Language and Cognition (2026), 18, e48, 1–29
doi:10.1017/langcog.2026.10097

ARTICLE

Diachronic competition and functional differentiation in object-involving complex directional complement structures: a multifactorial case study of *qǐlái*

Yuhang Yang and Fuyin Thomas Li

School of Foreign Languages, Beihang University, Beijing, China
Corresponding author: Fuyin Thomas Li; Email: thomasli@buaa.edu.cn

(Received 18 August 2025 UTC; Revised 25 May 2026 UTC; Accepted 05 June 2026 UTC)

Abstract

This study investigates the diachronic evolution of four object-involving syntactic structures formed with the Mandarin complex directional complement (CDC) *qǐlái* 'get up' from the 12th to early 20th centuries, using data from the Center for Chinese Linguistics (CCL) corpus. Frequency distributions and multiple correspondence analysis (MCA) show that these structures did not undergo simple replacement but instead developed through gradual functional differentiation. Mid-object (mid-O) emerged earliest and remained the most frequent pattern, developing as a relatively neutral and processual structure more strongly associated with event unfolding and with new, indefinite and non-patient objects. Pre-object (pre-O) appeared later, peaked in the 17th century and was increasingly restricted to state-change events, a development related to its verb–object (V–O) adjacency and subsequent competition from ba-O. Ba object (ba-O) became increasingly prominent over time and reached its highest proportion in the 19th century, becoming specialized for foregrounding cognitively accessible and highly affected objects. Although post-O remained rare, qualitative evidence suggests a degree of specialization in evaluative uses, where the postverbal object serves as an anchor for subsequent judgment or description. These findings contribute to a usage-based understanding of syntactic change by showing how formally related variants come to be differentiated through the interaction of structural, cognitive and language-specific factors.

Keywords: Chinese complex directional complements (CDCs); diagrammatic iconicity; language competition; multifactorial analysis; object positioning

1. Introduction

Directional complements in Mandarin, defined as directional verbs serving as complements to verbs or adjectives (Liu, 1998), have been extensively studied since the early 20th century (e.g., Chao, 1968; Fan, 1963; Li, 1924; Li & Thompson, 1981;

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

Liang, 2007; Lyu, 1980) and continue to attract scholarly attention from diverse theoretical perspectives (e.g., Cao, 2023; Huang & Hsieh, 2008; Peyraube, 2006; Song, 2007; Wang, 2005; Yang & Li, 2025). One major line of research focuses on syntactic structures in which objects co-occur with simplex or complex directional complements (CDCs). With respect to CDCs, four syntactic patterns are typically identified (Lyu, 1980; Zhang, 1991b; Zhang & Fang, 1996): pre-object (pre-O), mid-object (mid-O), post-object (post-O) and *ba*-object (ba-O). These structures share broadly comparable syntactic and semantic properties but differ primarily in word order (Zhang, 1991a, *p.* 5). Crucially, however, they are not always interchangeable. For example, in (1), the object *huà* 'word' must appear in the mid-O pattern (see [1a]); the post-O pattern is marginal (see [1b]), while the pre-O and ba-O patterns are unacceptable (see [1c–d]).

(1) a. mid-O: *shuō qǐ huà lái* 'start to speak'
b. post-O: *shuō qǐlái huà* 'start to talk'
c. pre-O: *shuō huà qǐlái** (ungrammatical)
d. ba-O: *bǎ huà shuō qǐlái** (ungrammatical)
(Zhang, 1991a, p. 8)

Previous research has examined these patterns from multiple perspectives, including their semantic and pragmatic contrasts (Guo, 2003; Yang, 2005), usage constraints (Chen, 1980, 2007; Jia, 1998; Lu, 2002; Yang, 2006; Zhang & Fang, 1996), diachronic developments (Hao, 2018; Hiroshi, 2021; Liang, 2007; Zhang, 1991b) and prosodic properties (Dong, 1998; Song, 2019). However, relatively few studies have integrated large-scale diachronic corpus analysis and advanced statistical modeling within a cognitive framework to examine how the usage preferences and functional distinctions among the four structures have evolved over time. Such an approach can provide a more empirically grounded and methodologically rigorous account of how syntactic differences emerge from a usage-based perspective, while also contributing Chinese evidence to broader discussions of language change. Against this background, the present study examines the diachronic development of the four object-involving syntactic patterns of the CDC *qǐlái* 'get up' from the 12th century to the early 20th century, drawing on data from the Center for Chinese Linguistics (CCL) corpus and adopting a mixed-methods approach that combines quantitative modeling with qualitative analysis.

The paper is organized as follows. Section 2 introduces the basic concepts of Mandarin CDCs, with particular attention to *qǐlái* 'get up', and reviews previous studies on object-involving CDC structures. Section 3 describes the methodology. Section 4 presents the empirical findings, combining quantitative results with qualitative analysis. Section 5 offers a cognitive account of the competition and functional differentiation among the structures. Finally, Section 6 summarizes the main findings, discusses the limitations of the study and suggests directions for future research.

2. Research on Mandarin CDCs and object-involving CDC structures

2.1. Overview of Mandarin CDCs and qǐlái 'get up'

Liu (1998) identifies a closed set of 28 Mandarin directional complements, including 11 simplex and 17 complex forms (see Supplementary Appendix A; see also Li, 2018,

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

p. 595). CDCs are typically formed by combining a simplex directional complement ranging from C_3 to C_{11} with either *lái* 'come' (C_1) or *qù* 'go' (C_2). An exception is *qǐqù* 'rise-go', which deviates from conventional compounding patterns and is generally regarded as non-standard (Zhu, 1982). Among them, *qǐlái* 'get up' – the primary focus of this study – originally denotes 'movement from a lower position to a higher one' (Liu, 1998, p. 341), as illustrated by expressions such as *zhàn-qǐlái* 'stand up'. Through progressive grammaticalization, *qǐlái* has developed several extended meanings when functioning as a complement. According to the *Modern Chinese Dictionary* (Dictionary Editorial Office, Institute of Linguistics, Chinese Academy of Social Sciences, 2016), *qǐlái* can (1) indicate the initiation or continuation of an action or state, as in *dàxiào-qǐlái* 'burst into laughter'; (2) denote the completion of an action or the attainment of a result, as in *xiǎng-qǐlái* 'remember-PFV'; and (3) express estimation or draw attention to a particular aspect, as in *kàn-qǐlái* 'it seems'. Scholars have noted that although all directional complements generally undergo semantic extension toward resultative meanings, only some can additionally express inchoative or continuative meanings related to a state or event (Guan & Guo, 2024; Liu, 1998), a property that distinguishes *qǐlái* from other complements.

Due to its semantic diversity and high degree of grammaticalization, *qǐlái* has attracted substantial scholarly attention, including studies of its syntactic structures (e.g., Song, 1980; Wang & Wen, 2015), semantics (e.g., Cao, 2023; Fang, 1992; Liu, 1998; Qi & Zeng, 2009; Song, 2007), diachronic developments (e.g., Liang, 2007; Wang, 2005; Zhang & Fang, 1996) and acquisition patterns and usage errors (e.g., Yang, 2003; Zhu, 2019). However, its co-occurrence with objects has received comparatively little attention. As previous studies have rarely examined this issue with a focus on a single CDC, Section 2.2 provides a general review of object-involving CDC syntactic structures to motivate and contextualize the present study.

2.2. Previous research on four object-involving CDC structures

A CDC may occur either without an object (e.g., *zuò-qǐlái* 'sit up') or with an object (e.g., *zuò qǐ shēn lái* 'sit up'). When an object is present, however, its position is variable. Previous studies (Lyu, 1980; Zhang, 1991b; Zhang & Fang, 1996) have identified four distinct syntactic patterns:

Pre-object (pre-O): $VOC_{3\text{-}11}C_{1/2}$ (e.g., *tái tóu qǐlái* 'lift one's head').
Mid-object (mid-O): $VC_{3\text{-}11}OC_{1/2}$ (e.g., *tái qǐ tóu lái* 'lift one's head').
Post-object (post-O): $VC_{3\text{-}11}C_{1/2}O$ (e.g., *tái qǐlái tóu* 'lift one's head').
Ba-object (ba-O): *bǎ*$OVC_{3\text{-}11}C_{1/2}$ (e.g., *bǎ* tóu tái qǐlái 'lift one's head').

While some scholars (e.g., Zhang, 1991a) argue that these four structures share similar syntactic and semantic properties and differ primarily in word order, others (e.g., Song, 2019) maintain that they vary in typicality and markedness. Among them, mid-O, which emerged during the Song–Yuan period (roughly from the 10th to 14th centuries) (Wang, 2005), is generally regarded as the prototypical form: it is the most frequent, subject to fewer constraints (Chen, 1980; Jia, 1998; Zhang, 1991b) and often serves as the basis for analyzing other structures (Hiroshi, 2021). Post-O developed later but only became established in Modern Chinese (Lyu, 1985; Wang, 2005; Zhang & Fang, 1996); it favors indefinite objects (Jia, 1998; Zhang, 1991b) and readily licenses long clausal objects (Zhang, 1991a). Pre-O, by contrast, is considered the

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

earliest type (Zhang, 1991b), though it is relatively uncommon in the Beijing dialect at present. Finally, ba-O, a subtype of the narrow disposal *ba* construction (Liu, 2025), expresses affectedness and marks the object as a patient undergoing an explicit change of state (Liu & Cuyckens, 2025). Like the broad disposal construction (Wang, 2014; Wu, 2003; Jing-Schmidt & Tao, 2009), it is characterized by OV order and exhibits a strong preference for definite objects (Hao, 2018; Jia, 1998; Zhang & Fang, 1996) and given information (Fang & Liu, 2021).

Early descriptive work (e.g., Chen, 1980; Jia, 1998; Liu, 1998; Zhang, 1991a, 1991b; Zhang & Fang, 1996; Zhu, 1982) laid the theoretical groundwork for subsequent investigations by outlining the syntactic constraints and semantic–pragmatic distinctions among the four structures. Zhang and Fang (1996, Chapter 7), for instance, offered a systematic analysis showing that syntactic choice is conditioned by multiple factors, including the syntactic and semantic types of the object, the semantics of the CDC and clausal mood. Building on this line of research, Lu (2002) classified the main verbs occurring in object-involving structures and examined their interaction with different object types, numeral modifications and the perfective marker *le*. Subsequent studies (e.g., Guo, 2003; Hao, 2018) further confirm that the four patterns are constrained by a range of factors, such as object length, definiteness, thematic roles, the presence of *le*, verb semantics and the meaning of the CDC (for further details, see Section 3.2). Scholars have also examined the diachronic development of these structures, identifying their periods of emergence and order of appearance (Liang, 2007; Wang, 2005; Zhang, 2018), as well as the forces motivating their change (Hao, 2018; Hiroshi, 2021). To account for their distinctions and constraints, explanations have been proposed from both functional–grammatical perspectives (Hao, 2018; Li, 2017; Zhang & Fang, 1996) and cognitive–linguistic approaches (Yang, 2005, 2017; Yang, 2006; Zhao, 2004). The latter draws on notions such as tense iconicity (Haiman, 1983) and scanning (Langacker, 1987). Yang (2005), for example, applies Dai and Huang's (1988) principle of temporal sequence to argue that object position reflects different cognitive processes. In addition, prosodic accounts (e.g., Dong, 1998; Song, 2019) have analyzed rhythmic and structural factors to explain the unmarked status and high frequency of mid-O.

