, Politzer-Ahles S (2021b) A study of complement coercion in Mandarin Chinese: evidence from an acceptability judgment task. In: Liu M, Kit C, Su Q (eds) Chinese lexical semantics. pp 775–784. Switzerland: Springer
Xue W, Liu M, Politzer-Ahles S, Tzeng OJ-L (2024) Verbal effect on the processing of complement coercion: Distinguishing between aspectual verbs and psych verbs. Lingua 306:103754
Yang Y, Wu F, Zhou X (2015) Semantic processing persists despite anomalous syntactic category: ERP evidence from Chinese passive sentences. PLoS ONE 10(6):e0131936
Zarcone A, McRae K, Lenci A, Padó S (2017) Complement coercion: The joint effects of type and typicality. Front Psychol 8:1987
Zhang Y, Li P, Piao Q, Liu Y, Huang Y, Shu H (2013) Syntax does not necessarily precede semantics in sentence processing: ERP evidence from Chinese. Brain Lang 126(1):8–19
Zhang Y, Yu J, Boland JE (2010) Semantics does not need a processing license from syntax in reading Chinese. J Exp Psychol Learn Mem Cogn 36(3):765

Acknowledgements

The authors would like to thank Bomiao Liang, Songquan Wei, Zewen Liu, Wanqi Luo for their help on data collection of this study. This work was supported by grants from the Hubei Provincial Social Science Fund Early-stage Funding Project (grant number: 23ZD238) and the Fundamental Research Funds for the Central Universities HUST (grant number: 2025WKQN007). This work was supported by grants from the Humanities and Social Science Youth Foundation of Chinese Ministry of Education (grant number: 2025YJC740060), the Fundamental Research Funds for the Central Universities HUST (grant number: 2025WKQN007), and the Hubei Provincial Social Science Fund Early-stage Funding Project (grant number: 23ZD238).

Author contributions

Wenting Xue designed the experiment, performed the data analysis, wrote the original draft, and revised the manuscript; Jinmeng Dou performed the data analysis, wrote the original draft, and revised the manuscript; Wenxin Qin collected the data, and edited the original draft and the revised manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Ethical approval

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of the Huazhong University of Science and Technology (protocol code 20230621S440) on June 21, 2023.

Informed consent

Written informed consent was obtained from all participants prior to their involvement in the study. The acceptability judgment task was conducted from July 1st to July 25th, 2023. The cloze test was carried out between August 1st and September 15th, 2023. The self-paced reading experiment took place from October 15th to November 15th, 2023. Participants were informed that they had the right to withdrawal from the study at any time without any negative consequences, penalty, or loss of benefits to which you would otherwise be entitled. They were also assured that their data would remain anonymous and be treated with strict confidentiality. Specifically, the data would be securely stored on a password-protected laptop, accessible only to the researcher. No participants' name or identifying information would be released or included in any report or publications.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-025-06086-y.

Correspondence and requests for materials should be addressed to Wenting Xue.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.