Although previous studies have yielded valuable insights, several limitations remain. First, much of the existing research (e.g., Hao, 2018; Lu, 2002; Zhang & Fang, 1996) treats the 17 CDCs as a homogeneous set and examines them collectively, which risks producing overly general conclusions that obscure the distinctive properties of individual CDCs. Second, most studies rely on single-genre corpora (e.g., literary texts such as novels) and primarily employ simple frequency counts, while relatively few draw on large-scale, diverse corpora. As a result, the proposed constraints have been based largely on qualitative analysis and limited quantitative evidence, without exploiting corpus data and advanced multivariate exploratory techniques to systematically capture usage preferences and differences across object-involving structures. Third, cognitive accounts of the diachronic evolution of the four structures and their semantic and functional distinctions remain limited. To address these gaps, this study seeks to answer the following research questions:

(1) How did the frequency distributions of the four object-involving *qǐlái* syntactic structures change from the 12th century to the early 20th century?

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

(2) What usage features were associated with each structure, and did these preferences remain stable or change over time?
(3) What cognitive factors motivate their diachronic development and account for their semantic and functional differences?

3. Methodology

3.1 Data collection

Data were drawn from the CCL Corpus (Zhan et al., 2019), a 5.84-billion-token resource covering multiple genres, including colloquial language, literature, newspaper, practical writing and drama. The corpus has been widely used in research on Chinese linguistics (e.g., Chen, 2025; Liu & Li, 2024; Yang & Li, 2025). More specifically, the present study draws on the Corpus of Historical Chinese (CHC), a sub-corpus containing over 1.09 billion tokens. Among these, 41.05% are classified by dynasty, spanning 14 dynasties from the Western Zhou (1046 BC – 771 BC) to the Qing dynasty (1636–1911).

Following Li et al. (2023) and Tian et al. (2025), the texts were further categorized by century to achieve greater temporal precision than that provided by dynasty-based classification. As object-involving *qǐlái* structures became relatively frequent only from the Southern Song period onward, data retrieval was limited to texts from the Southern Song dynasty (12th–13th centuries) through the Qing dynasty (17th–early 20th centuries), in accordance with the dynasty-based search design of the CHC. Since the CHC does not consistently provide precise information on authorship or date of composition, these details were supplemented with reference to other authoritative sources. Following Li et al. (2023), when the year of composition was unavailable, it was estimated as the midpoint between the author's birth and death years. Texts spanning multiple centuries or lacking a clearly identifiable century of composition were excluded. The final dataset consists of 311 books, whose distribution across centuries and genres is shown in Supplementary Appendix B. Because of limited data in the early periods, the first three centuries were collapsed into a single period.

To retrieve the relevant data as comprehensively as possible, four query patterns available through the CHC pattern-search function – *(v)*起来*(X)*, *(v)*起*(X)*来, *(v)(X)*起来 and 把*(X)(v)*起来 – were used. All retrieved instances were then manually inspected, and spurious hits were excluded. These included cases in which *(X)* was a function word (e.g., *de*, *bù*, *le*, *jiāng*) rather than a lexical noun phrase, as well as cases in which *(v)* did not function as an independent lexical verb (e.g., modal or auxiliary uses). After filtering, a total of 4,641 valid tokens were retained, as shown in Table 1.

Moreover, since mid-O was substantially more frequent than the other structures, especially from the 18th to 20th centuries, 300 mid-O tokens were randomly sampled from each century between the 17th and 20th centuries to make manual annotation feasible for subsequent quantitative analyses, following established practices in corpus-based diachronic research (e.g., Chen, 2025; Liu & Li, 2024). Given the extreme rarity of post-O in the corpus, it was excluded from the statistical modeling to mitigate the effects of data sparsity. In total, 1,991 concordance lines (see Table 2) were annotated for subsequent modeling, as discussed in the following section.

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

Table 1. Raw frequency of valid tokens across centuries

Century	Pre-O	Mid-O	Post-O	Ba-O	Total
12th	0	5	0	0	5
12th–13th	1	9	0	0	10
13th	0	4	0	0	4
13th–14th	9	107	3	4	123
14th	0	15	0	0	15
16th	3	85	1	2	91
17th	58	430	3	24	515
18th	75	1161	4	91	1331
19th	44	968	7	148	1167
20th	51	1269	4	56	1380
Total	241	4053	22	325	4641

Table 2. Distribution of annotated data across centuries for quantitative analyses

Century	Pre-O	Mid-O	Ba-O	Total
12th–16th	13	225	6	244
17th	58	300 (sampled)	24	382
18th	75	300 (sampled)	91	466
19th	44	300 (sampled)	148	492
20th	51	300 (sampled)	56	407
Total	241	1425	325	1991

3.2. Data annotation

To examine the diachronic dynamics of the object-involving structures in terms of their usage features, the dataset was further annotated on the basis of previous findings and the characteristics of the data. In total, ten variables capturing object-related, verb-related, CDC-related and other relevant properties were coded. The coding scheme is presented in Table 3, and each variable is explained in detail below.

(1) **Object length (Label: *Oblength*)** is a well-established factor that influences word order in Chinese. Longer objects, especially clausal ones (Zhang, 1991a), tend to occur in post-O, whereas shorter objects favor pre-O, followed by mid-O and ba-O (Song, 2019), aligning with the 'short-before-long' preference (Chao, 1968; Hawkins, 1983). For multiple correspondence analysis (MCA), object length was categorized as short (1–2 characters), medium (3–5 characters) and long (6–23 characters).

(2) **Object semantics (Label: *Obsemantics*)** were classified into three types:[1] animate entities, including humans and animals; concrete entities, referring to physically perceptible objects; and abstract entities, including intangible concepts such as affective states (e.g., *nù* 'anger', *diāo* 'cunning') and events or actions (Chen, 2025; Croft, 2022; Zaenen et al., 2004). This classification not only refines earlier typologies of object types (e.g., Liu, 1998; Lu, 2002; Lyu, 1985; Zhu, 1982), which often conflated semantic categories with thematic roles and employed relatively coarse distinctions (e.g., agentive, patient, locative categories), but also captures animacy

[1]Since *qǐ* is intransitive, the locative-object use frequently reported in earlier studies (Lyu, 1980; Zhang, 1991b) is not attested in the present data; accordingly, no separate locative category was included in the semantic classification.

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

Table 3. Coding scheme

Constraints		Label	Categories
Object	Length	Oblength	Short, medium, long
	Semantics	Obsemantics	Animate beings, concrete entities, abstract entities
	Definiteness	Obdefiniteness	Yes, no
	Information status	Obinformationstatus	New, given
	Personal pronoun?	Obpersonpronoun	Yes, no
	Thematic role	Obthematicrole	Patient, non-patient (instrument, theme, agent, experiencer, goal, location, recipient, role, source)
Verb	Semantics	Vsemantics	Activity, movement, physical causative, state-change causative
Le	Appearance	Leappearance	Yes, no
CDC	Event type	Eventtype	Motion, state change, realization, temporal contouring (abbreviated as temporal)
Others	Mood	Mood	Declarative, imperative, subjunctive, interrogative

and concreteness, two factors widely recognized in studies of word order variation (Branigan et al., 2008; Vogels & Van Bergen, 2017).

(3) **Object definiteness (Label: *Obdefiniteness*)** has been shown to affect syntactic choices (Hao, 2018; Zhang, 1991b; Zhu, 1982). Pre-O and post-O typically co-occur with indefinite objects (Hiroshi, 2021; Zhang, 1991a, 1991b), ba-O strongly favors definite objects (Chen, 2004; Lyu, 1980), and mid-O accommodates both but tends toward indefinite ones (Zhang, 1991b). As Mandarin lacks morphological definiteness marking, definiteness was inferred from lexical and contextual cues. Following Chen's (2004) widely adopted classification (e.g., Li et al., 2023; Liu & Cuyckens, 2025), objects were coded for definiteness as a binary yes/no variable.

(4) **Object information status (Label: *Obinformationstatus*)** was annotated following established practices (e.g., Fang & Liu, 2021; Xu et al., 2024). A referent was coded as given if it had been mentioned within the preceding ten lines or realized as a pronoun; otherwise, it was coded as new. Previous studies suggest that syntactic choices tend to follow the given-before-new principle (Arnold et al., 2000): post-O and mid-O often introduce new information, whereas ba-O predominantly encodes given information (Hao, 2018;Zhang, 1991b; Zhang & Fang, 1996).

(5) **Object pronominality (Label: *Obpersonpronoun*)** was annotated as a binary variable, indicating whether the object was a personal pronoun. Although this factor has received limited attention in previous research, manual inspection of the data indicates that personal pronouns tend to appear more frequently in pre-O and ba-O than in mid-O and post-O, a pattern consistent with the pronoun precedence principle (Bresnan et al., 2007).

(6) **Object thematic role (Label: *Obthematicrole*)** was coded into two categories, patient and non-patient, based on the dataset. As previously mentioned, the boundary between semantic class and thematic role in object classification is often blurred (e.g., Chen, 1980; Lu, 2002). Although an initial fine-grained taxonomy, including instrument, theme, patient, and agent (Yang, 2011), was adopted, these roles were subsequently merged due to data sparsity.

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

(7) **Verb semantics (Label: *Vsemantics*)** was coded using a four-way classification: activity verbs, which are atelic and lack affectedness or causative semantics (e.g., *shuō* 'speak', *hē* 'drink', *xiǎng* 'think'); verbs of movement, which describe spontaneous, self-initiated movement in physical space (e.g., *pǎo* 'run', *zǒu* 'walk'); physical causative verbs, which encode the application of physical force resulting in spatial displacement (e.g., *sòng* 'deliver', *lā* 'pull'); and state-change causative verbs, which denote agent-induced changes of state or property (e.g., *suǒ* 'lock', *shā* 'kill').

(8) **The presence of the perfective marker *le* (Label: *Leappearance*)** was coded as a binary variable (yes/no), given its interaction with mood reported in previous studies (Lu, 2002; Zhang, 1991b). For instance, Zhang (1991b) observes that *le* is incompatible with post-O in declarative contexts and with mid-O in imperative contexts. In pre-O, the presence of *le* is typically associated with declarative mood, whereas its absence is generally associated with imperative readings.

(9) CDC semantics has been shown to influence syntactic choice and is traditionally classified into three types: directional, resultative and stative (Liu, 1998; Song, 2007; Wang, 2005). However, this tripartite classification remains relatively coarse. Recent studies adopting a macro-event perspective (Li, 2018; Yang & Li, 2025) demonstrate that CDCs can encode a wider range of macro-event types. Accordingly, this study adopts Talmy's (Talmy, 2000, p. 214) macro-event typology **(Label: *Eventtype*)**, distinguishing motion, temporal contouring, state change and realization events. This framework subsumes the traditional categories while providing a more nuanced basis for capturing the semantic diversity of *qǐlái*.

(10) **Mood (Label: *Mood*)** was coded into four categories: declarative, imperative, subjunctive and interrogative, given its interaction with *le* in shaping syntactic choice (Zhang, 1991b; Zhang & Fang, 1996).

3.3. Data analysis

First, the full dataset before sampling (see Table 1) was used to trace the emergence and diachronic distribution of the four object-involving *qǐlái* structures, and a chi-square test was conducted to assess the statistical significance of distributional differences across periods. All statistical analyses and visualizations were conducted in R version 4.4.2 (R Core Team, 2024). Frequency distributions were visualized using *ggplot2* version 3.5.2, with data preparation carried out using *dplyr* version 1.1.4 and *tidyr* version 1.3.1. Subsequently, MCA was performed using the *FactoMineR* package version 2.11 in R, while adjusted percentages of inertia were obtained using the *ca* package version 0.71.1. MCA offers several advantages for the present study: (i) it is well suited to datasets with multiple categorical variables and rich annotation schemes; (ii) it reduces complex multi-way associations to a small number of interpretable dimensions; and (iii) it provides powerful visualizations that map relations of proximity and distance onto configuration plots, making clusters, oppositions and usage tendencies readily observable (Brezina, 2018; Xu, 2023). For these reasons, MCA has been widely adopted in diachronic linguistic research (e.g., Liu, 2023; Zehentner, 2021). It is well suited to exploratory contexts in which data imbalance and the rarity of certain structures across periods are common, as it allows both the relative relations among syntactic structures and their usage preferences to be visualized within multidimensional spaces. Nevertheless, categories with very low frequencies were interpreted with caution, and post-O was excluded from the MCA because of its extreme rarity.

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

4. Results

4.1. Diachronic frequency distribution of the four object-involving qǐlái structures

Since frequent use contributes to cognitive entrenchment of linguistic structures (Bybee, 2010; Langacker, 1987), frequency and proportion provide an important basis for identifying prototypical syntactic patterns and competitive relations among the structures. To better illustrate these trends, Figure 1 visualizes the distribution of the four syntactic forms across centuries.

Mid-O was the earliest-attested form and maintained the highest frequency across all historical periods. Both post-O and ba-O were first attested in the 13th–14th-century transitional period. Although mid-O remained overwhelmingly dominant, its proportion shows a general downward tendency after the 14th century, with fluctuations and a noticeable rebound in the 20th century. This pattern suggests that mid-O increasingly coexisted with other emerging structures rather than being replaced by them. Pre-O reached its highest proportion in the 17th century, whereas ba-O became increasingly prominent from the 16th–19th centuries and reached its highest proportion in the 19th century before declining in the early 20th century. By contrast, post-O remained marginal throughout the present dataset. A chi-square test further confirmed a significant association between syntactic structure and century (χ^2 = 162.31, df = 27, p < 0.001), indicating substantial diachronic change in distribution. To further examine patterns of competition among the structures and their distinct diachronic trajectories with respect to usage features, Section 4.2 turns to the results of the MCA.

4.2. MCA visualization of the diachronic evolution of object-involving qǐlái structures

To further examine diachronic changes in the usage features of pre-O, ba-O and mid-O, MCA was conducted. Given their uneven distribution across centuries, the data were further consolidated into three broader stages: the 12th–16th centuries were retained as Stage I, the 17th and 18th centuries were collapsed into Stage II, and the 19th and early 20th centuries were collapsed into Stage III. This step was taken to

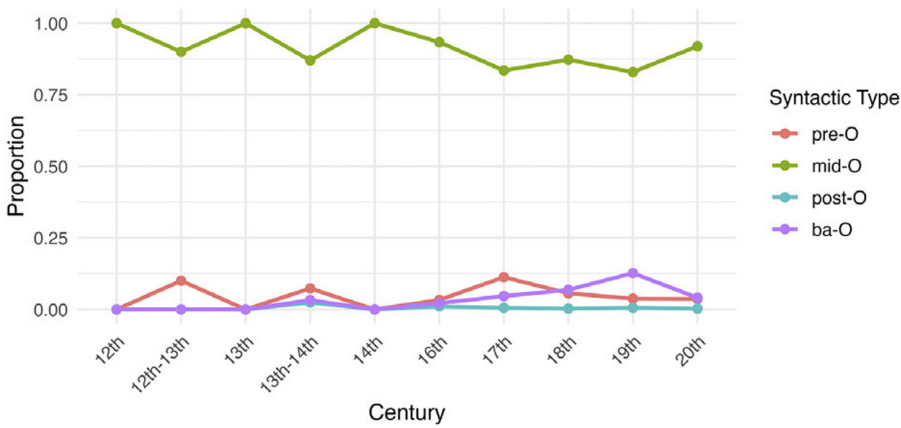


Figure 1. Proportion of the four syntactic forms across centuries.

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

reduce the influence of data imbalance and low-frequency categories on the MCA results (Glynn, 2014). A comparison with separate century-by-century analyses confirms that the merged-stage analysis captures diachronic trends in the usage features of the three structures more clearly.

4.2.1. Overall trend

Figure 2 presents the 95% confidence ellipses for the three structures across the historical stages, revealing a rough but clear evolutionary trend. The raw inertia values for the first two dimensions indicate moderate explanatory power across stages. However, since standard MCA tends to underestimate the proportion of explained variance, these raw values are not directly interpretable. We therefore report the adjusted percentages of inertia for the first five dimensions in each stage in Table 4. These adjusted values provide a more realistic estimate of the information captured by each dimension and thus allow for a more reliable evaluation of how well the leading dimensions represent the structure of the data (Glynn, 2014; Levshina, 2015; Zehentner, 2021). Across all stages, the first two dimensions consistently account for approximately 71.2%–74.7% of the total variance, indicating relatively strong explanatory power. Moreover, the scree plots show a clear drop after Dimension 2 in every stage. The following analysis therefore focuses on the first two dimensions, which capture the most substantial and stable associations among the variables (Desagulier, 2017).

Having established that the first two MCA dimensions adequately capture the structural variation in the data, we now turn to the broader diachronic patterns revealed by the analysis. Overall, the positions of all three structures change across the three stages, indicating substantial shifts in their usage profiles over time. In Stage I (see Figure 2a), pre-O and mid-O are clearly separated in the MCA space, whereas ba-O is too sparsely represented to allow a robust interpretation. In Stage II (see Figure 2b), as the frequencies of pre-O and ba-O increase, the spatial distribution of the three structures changes noticeably. Pre-O shifts leftward along Dimension 1, while remaining mainly on the positive side of Dimension 2. At the same time, the confidence ellipse of ba-O moves markedly closer to that of pre-O, and the substantial overlap between the two suggests a temporary convergence in their usage profiles. In Stage III (see Figure 2c), the three structures become more clearly differentiated. Mid-O is separated from ba-O, and to a lesser extent from pre-O, primarily along Dimension 1, whereas the contrast between pre-O and ba-O is more strongly reflected along Dimension 2. Notably, mid-O remains relatively close to the x-axis, suggesting that its distribution is less strongly shaped by the features captured by Dimension 2. Taken together, this evolving spatial configuration points to a further redistribution of usage features across the three structures and suggests increasing functional differentiation over time.

To further examine how the usage profiles of the three structures evolve, Section 4.2.2 provides a more detailed inspection of the MCA biplots, with particular attention to diachronic variation in their usage patterns. The contributions of individual variables to Dimensions 1 and 2 across different centuries are summarized in Supplementary Appendix C.

4.2.2. Stage I: Initial differentiation

During the 12th–16th centuries, mid-O is the dominant structure, whereas the other two structures are still emerging. This is especially the case for ba-O, which

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

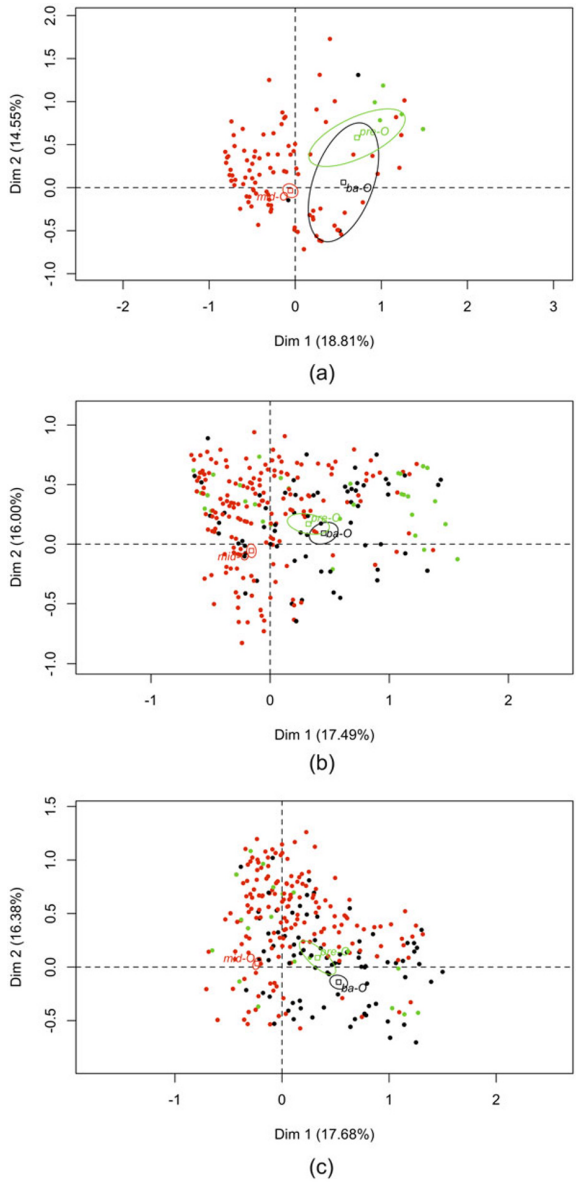


Figure 2. Confidence ellipses for pre-O, mid-O and ba-O across three historical stages. Individual instances are color-coded by syntactic type: pre-O in green, mid-O in red and ba-O in black. Panels (a), (b) and (c) correspond to Stage I, Stage II and Stage III, respectively.

remains relatively infrequent at this stage. The MCA biplot for Stage I is presented in Figure 3.

As shown in Figure 3, mid-O is located very close to the center of the MCA map. This central position suggests that, on the first two dimensions, mid-O is not strongly

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

Table 4. Principal inertias (eigenvalues) of the first five MCA dimensions across stages

Stage I: 12th–16th				
Dim	Value	% Adjusted	% Cumulative adjusted	Scree plot adjusted
1	0.070	48.8	48.8	**********
2	0.032	22.5	71.2	*******
3	0.011	7.9	79.1	**
4	0.002	1.3	80.4	
5	0.000	0.3	80.7	
Stage II: 17th–18th				
Dim	Value	% Adjusted	% Cumulative adjusted	Scree plot adjusted
1	0.057	42.2	42.2	**********
2	0.044	32.4	74.6	**********
3	0.007	5.0	79.6	**
4	0.001	0.6	80.2	
5	0.000	0.2	80.4	
Stage III: 19th–20th				
Dim	Value	% Adjusted	% Cumulative adjusted	Scree plot adjusted
1	0.059	41.5	41.5	**********
2	0.047	33.1	74.7	**********
3	0.007	4.6	79.3	*
4	0.001	0.8	80.1	
5	0.001	0.5	80.5	

associated with a specific cluster of usage features but instead covers a relatively broad range of contexts. This distributional pattern is consistent with its high frequency in Stage I and suggests that mid-O had not yet developed a highly specialized usage profile at this stage. By contrast, although pre-O is considerably less frequent than mid-O, it shows a clear association with object-related features. It is positioned close to the categories *patient* and *animate* and, more broadly, to *Obpersonpronoun_yes* in the upper-right region of the plot. This distribution indicates that pre-O is preferentially associated with contexts involving animate patient objects, especially pronominal objects. Example (2) illustrates this pattern: the object *wǒ* 'I' is a first-person pronoun referring to a human patient who is physically assisted into an upright position.

(2) 你 且 扶 我 起来, 取出 我的 纸笔墨…
Nǐ qiě fú wǒ qǐlái, qǔ-chū wǒ-de zhǐ-bǐ-mò…
2SG first hold 1SG QILAI take-out 1SG-GEN paper-brush-ink
'First help me up, and take out my paper, brush, and ink'.
(*Xi You Ji*, 12th–16th century)

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

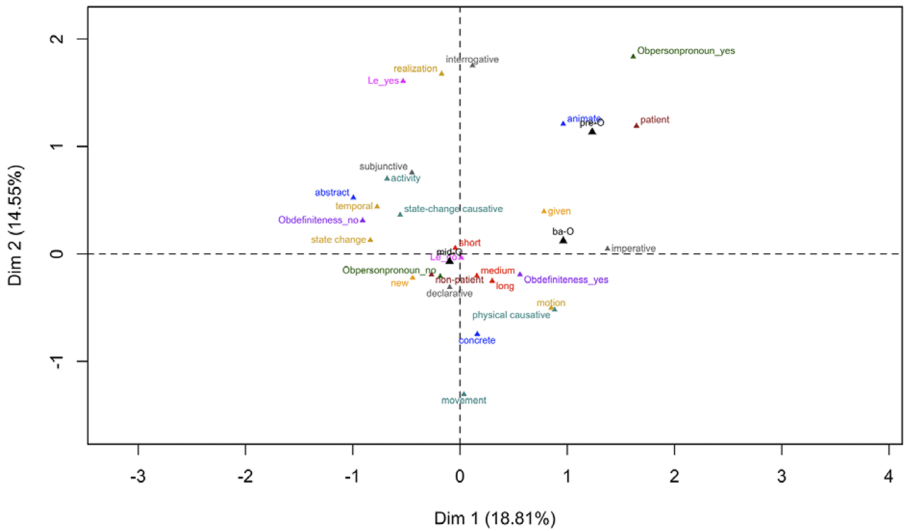


Figure 3. MCA biplot of Dimensions 1 and 2 for Stage I.
Note: In Figures 3 to 5, the categories of the supplementary variable *SyntacticType* (pre-O, mid-O and ba-O) are represented by black triangles in the MCA biplots, while the categories of the explanatory variables are distinguished by different colors to enhance readability.

4.2.3. Stage II: Functional overlap and initial specialization

During Stage II, both pre-O and ba-O increase in frequency and become more fully represented in the MCA space. As a result, the usage profiles of the three structures show a clearer reorganization than in Stage I. The MCA biplot for this stage is presented in Figure 4.

Compared with Figure 3, the distribution of categories in Figure 4 changes in several important respects. Dimension 1 is still structured primarily by object-related

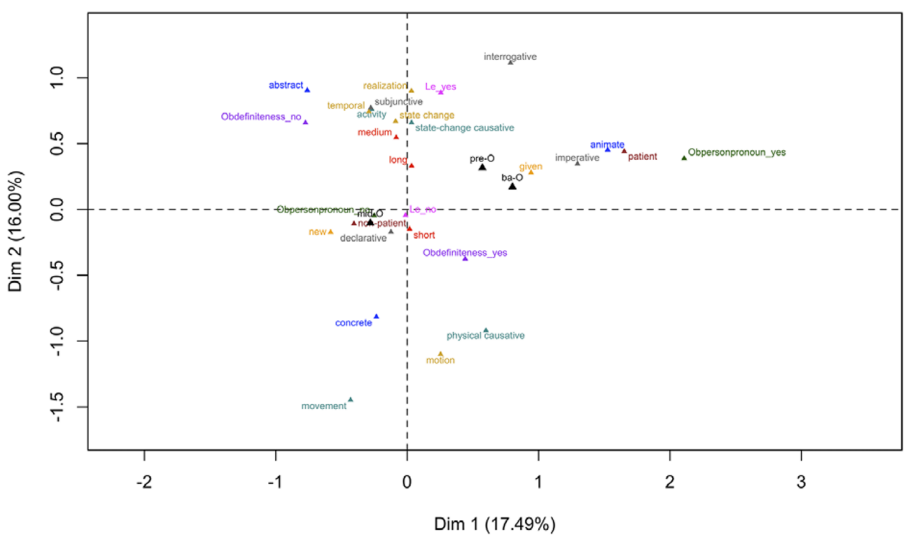


Figure 4. MCA biplot of Dimensions 1 and 2 for Stage II.

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

variables. Its positive side is associated with features such as *given*, *patient*, *animate* and *Obpersonpronoun_yes*, whereas its negative side is associated with *Obdefiniteness_no*, *new*, *abstract* and *non-patient*. This dimension now separates pre-O and ba-O more clearly from mid-O, suggesting that the former two structures are more closely associated with cognitively accessible and affected object profiles. More specifically, pre-O and ba-O move noticeably closer to each other in the upper-right region of the plot, where both are associated most directly with *given* objects. This pattern is illustrated in (3), where the pronominal object *tā* 'him' and the definite nominal object *fùrén* 'the woman' both refer to discourse-accessible participants.

(3) a. 秦王 自 下 阶 来 搀 他 起来…
Qínwáng zì xià jiē lái chān tā qǐlái…
King.Qin from descend step come support 3SG QILAI
'The King of Qin came down the steps and helped him up'.
(*Sui Tang Yan Yi*, 17th century)

b. 用 手 把 妇人 拉将起来.
Yòng shǒu bǎ fùrén lā-jiāng-qǐlái.
use hand BA woman pull-JIANG-QILAI
'Using his hands, (he) pulled the woman up'.
(*Jin Ping Mei*, 17th century)

Meanwhile, *patient*, *animate* and *Obpersonpronoun_yes* continue to make substantial contributions to the positive side of Dimension 1. These features are also relevant to the two examples in (3), both of which involve causative spatial motion affecting highly accessible objects. Importantly, however, these object-related features are now positioned somewhat closer to ba-O than in Stage I. Although this association should not be overstated, the shift suggests that, as ba-O becomes more frequent, it begins to share part of the object-related profile that was previously more strongly associated with pre-O. This creates a greater degree of functional overlap between the two structures.

Dimension 2, by contrast, is better interpreted as distinguishing between more concrete, spatially grounded events and more abstract or aspectually construed event types. The negative pole is defined mainly by features such as *movement*, *motion*, *physical causative* and *concrete*, whereas the positive pole is associated with *state change*, *realization*, *temporal*, *activity* and *state-change causative*. In this sense, Dimension 2 captures a contrast between events grounded in visible spatial motion and those involving state change, realization or temporal contouring.

Against this background, pre-O shows a shift relative to its position in Stage I. It is no longer concentrated as strongly around the highly accessible object features in the upper-right region but instead lies closer to upper-central categories such as *state change* and *state-change causative*. This does not indicate a one-to-one association between pre-O and any single event-type feature. Rather, it suggests that pre-O retains part of its earlier object-related profile while beginning to overlap more clearly with features related to event type and verbal semantics. This pattern can be seen in (4a), where the pre-O structure *suǒ wǒ qǐlái* 'lock me up' involves a pronominal human patient, thus preserving the object-related profile associated with pre-O. At the same time, the verb *suǒ* 'lock' profiles a causative change of state: the affected

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

participant is brought into a constrained or enclosed condition. In other cases, pre-O is used to mark a change in the state of the event participant. For instance, in (4b), the object *nǎo* 'anger' is an abstract emotion initiated by the verbal action and the whole expression denotes the onset of an angry state. These examples suggest that, during Stage II, pre-O retains its earlier association with affected patient objects while also extending toward more abstract state-change uses.

(4) a. 本官 焉 敢 锁 我 起来?
Běnguān yān gǎn suǒ wǒ qǐlái?
this.official how dare lock 1SG QILAI
'How dare this official lock me up?'
(*Shuo Tang Quan Zhuan*, 18th century)

b. 佛印 就 发 恼 起来.
Fó.Yìn jiù fā nǎo qǐlái.
PN then become angry QILAI
'Fo Yin then became angry'.
(*Yu Shi Ming Yan*, 17th century)

The proximity of pre-O to ba-O also indicates that the two structures now share a larger portion of the MCA feature space than before. A comparable profile is found in ba-O, as exemplified in (5), where *bǎ zhè nǚzǐ bǎngfù qǐlái* 'tie this woman up' involves a definite, animate and highly affected object. The verb *bǎngfù* 'bind' encodes a causative change of state, in which the woman is brought into a physically constrained condition. Thus, the closer positioning of pre-O and ba-O in the MCA space reflects not only overlap in object-related features but also overlap in event types.

(5) 把 这 女子 绑缚起来.
Bǎ zhè nǚzǐ bǎngfù-qǐlái.
BA this woman bind-QILAI
'(He) tied this woman up'.
(*Xing Shi Heng Yan*, 17th century)

Finally, mid-O remains closest to the origin, suggesting that it continues to be the least specialized structure at this stage. At the same time, its shift toward the negative side of Dimension 1 is more visible than in Stage I, indicating that this dimension now contributes more clearly to its differentiation from pre-O and ba-O. Compared with the other two structures, mid-O is located closer to the central-to-left cluster defined by *non-patient*, *new* and *declarative*, and, more weakly, by *Obdefiniteness_no*, rather than by the cognitively accessible and affected object profile associated with pre-O and ba-O. The instances in (6) further illustrate the broader distributional profile of mid-O. The objects *bìng* 'illness', *yī-gè dà-báihè* 'a large white crane', *yuè* 'music' and *yì-jiān cháipéng* 'a firewood shed' are all discourse new and non-patient in declarative clauses, yet they occur in different event types. This suggests that mid-O does not show a strong preference for any particular event type on Dimension 2, further supporting its relatively unspecialized and prototypical status across event types.

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

(6) a.

受了	暑气,	在	山东	台儿庄	生	起	病
Shòu-le	*shǔqì,*	*zài*	*Shāndōng*	*Tái'érzhuāng*	*shēng*	*qǐ*	*bìng*
suffer-PFV	summer.heat	at	Shandong	Tai'erzhuang	grow	QI	illness

来.
lái.
LAI

'Having suffered from summer heat, [he] fell ill in Tai'erzhuang, Shandong'.
(Event of state change, *Ye Sou Pu Yan*, 18th century)

b.

只	听	那	黑影	里	嘎然	一	声,
Zhǐ	*tīng*	*nà*	*hēiyǐng*	*lǐ*	*gārán*	*yī*	*shēng,*
only	hear	that	shadow	in	sudden. crack	one	sound

却	飞	起	一个	大白鹤	来.
què	*fēi*	*qǐ*	*yī-gè*	*dà-báihè*	*lái.*
but	fly	QI	one-CLF	big-crane	LAI

'[One] only heard a sudden cracking sound in the shadow, and then a large white crane flew up'.
(Motion event, *Hong Lou Meng*, 18th century)

c.

不多时,	军中	奏	起	乐	来.
Bùduōshí,	*jūnzhōng*	*zòu*	*qǐ*	*yuè*	*lái.*
before. long	army. inside	play	QI	music	LAI

'Before long, music began to be played in the army camp'.
(Event of temporal contouring, *Lyu Ye Xian Zong*, 18th century)

d.

在	大树	之间,	搭	起	一间	柴棚	来.
Zài	*dàshù*	*zhījiān,*	*dā*	*qǐ*	*yī-jiān*	*cháipéng*	*lái.*
at	big.tree	between	build	QI	one-CLF	firewood.shed	LAI

'[They] built a firewood shed among the big trees'.
(Event of realization, *Chu Ke Pai An Jing Qi*, 17th century)

Overall, Stage II appears to represent a transitional phase in the historical development of the three structures. With the increase in frequency of ba-O, its distributional profile becomes more salient in the MCA space and begins to overlap with the profile previously more strongly associated with pre-O. At the same time, pre-O no longer remains confined to its earlier object-related profile but extends toward other parts of the feature space. These developments suggest increasing competition between pre-O and ba-O, while also laying the groundwork for the more explicit functional differentiation among pre-O, mid-O and ba-O in the subsequent stage.

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

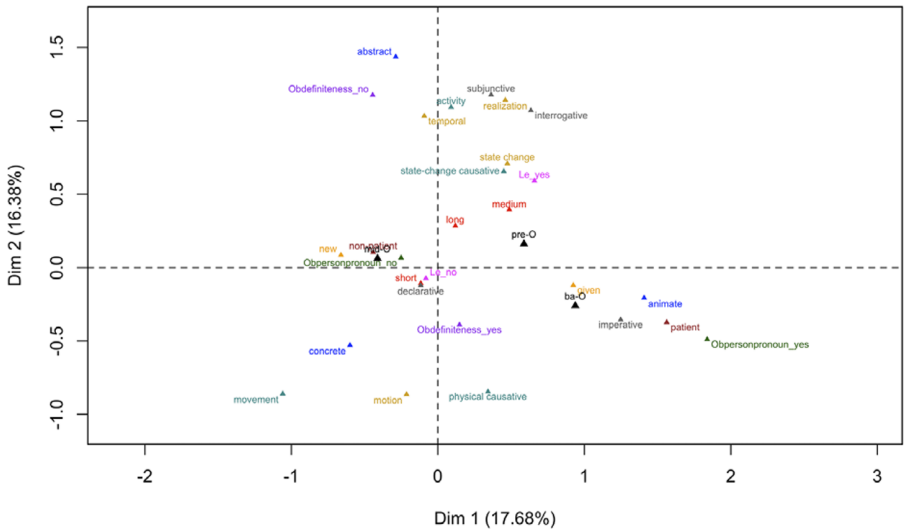


Figure 5. MCA biplot of Dimensions 1 and 2 for Stage III.

4.2.4. Stage III: Clear specialization with continuing overlap

Finally, during Stage III, the three structures become more clearly separated in the MCA space, and their distributional profiles show the strongest differentiation observed across the three stages. Compared with Stage II, where pre-O and ba-O still displayed substantial overlap, Stage III is characterized by a more distinct redistribution of usage features across the structures. The MCA biplot for this stage is presented in Figure 5.

Generally, mid-O remains on the left and relatively close to the origin, pre-O is located in the upper-right quadrant, and ba-O is located in the lower-right quadrant. This configuration suggests that the three structures reach their highest degree of functional differentiation in this stage. However, this differentiation should be understood as gradient rather than absolute, as their distributions still partially overlap.

As in the preceding stage, Dimension 1 is structured mainly by object-related variables, especially information status, affectedness, animacy and pronominality. From left to right, object profiles become increasingly cognitively accessible and strongly affected: the negative side is associated more with *new* and *non-patient* objects, while the positive side is associated more with *given*, *patient*, *animate* and *Obpersonpronoun_yes*. Dimension 2, by contrast, differentiates the data points more clearly in terms of event types. Its positive side is associated with relatively more abstract, temporal or aspectually construed event types, including *abstract*, *temporal*, *activity*, *realization*, *state change* and *Le_yes*. Its negative side is associated with more concrete and spatially grounded event types, such as *movement*, *motion*, *physical causative* and *concrete*.

Against this background, ba-O occupies the lower-right quadrant and is most closely associated with *given*, *patient* and *animate*, while also remaining relatively close to *imperative*. This distribution indicates that ba-O is the structure most strongly associated with cognitively accessible and strongly affected objects. For example, in (7a), the object *tā* 'him' is a pronominal, discourse-given and animate patient who is physically assisted into an upright position, and in (7b), the object *tāmen sānrén* 'the three of them' is likewise definite, animate and affected, and the

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

clause occurs in an imperative context, where the speaker orders the addressee to keep the three participants in custody. These examples show that ba-O foregrounds the affectedness of a salient object, often under a causative or directive construal.

(7) a. 太夫人 把 他 搀了起来.
Tài-fūrén bǎ tā chān-le-qǐlái.
Madam.Dowager BA 3SG support-PFV-QILAI.
'The Madam Dowager helped him up'.
(*Jiu Wei Gui*, 20th century)

b. 老爷 吩咐: '暂且 把 他们 三 人 看押起来'.
Lǎoyé fēnfù: 'zànqiě bǎ tāmen sān rén kànyā-qǐlái'.
Master order temporarily BA 3PL three person detain-QILAI
'The Master ordered, "Detain the three of them for the time being".'
(*Peng Gong An*, 19th century)

In addition, the position of ba-O in the lower half of the plot suggests that, relative to pre-O, it is less closely associated with abstract and temporally construed event types. This does not mean that ba-O preferentially encodes motion events as such, since it still remains some distance from motion and physical causative. Rather, the contrast is one of relative degree: compared with pre-O, ba-O tends to be associated with a more concrete event space in which a definite and accessible object undergoes physical manipulation, control or confinement. In this sense, ba-O is distinguished not simply by the type of event it encodes, but by the way it construes the object as a salient affected participant within a causative or disposal-like event frame.

By contrast, pre-O occupies the upper-right quadrant on its own. Relative to Stage II, it is now more clearly separated from ba-O along Dimension 2, suggesting that the two no longer share the same distributional profile to the same extent. The position of pre-O indicates a moderate association with the positive side of Dimension 1, but its distinctiveness in this stage lies more clearly in its association with the upper-half feature space, especially relatively close to categories such as *state change*, *state-change causative* and *Le_yes*. These associations are not especially strong individually; taken together, however, they suggest that pre-O is now relatively more inclined toward events of state change or, more broadly, toward aspectually or temporally profiled events. In this sense, what becomes more salient in Stage III is not simply the persistence of its earlier object-related profile, but the emergence of a more distinctive role in event representation. For instance, in (8), *yōudài tā qǐlái* 'begin to treat him favorably' foregrounds the beginning of a new course of action or attitude toward the pronominal object *tā* 'him'.

(8) 因 贪吃 蟠桃 的 缘故,
Yīn tānchī pántáo de yuángù,
because greedily.eat flat.peaches DE reason
反 格外 优待 他 起来.
fǎn géwài yōudài tā qǐlái.
instead especially treat.well 3SG QILAI
'Because of his greed for the peaches of immortality, [they] instead began to treat him especially well'.
(*Ba Xian De Dao*, 19th century)

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

Finally, mid-O remains on the left side of the plot and closest to the origin. Its position indicates a closer relation to features such as *new*, *non-patient* and *Obpersonpronoun_no*, which are still less affected and non-topical. At the same time, it does not cluster strongly with any single group of event-type features on Dimension 2. This suggests that mid-O continues to be the least specialized of the three structures, with a broader and more neutral distribution across different kinds of scenes. Its relative centrality therefore remains consistent with its prototypical status and its wider range of use.

Overall, the MCA results for Stage III suggest that the three structures display their most pronounced functional specialization in this period. However, such specialization does not imply sharply bounded contexts of use, and they continue to overlap to some extent. Therefore, the differentiation is best understood not as categorical division but as a gradual redistribution of usage preferences across object accessibility, affectedness and event representation.

4.2.5. *Qualitative analysis of post-O*

Due to its extreme rarity, post-O was excluded from the MCA modeling. Instead, a qualitative analysis was conducted to examine its distinctive usage features. Unlike the other three structures, post-O shows a strong preference for activity verbs such as *xiǎng* 'think', *kàn* 'look' and *shuō* 'speak'. These verbs typically introduce a subsequent act of evaluation, judgment or description, with the postverbal object functioning as the entity to be evaluated or further elaborated. For example, in (9a), *kàn-qǐlái zhège fùrén* 'looking at this woman' introduces *zhège fùrén* 'this woman' as the entity to be evaluated, which is then followed by the speaker's judgment, *shì gè bùliáng de* '[she] is not a decent person'. Similarly, in (9b), *xiǎng-qǐlái nà yì tiān* 'thinking back to that day' does not simply denote a cognitive activity. Rather, it establishes *nà yì tiān* 'that day' as a discourse anchor, which provides the reference point for the following evaluative comment.

(9) a. 看起来 这个 妇人, 是 个 不良 的.
Kàn-qǐlái zhège fùrén, shì gè bùliáng de.
look-QILAI this woman be CLF immoral ADJ
'Looking at this woman, [one can tell that] she is not a decent person'.
(*Quan Yuan Qu*, 12th–16th century)

b. 想起来 那 一 天, 因 行令 戒酒,
Xiǎng-qǐlái nà yì tiān, yīn xínglìng jièjiǔ,
think-QILAI that one day because drinking-game abstain.from.alcohol
于 立志 虽 佳, 于 戒酒 稍 欠.
yú lìzhì suī jiā, yú jièjiǔ shāo qiàn.
regarding resolve although good regarding abstaining.from.alcohol somewhat lacking
'Thinking back to that day, when the drinking game led to a pledge of abstinence, the intention was admirable, though the actual abstinence was less than complete'.
(*Xia Nü Qi Yuan*, 19th century)

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

In other cases, post-O can also encode a basic motion event, as in (10). The object *yí kuài xiǎozhuān* 'a small brick' is an indefinite, new, concrete entity, and the structure describes its upward motion.

(10) 由 地下 飞起来 一 块 小砖.
Yóu dìxià fēi-qǐlái yí kuài xiǎozhuān.
from ground fly-QILAI one CLF small.brick
'A small brick flew up from the ground'.
(*Ji Gong Quan Zhuan*, 18th century)

Overall, post-O remains rare throughout the period under investigation, and its range of use is therefore relatively limited. For this reason, its distributional profile should be interpreted with caution. In the following section, we further examine how differences in surface syntactic structure give rise to these distributional patterns and how they help explain the historical competition and functional differentiation among the object-involving *qǐlái* structures from a cognitive perspective.

5. Discussion

The analyses presented above reveal a clear diachronic reorganization in the distribution and usage profiles of the four object-involving V *qǐlái* structures. Overall, the frequency distributions of these structures changed significantly across centuries. The decline of pre-O, the marked rise of ba-O, the persistent rarity of post-O and the consistently high frequency of mid-O are broadly consistent with previous observations on object-involving CDC structures (Hao, 2018; Hiroshi, 2021; Wang, 2005; Zhang, 1991b; Zhang, 2018). At the same time, the present data also reveal an important complement-specific pattern. Contrary to the claim that pre-O represents the earliest object-involving form in CDC constructions more generally (Wang, 2005; Zhang, 1991b), the present data suggest that, in the case of *qǐlái*, mid-O is attested earlier and remains the most frequent pattern throughout the period under investigation. This finding supports the view that mid-O constitutes the prototypical configuration of object-involving *qǐlái* structures (Hiroshi, 2021; Jia, 1998). The diachronic development of object-involving *qǐlái* structures therefore both conforms to broader tendencies in CDC constructions and follows a complement-specific trajectory of its own.

The MCA results further show that the usage profiles of pre-O, mid-O and ba-O changed substantially across the three diachronic stages. This development points not to a simple process of replacement but to a more complex process of competition involving retention, redistribution and increasing functional differentiation. The three structures continued to coexist, and their usage profiles became increasingly differentiated over time. This pattern is consistent with recent accounts of language change showing that competing variants may be retained over long periods through partial functional differentiation, stable overlap or the formation of probabilistic functional niches (Fonteyn, 2017, 2019; Fonteyn & Nini, 2020; De Smet, 2008; Torres Cacoullos & Walker, 2009; Zehentner, 2021). More specifically, the MCA results suggest that this redistribution was shaped by several interrelated factors, especially object accessibility, object affectedness and event representation. In what follows, this

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

diachronic redistribution is explained from a cognitive perspective, focusing on language processing, conceptual accessibility and diagrammatic iconicity.

In Stage I, the main contrast was between mid-O and pre-O. Previous psycholinguistic and usage-based studies have shown that conceptually accessible elements tend to occur earlier in the clause. Such accessibility may derive from discourse status, such as givenness and definiteness, or from inherent semantic and referential properties, such as animacy and pronominality (Bock & Warren, 1985; Branigan et al., 2008; Bresnan et al., 2007; MacDonald, 2013; V. Nice & Dietrich, 2003; Vogels & Van Bergen, 2017). This helps explain why pre-O was associated with animate objects and personal pronouns in the early stage. Compared with mid-O, the [V–O–C] order of pre-O places the object immediately after the verb and before the complement, thereby giving it a more prominent position. This ordering effect can also be interpreted through the iconic principle of proximity. According to diagrammatic iconicity, elements that are conceptually or functionally close tend to be placed close to each other in linguistic form (Bolinger, 1977; Givón, 1991; Haiman, 1985; Radden, 2021). The verb and its object constitute a particularly clear case of such conceptual proximity, since the object is the participant most directly acted upon by the verbal event. In pre-O, the adjacency of V and O strengthens the conceptual link between the action and the affected participant. This explains why pre-O was strongly associated with patient objects: it is well suited to conceptualizing the object as directly affected by the event.

In Stage II, ba-O increased in frequency and began to display a usage profile similar to that of pre-O. It preferred discourse-given objects and was also weakly associated with animate and patient objects, a tendency broadly consistent with previous descriptions of object-involving *qǐlái* structures (Hao, 2018; Zhang, 1991a, 1991b; Zhu, 1982). This convergence is also structurally motivated. The *ba* construction places the object before the verb and marks it overtly, thereby making the object syntactically prominent and construing it as semantically affected. Previous studies of the *ba* construction have repeatedly emphasized its association with disposal, affectedness and causation (Wang, Wang, 2014; Li & Thompson, 1981; Jing-Schmidt & Tao, 2009; Peyraube & Wiebusch, 2020; Liu & Cuyckens, 2025). From this perspective, the similarity between pre-O and ba-O in Stage II is not accidental: both structures can foreground cognitively accessible and highly affected objects. However, they achieve this foregrounding through different formal means. Ba-O relies on an explicitly marked object-preposing strategy, whereas pre-O relies on the adjacency between V and O.

The growing frequency of ba-O also helps explain the subsequent reorganization of pre-O. Since ba-O is structurally specialized for affected and cognitively accessible objects, it may have gradually encroached on part of the functional domain previously associated with pre-O. Pre-O, however, was not simply replaced. Instead, it developed a more specific association with events of state change. This development can again be linked to its surface configuration. Because V and O remain adjacent in pre-O, the verb and the object are more tightly integrated into a conceptual unit. When the verb encodes state-change causation, the structure is therefore particularly suitable for presenting the emergence of a new state in which the object is directly involved. In this respect, pre-O does not simply compete with ba-O for affected objects; rather, it comes to favor a more specific representation that highlights the emergence of a change of state involving the object (see [4]). By contrast, in mid-O, V and O are separated by *qǐ*, yielding a constituent structure of [V–*qǐ*] [O–*lái*]. Since *qǐ*

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

can express an inchoative meaning, the former unit can encode event initiation, while [O–*lái*] marks the progression of this initiated event. It is this constituent organization that makes mid-O more compatible with processual or aspectual construals that foreground the unfolding of the event. In addition, the weaker V–O adjacency also makes mid-O less suited to highlighting direct object affectedness, and its object tends to be less patient-like. Moreover, since the object occurs later in the clause, mid-O also tends to occur with indefinite, new, inanimate and non-pronominal objects, which are typically less cognitively accessible.

By Stage III, the three structures had developed more clearly differentiated profiles under the constraints of their respective formal configurations. Ba-O became strongly associated with cognitively accessible and highly affected objects. Its marked [BA–O–V–C] configuration made it especially appropriate when the speaker wished to foreground a definite, given, animate or otherwise salient object before presenting the action that affected it. Pre-O, benefiting from the close proximity between V and O, remained associated more strongly with events of state change. This profile is particularly compatible with relatively short objects, since longer objects are more likely to disrupt the prosodic and processing integration among V, O and the following complement. At the same time, the frequency and proportion of pre-O declined in this stage, and its usage profile came to occupy an intermediate position between mid-O and ba-O. This suggests that pre-O was increasingly constrained by both structural and usage-based pressures: it retained a specialized function, but its distributional range became more limited.

Mid-O, by contrast, retained a broader and more processual profile. Its surface configuration, organized as [V–*qǐ*] [O–*lái*], makes it suitable for events represented in terms of temporal unfolding, continuation or aspectual development. Importantly, its stability does not mean that its usage profile remained unchanged. As pre-O and ba-O developed stronger associations with accessible and affected objects, mid-O was gradually redefined by contrast within the system of competing structures. It came to favor less accessible and non-patient objects, while still covering a wide range of event types. This suggests that even the most basic or unmarked variant may adjust its distributional profile in response to structural competition. The prototypicality of mid-O may also be related to broader prosodic constraints in Chinese. Dong (1998) argues that object-involving verb–complement constructions are strongly constrained by Chinese prosodic organization: the verb–complement sequence tends to form a disyllabic prosodic word. In this light, mid-O can be understood as a prosodically balanced configuration, since it allows the structure to be parsed into two relatively compact units, namely [V–*qǐ*] [O–*lái*] (Feng, 1996; Song, 2019). This prosodic compatibility, together with general production preferences, may have facilitated its repeated selection in use (MacDonald, 2013). However, its neutrality and unmarkedness do not mean that it lacks a function of its own. Rather, mid-O provides a relatively neutral, processual and less object-prominent way of representing an event. Precisely because mid-O offers this default representation, the choice of pre-O or ba-O becomes more meaningful, as these forms highlight more specific features such as object accessibility and affectedness. In this sense, competition among syntactic variants is not only a matter of local functional redistribution but also reflects their differential alignment with broader prosodic and production-based tendencies in the language.

The post-O structure requires separate discussion because of its persistent rarity and distinct developmental pathway. Previous studies have proposed that post-O tends to disfavor definite objects, introduce new information and require numeral

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

modification of the object (Lu, 2002; Zhang, 1991b; Zhu, 1982). The present data, however, show that most attested instances of post-O involve definite objects and that numeral modification is not obligatory (see [9a–b]). This discrepancy may reflect different stages in the diachronic development of post-O. In the earlier stage, post-O had not yet developed a fully specialized function: approximately half of the tokens encoded basic motion events (see [10]), while the remaining cases showed an emerging evaluative use. At this stage, evaluative meanings were still largely anchored to definite or discourse-given referents, which provided cognitively accessible reference points before further evaluative information was introduced. As illustrated in (9a–b), *zhège fùrén* 'this woman' and *nà yī tiān* 'that day' are both definite and discourse-old; they first establish the entity or situation to be evaluated, and the following discourse then supplies new evaluative or descriptive information about them. This suggests that, in modern Chinese, with the further grammaticalization of *qǐlái* toward more subjective meanings (Qi & Zeng, 2009; Traugott, 1995), and with the tendency of the [V–C–O] order to accommodate cognitively less accessible information, such as new, indefinite or inanimate objects (Hao, 2018; Li, 2017; Zhang & Fang, 1996), post-O may have become less dependent on such anchoring points and gradually extended its evaluative function to newly introduced referents. From this perspective, post-O can be seen as a transitional structure that helped pave the way for the subsequent development of V *qǐlái* into a discourse marker (Jiang, 2020; Wu, 2012), as exemplified by (11).

(11)

现在	说起来,	人们	很	难	相信	
Xiànzài	*shuō-qǐlái,*	*rénmen*	*hěn*	*nán*	*xiāngxìn*	
now	say-QILAI	people	very	hard	believe	
这	曾	是	一个	白手起家	的	乡镇
zhè	*céng*	*shì*	*yí-gè*	*báishǒu-qǐjiā*	*de*	*xiāngzhèn*
this	once	be	one-CLF	self.start	ATTR	township

企业。
qǐyè.
enterprise

'Now, speaking of it, people find it hard to believe that this was once a township enterprise built up from nothing'.
(News report from *Xinhua News Agency*, 2010)

Overall, the diachronic development of the object-involving *qǐlái* structures shows that competition among related syntactic structures does not necessarily result in the replacement of one form by another. Rather, pre-O, mid-O and ba-O continued to coexist while gradually developing different distributional preferences and functional profiles. This pattern is compatible with usage-based accounts of grammatical change, which emphasize gradualness, variation, frequency effects and the cumulative impact of repeated language use (Bybee, 2010; Traugott & Trousdale, 2013). It is also consistent with studies of variant competition showing that formally related variants may be retained through functional redistribution rather than categorical substitution (Fonteyn, 2017, 2019; Fonteyn & Nini, 2020; De Smet, 2008; Torres Cacoullos & Walker, 2009). The choice between pre-O, mid-O and ba-O should therefore not be treated as a mechanically determined alternation. Even when these structures can represent comparable event scenes, they impose different ways of representing the relation between the event and the object, especially with respect to object accessibility, affectedness and event unfolding, which corresponds to the

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

cognitive operation of construal (Radden & Dirven, 2007). This interpretation is also consistent with MacDonald's (2013) production-oriented account of word order choice, in which speakers select forms that help manage conceptual accessibility, production efficiency and communicative clarity. Overall, the case of object-involving *qǐlái* structures contributes to a broader understanding of language change by showing that structural competition is not merely a matter of changing frequencies or formal redistribution. More fundamentally, it involves the gradual reorganization of how speakers recruit different syntactic patterns to encode event structure, object prominence and discourse accessibility. In this sense, the present study provides a concrete case of how syntactic change emerges from the interaction of multiple factors, including surface structural constraints, processing-based salience and broader language-specific preferences.

6. Conclusions

This study examined the diachronic evolution of four object-involving *qǐlái* structures from the 12th to early 20th centuries using large-scale corpus data and multifactorial quantitative analysis. The results reveal, first, substantial diachronic changes in their frequency distributions: mid-O was the earliest-attested and consistently most frequent structure, pre-O peaked in the 17th century and subsequently declined, ba-O increased in prominence and reached its highest proportion in the 19th century, and post-O remained marginal throughout. These findings help revise earlier claims about the order of emergence and support the prototypical status of mid-O, underscoring the importance of analyzing object-involving structures at the level of individual CDCs. Beyond frequency, MCA visualizations reveal the diachronic competition and functional differentiation of pre-O, mid-O and ba-O in terms of their usage preferences. These patterns were further interpreted from a cognitive perspective, with particular reference to language processing, conceptual accessibility and diagrammatic iconicity. Over time, ba-O became increasingly associated with cognitively accessible and highly affected objects, pre-O retained a more specialized link with state-change events due to V–O adjacency, and mid-O remained the most frequent and relatively neutral pattern, being more closely associated with less accessible or non-patient objects and foregrounding event unfolding. The persistent rarity of post-O suggests a distinct developmental pathway, in which its evaluative use may have paved the way for the later discourse-marker function of V *qǐlái*. Overall, the findings show that the diachronic development of the object-involving *qǐlái* structures was not a process of simple replacement, but one of long-term coexistence, competition and functional redistribution. More broadly, syntactic competition in language change involves not only frequency shifts but also the gradual reorganization of form-function mappings through which speakers encode object prominence, event structure and discourse accessibility.

Taken together, this study contributes to research on Mandarin object-involving CDC structures and syntactic change more broadly. Theoretically, it demonstrates that competition among formally related structures may lead to gradual functional differentiation rather than simple replacement, a process shaped by multiple interacting factors, including surface configuration, processing-related cognitive tendencies and prosodic preferences in the broader language system. It also underscores the need to examine object-involving structures at the level of individual CDCs rather than treating CDCs as a homogeneous category. Methodologically, by integrating large-

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

scale diachronic corpus data with multifactorial quantitative analysis, the study provides a replicable framework for capturing subtle usage preferences and diachronic reorganization in syntactic systems characterized by imbalance and variation.

Several limitations should nonetheless be noted. Due to the extreme rarity of post-O, its analysis necessarily relied primarily on qualitative evidence, which limits direct comparison with the three quantitatively modeled structures. Moreover, although the historical corpus data provide extensive diachronic coverage, they do not capture recent developments or spoken uses of *qǐlái*, where additional functional extensions may have emerged. Future research should therefore incorporate larger and more diverse datasets, including modern and spoken corpora, to assess the robustness of the usage patterns identified here. Extending the present multifactorial and exploratory approach to other CCs, as well as to cross-linguistic data, would further clarify whether the observed pathways of competition and functional differentiation reflect CDC-specific tendencies or more general cognitive operations underlying syntactic change.

Data availability statement. The datasets generated and/or analyzed during the current study, together with the supplementary material, are available in the OSF repository at https://osf.io/qwuv2/overview?view_only=0c9c33a51b7a4f238a4226739304a3d1.

Acknowledgments. We would like to express our sincere gratitude to the anonymous reviewers for their valuable comments and constructive suggestions. Their insightful feedback has significantly improved the quality, clarity and overall presentation of this article.

Competing interests. The authors declare no conflict of interest.

Declaration of AI-assisted language editing. The authors used ChatGPT, developed by OpenAI and available at chatgpt.com, to assist with English-language editing, including improving grammatical accuracy, sentence fluency and stylistic clarity in selected parts of the manuscript. All AI-assisted suggestions were checked, revised and approved by the authors, who take full responsibility for the final content.

References

Arnold, J. E., Losongco, A., Wasow, T., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, 76(1), 28–55. https://doi.org/10.2307/417392

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21(1), 47–67. https://doi.org/10.1016/0010-0277(85)90023-X

Bolinger, D. (1977). *Form and meaning*. Longman.

Branigan, H. P., Pickering, M. J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, 118(2), 172–189. https://doi.org/10.1016/j.lingua.2007.02.003

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In G. Bouma, I. Krämer, & J. Zwarts (Eds.), *Cognitive foundations of interpretation* (pp. 69–94). Royal Netherlands Academy of Science.

Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge University Press. https://doi.org/10.1017/9781316410899

Bybee, J. (2010). *Language, usage and cognition*. Cambridge University Press. https://doi.org/10.1017/CBO9780511750526

Cao, H. (2023). 现代汉语显赫范畴扩张效应的测量–以复合趋向补语“起来”为例 [Measurement of the expanding effect of the prominent category in modern Chinese – A case study of the compound directional complement “*qilai*”]. *Chinese Language Learning*, (3), 63–71.

Chao, Y. (1968). *A grammar of spoken Chinese*. University of California Press.

Chen, J. (1980). “拿出一本书来” 的同义句式 [Synonymous structures of “*na chu yiben-shu lai* ‘take out a book’ ”]. *Chinese Language Learning*, (2), 15–19.

Chen, P. (2004). Identifiability and definiteness in Chinese. *Linguistics*, 42(6), 1129–1184. https://doi.org/10.1515/ling.2004.42.6.1129

Chen, Q. (2025). Profiling analytic causative construction in Chinese: A multifactorial analysis of diachronic change. *Corpus Linguistics and Linguistic Theory*, 21(1), 201–236. https://doi.org/10.1515/cllt-2023-0062

Chen, Z. (2007). 复合趋向补语中"来/去"的句法分布顺序及其理据 [The distributional sequence and motivation of "*lai*/*qu*" as complex directional complements]. *Contemporary Linguistics*, 9(1), 39–43.

Croft, W. (2022). *Morphosyntax: Constructions of the world's languages*. Cambridge University Press. https://doi.org/10.1017/9781316145289

De Smet , H. (2008). Functional motivations in the development of nominal and verbal gerunds in middle and early modern English. *English Language and Linguistics*, 12(1), 55–102. https://doi.org/10.1017/S1360674300700250X

Dai, H., & Huang, H. (1988). 时间顺序和汉语的语序 [Temporal sequence and word order in Chinese]. *Foreign Linguistics*, (1), 10–20.

Desagulier, G. (2017). *Corpus linguistics and statistics with R: Introduction to quantitative methods in linguistics*. Springer. https://doi.org/10.1007/978-3-319-64572-8

Dictionary Editorial Office, Institute of Linguistics, Chinese Academy of Social Sciences. (2016). 现代汉语词典 [*Modern Chinese dictionary*] (7th ed.). The Commercial Press.

Dong, X. (1998). 述补带宾句式中的韵律制约[Prosodic constraints in verb-complement constructions with objects]. *Studies in Language and Linguistics*, (1), 55–62.

Fan, J. (1963). 动词和趋向性后置成分的结构分析 [A structural analysis of verbs and postverbal directional complements]. *Studies of the Chinese Language*, (2), 136–160.

Fang, Y. (1992). "起来"的分布和语义特征 [The distribution and semantic features of "*qilai*"]. *Chinese Teaching in the World*, (1), 23–28.

Fang, Y., & Liu, H. (2021). Predicting syntactic choice in mandarin Chinese: A corpus-based analysis of *ba* sentences and SVO sentences. *Cognitive Linguistics*, 32(2), 219–250. https://doi.org/10.1515/cog-2020-0005

Feng, S. (1996). 论汉语的韵律结构及其对句法构造的制约 [On the prosodic structure of Chinese and its constraints on syntactic construction]. *Linguistic Research*, (1), 108–122.

Fonteyn, L. (2017). The aggregate and the individual: Thoughts on what non-alternating authors reveal about linguistic alternations – A response to Petré. *English Language and Linguistics*, 21(2), 251–262. https://doi.org/10.1017/S1360674317000119

Fonteyn, L. (2019). *Categoriality in language change: The case of the English gerund*. Oxford University Press. https://doi.org/10.1093/oso/9780190917579.001.0001

Fonteyn, L., & Nini, A. (2020). Individuality in syntactic variation: An investigation of the seventeenth century gerund alternation. *Cognitive Linguistics*, 31(2), 279–308. https://doi.org/10.1515/cog-2019-0040

Givón, T. (1991). Isomorphism in the grammatical code. *Studies in Language*, 15(1), 85–114. https://doi.org/10.1075/sl.15.1.04giv

Glynn, D. (2014). Techniques and tools. In D. Glynn & J. Robinson (Eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy* (pp. 307–341). John Benjamins Publishing Company. https://doi.org/10.1075/hcp.43.12gly

Guan, Y., & Guo, R. (2024). 略论类时体标记在汉语口语互动中的功能–以"起来" "下去" "来着" 为例 [On the functions of quasi-tense-aspect markers in spoken Chinese interactions: Taking "*qilai*, *xiaqu*, and *laizhe*" as examples]. *Lexicographical Studies*, (3), 68–79.

Guo, C. (2003). 复合趋向补语与非处所宾语的位置问题补议 [Further discussion on the placement of compound directional complements and non-locative objects]. *Chinese Teaching in the World*, (3), 37–45.

Haiman, J. (1983). Iconic and economic motivation. *Language*, 59(4), 781–819. https://doi.org/10.2307/413373

Haiman, J. (Ed.). (1985). *Iconicity in syntax*. John Benjamins Publishing Company. https://doi.org/10.1075/tsl.6

Hao, Q. (2018). 清中叶以来北京话述趋式带宾语语序的历时变化 [The diachronic change of word order for the expression of verb-directional complement constructions with objects in Beijing dialect since the middle of qing dynasty]. *Linguistic Series*, 58, 191–222.

Hawkins, J. A. (1983). *Word order universals*. Academic Press.

Hiroshi, I. (2021). 汉语"复合动趋式+宾语"的语序问题 [Word order issues of "verb-compound directional verb constructions" in mandarin Chinese]. *Language Teaching and Research*, (4), 44–55.

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

Huang, S., & Hsieh, S. C. (2008). Grammaticalization of directional complements in mandarin Chinese. *Language and Linguistics*, 9(1), 49–68.

Jia, Y. (1998). "来/去"作趋向补语时动词宾语的位置 [The position of the verb object when "lai/qu" functions as a directional complement]. *World Chinese Teaching*, (1), 41–46.

Jiang, H. (2020). 评注类V起来话语标记的功用与成因 [The use and the formation of the commentary "V *qilai*" discourse marker]. *Journal of Xinjiang University (Philosophy, Humanities & Social Science)*, 48(2), 149–156.

Jing-Schmidt, Z., & Tao, H. (2009). The mandarin disposal constructions: Usage and development. *Language and Linguistics*, 10(1), 29–58.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* . Stanford University Press.

Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. John Benjamins Publishing Company. https://doi.org/10.1075/z.195

Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar*. University of California Press. https://doi.org/10.1525/9780520352858

Li, J. (1924). 新著国文语法 [*Newly written national language grammar*]. The Commercial Press.

Li, J. (2017). "宾补争动"的焦点实质 [On the focal essence of verb-object and verb-complement competition]. *Chinese Language Learning*, (5), 12–19.

Li, T. F. (2018). Extending the Talmyan typology: A case study of the macro-event as event integration and grammaticalization in mandarin. *Cognitive Linguistics*, 29(3), 585–621. https://doi.org/10.1515/cog-2016-0050

Li, Y., Szmrecsanyi, B., & Zhang, W. (2023). The theme-recipient alternation in Chinese: Tracking syntactic variation across seven centuries. *Corpus Linguistics and Linguistic Theory*, 19(2), 207–235. https://doi.org/10.1515/cllt-2021-0048

Liang, Y. (2007). 汉语趋向动词的语法化 [*Grammaticalization of directional verbs in Chinese*]. Xue Lin Publishing House.

Liu, M. (2023). Towards a dynamic behavioral profile of the mandarin Chinese temperature term re: A diachronic semasiological approach. *Corpus Linguistics and Linguistic Theory*, 19(2), 289–321. https://doi.org/10.1515/cllt-2021-0046

Liu, M. (2025). *Modeling a network of the ba-constructions in contemporary mandarin: A construction grammar approach*. Advance online publication. https://doi.org/10.1075/rcl.00242.liu

Liu, M., & Cuyckens, H. (2025). Alternation in the mandarin disposal constructions: Quantifying their evolutionary dynamics across twelve centuries. *Corpus Linguistics and Linguistic Theory*, 21(1), 75–105. https://doi.org/10.1515/cllt-2023-0038

Liu, N., & Li, T. F. (2024). Event integration as a driving force of language change: Evidence from Chinese 使-shǐ-make. *Language and Cognition*, 16(2), 307–328. https://doi.org/10.1017/langcog.2023.36

Liu, Y. (1998). 趋向补语通释 [*An inventory of directional complements*]. Beijing Language and Culture University Press.

Lu, J. (2002). 动词后趋向补语和宾语的位置问题 [Concerning the verbal complement of direction and the position of object]. *Chinese Teaching in the World*, (1), 5–17.

Lyu, S. (1980). 现代汉语八百词 [*Eight hundred words in modern Chinese*]. The Commercial Press.

Lyu, S. (1985). 动趋式 + 宾语的顺序 [The order of directional complement constructions and object]. *Studies of the Chinese Language*, (3).

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, 4(226), 1–16. https://doi.org/10.3389/fpsyg.2013.00226

Peyraube, A. (2006). Motion events in Chinese. In M. Hickmann & S. Robert (Eds.), *Space in languages: Linguistic systems and cognitive categories* (pp. 121–135). John Benjamins Publishing Company. https://doi.org/10.1075/tsl.66.08pey

Peyraube, A., & Wiebusch, T. (2020). New insights on the historical evolution of differential object marking (DOM) in Chinese. In J. Z. Xing (Ed.), *A typological approach to grammaticalization and lexicalization* (pp. 101–130). De Gruyter. https://doi.org/10.1515/9783110641288-005

Qi, H., & Zeng, C. (2009). "V起来"的语义分化及相关问题 [The semantic divergence of "V *qilai*"]. *Chinese Language Learning*, (2), 3–11.

R Core Team. (2024). R: A language and environment for statistical computing (Version 4.4.2) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

Radden, G. (2021). Iconicity. In X. Wen & J. R. Taylor (Eds.), *The Routledge handbook of cognitive linguistics* (pp. 268–296). Routledge. https://doi.org/10.4324/9781351034708-19

Radden, G., & Dirven, R. (2007). *Cognitive English grammar*. John Benjamins Publishing Company. https://doi.org/10.1075/clip.2

Song, J. (2019). 从韵律角度谈VC1C2带宾语形成的句式 [A prosodic study on the patterns of "V C_1 C_2 + O"]. *Studies in Prosodic Grammar*, 4(2), 124–140.

Song, W. (2007). 现代汉语动结式的认知研究 [*A cognitive study of resultative constructions in modern Chinese*]. Peking University Press.

Song, Y. (1980). 说"起来"及与之有关的一种句式 [On "*qilai*" and an associated syntactic structure]. *Foreign Language Teaching and Research*, (1), 16–21.

Talmy, L. (2000). *Toward a cognitive semantics*. MIT Press.

Tian, X., Speelman, D., & Zhang, W. (2025). Unravelling the competing dynamics of Chinese causative markers shi 使, ling 令, jiao1 叫 and jiao2 教: A diachronic analysis. *Language Sciences*, 108, 101701. https://doi.org/10.1016/j.langsci.2024.101701

Torres Cacoullos, R., & Walker, J. A. (2009). The present of the English future: Grammatical variation and collocations in discourse. *Language*, 85(2), 321–354. https://doi.org/10.1353/lan.0.0110

Traugott, E. C. (1995). Subjectification in grammaticalisation. In D. Stein & S. Wright (Eds.), *Subjectivity and subjectivisation: Linguistic perspectives* (pp. 31–54). Cambridge University Press. https://doi.org/10.1017/CBO9780511554469.003

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional changes*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679898.001.0001

Nice V., K. Y., & Dietrich, R. (2003). Task sensitivity of animacy effects: Evidence from German picture descriptions. *Linguistics*, 41(5), 825–850. https://doi.org/10.1515/ling.2003.027

Vogels, J., & Bergen V., G. (2017). Where to place inaccessible subjects in Dutch: The role of definiteness and animacy. *Corpus Linguistics and Linguistic Theory*, 13(2), 369–398. https://doi.org/10.1515/cllt-2013-0021

Wang, G. (2005). 趋向问题研究 [*Analysis of directional issues*]. Huaxia Press.

Wang, L. (2014). 中国现代语法 [*Chinese modern grammar*]. Zhonghua Book Company. (Original work published 1943).

Wang, Y., & Wen, B. (2015). 评价性"V-起来"句的句法推导 [Syntax of the evaluative "NP + V-*qilai* + AP" construction]. *Foreign Language Research*, (5), 47–52.

Wu, F. (2003). 再论处置式的来源 [Further discussion on the origin of the disposal construction]. *Studies in Language and Linguistics*, 23(3), 1–14.

Wu, W. (2012). "V起来"构式的多义性及其话语功能: 兼论英语中动句的构式特征 [Polysemy of the construction "V *qilai*" and its discourse function: With a discussion on the constructional features of the middle voice in English]. *Chinese Language Learning*, (4), 3–31.

Xu, J. (2023). 语料库研究方法 [*Research methods in corpus linguistics*]. Foreign Language Teaching and Research Press.

Xu, M., Li, F., & Szmrecsanyi, B. (2024). Modeling the locative alternation in mandarin Chinese. *International Journal of Corpus Linguistics*, 29(2), 258–285. https://doi.org/10.1075/ijcl.22072.xu

Yang, D. (2003). 英语母语学习者趋向补语的习得顺序 — 基于汉语中介语语料库的研究 [Sequence of acquiring the directional complement by English-speaking learners of Chinese]. *Chinese Teaching in the World*, (2), 52–65+52–53.

Yang, D. (2005). "时间顺序原则" 与 "动词 + 复合趋向动词" 带宾语形成的句式 ["Temporal sequence principle" and sentence formation with "verb + compound directional verb" with objects]. *Chinese Teaching in the World*, (3), 56–65.

Yang, D. (2011). VC1OC2中宾语的语义类及范畴化现象 [On the case and categorization of the object in the structure]. *TCSOL Studies*, (1), 59–66.

Yang, D. (2017). 趋向补语的认知和习得研究 [*Cognitive and acquisition research on directional complements*]. Peking University Press.

Yang, K. (2006). 论趋向补语和宾语的位置 [The position of directional complements and objects]. *Journal of Chinese Linguistics*, (2), 55–61.

Yang, Y., & Li, T. F. (2025). A diachronic study on the mandarin complex directional complement *guòlái*'come over' from the macro-event perspective. *Cognitive Linguistics*, 36(2), 261–297. https://doi.org/10.1515/cog-2024-0026

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press

Zaenen, A., Carletta, J., Garretson, G., Bresnan, J., Koontz-Garboden, A., Nikitina, T., O'Connor, M. C., & Wasow, T. (2004). Animacy encoding in English: Why and how. In D. K. Byron & B. Webber (Eds.), *Proceedings of the workshop on discourse annotation (pp. 118–125)*. Association for Computational Linguistics.

Zehentner, E. (2021). Alternations emerge and disappear: The network of dispossession constructions in the history of English. *Corpus Linguistics and Linguistic Theory*, 17(3), 525–561. https://doi.org/10.1515/cllt-2020-0074

Zhan, W., Guo, R., Chang, B., Chen, Y., & Chen, L. (2019). 北京大学CCL语料库的研制 [Development of the Peking University CCL corpus]. *Corpus Linguistics*, 6(1), 71–86. Available online at http://ccl.pku.edu.cn:8080/ccl_corpus/.

Zhang, B. (1991a). 动趋式里宾语位置的制约因素 [Constraints on object position in directional complement constructions]. *Chinese Language Learning*, (6), 4–8.

Zhang, B. (1991b). 关于动趋式带宾语的几种语序 [Word orders of directional complement constructions carrying objects]. *Studies of the Chinese Language*, (3), 183–191.

Zhang, B., & Fang, M. (1996). 汉语功能语法研究 [Research on Chinese functional grammar]. Jiangxi Education Press.

Zhang, J. (2018). 功能和演变视角的汉语语法研究 [*The study of Chinese grammar from the perspective of function and evolution*]. SDX Joint Publishing Company.

Zhao, M. (2004). 也谈"复合动趋式 + 宾语"句式 [A further discussion on the "compound verb-directional + object" construction]. In *Proceedings of the 2004 Annual Conference of the Jiangxi Linguistic Society* (pp. 72–77).

Zhu, D. (1982). 语法讲义 [*Grammar lectures*]. The Commercial Press.

Zhu, J. (2019). "V起来"误代偏误的认知语义分析及教学应用 [Cognitive semantic analysis and pedagogical applications of the misrepresentation bias in "V *qilai*"]. *Applied Linguistics*, (2), 84–93.

Cite this article: Yang, Y., & Li, F. T. (2026). Diachronic competition and functional differentiation in object-involving complex directional complement structures: a multifactorial case study of *qǐlái*, *Language and Cognition*, 18, e48, 1–29. https://doi.org/10.1017/langcog.2026.10097

https://doi.org/10.1017/langcog.2026.10097 Published online by Cambridge University Press