

An Investigation into the Omission of Chinese Object Errors by Japanese and Korean Learners — An Analysis Based on the HSK Dynamic Writing Corpus

Wenpan Guo ^{1,2,*}

¹ School of Literature, Taiyuan Normal University, Taiyuan 030600, China

*** Correspondence:**

Wenpan Guo

2787360340@qq.com

Received: 1 June 2026/ Accepted: 10 September 2026/ Published online: 14 September 2026

Abstract

Using the HSK Dynamic Composition Corpus 3.0, this study examines 541 instances of object omission in Chinese compositions written by Japanese and Korean learners. The screened sample contains 277 instances from Japanese learners and 264 from Korean learners. Nominal objects account for 487 cases (90.02%), including 325 nouns and noun phrases (60.07% of the full sample), whereas predicative objects account for 54 cases (9.98%). The errors cluster around common verbs, recoverable referents, and constructions in which one noun phrase serves more than one syntactic or semantic function. These patterns suggest gaps in learners' knowledge of verb argument structure and the discourse conditions that license omission. Crosslinguistic influence may contribute, but the corpus does not support reducing the pattern to a simple SOV and SVO contrast. The paper proposes teaching verbs with typical objects and constructions, comparing acceptable ellipsis with incomplete written clauses, and extending correction from isolated sentences to connected discourse.

Keywords: Object Omission; Japanese Learners; Korean Learners; Learner Corpus; Chinese as a Second Language

1. Introduction

Objects are central to the argument structure of many Chinese verbs, yet their overt expression is sensitive to grammar, lexical meaning, and discourse. Mandarin permits unexpressed arguments when a referent is salient and recoverable, especially in conversational exchanges (Huang, 1984; Li & Thompson, 1981). This flexibility creates a practical problem in learner writing. A missing object can be interpretable and still make a clause structurally incomplete or stylistically unnatural.

Learner errors offer evidence about a developing linguistic system. Error analysis treats recurrent forms as patterned data instead of accidental failures (Corder, 1967), while the concept of

interlanguage explains why learners may build stable rules that differ from target-language usage (Selinker, 1972). In the present data, learners often select an appropriate verb but fail to realize its object. Examples include cannot answer in Chinese, can overcome, lack and suffer from hunger, and cannot concentrate. The intended meaning is generally clear, but the verb-object relation has not been fully encoded in the sentence.

Previous descriptions of Chinese grammar have discussed transitivity, argument structure, and the discourse licensing of null arguments (Huang et al., 2009; Huang & Liao, 2017). Research on Japanese and Korean has also shown that missing objects may receive interpretations supported by syntactic and discourse conditions (Hoji, 1998; Kim, 1999). These findings make first-language influence plausible, but transfer must be demonstrated through converging evidence. A difference in basic word order alone does not establish transfer (Jarvis, 2000; Odlin, 1989).

This study gives a combined descriptive account of Japanese and Korean learner data from the HSK corpus. It asks three questions. What grammatical categories occur among the 541 reconstructed objects? Which lexical, contextual, and syntactic environments recur in the errors? What explanations and teaching responses are consistent with the observed distribution? Because the available coding does not support a controlled group comparison, the paper reports shared tendencies and does not claim that the two learner groups differ in frequency or developmental path.

2. Theoretical Background

2.1. Chinese Object Omission and Discourse Licensing

Chinese allows subjects and objects to remain unpronounced under identifiable discourse conditions. The interpretation of these elements depends on syntactic position, discourse topic, and the availability of an antecedent (Huang, 1984; Huang et al., 2009). A short reply such as Have you eaten followed by Yes, I have is natural because the relevant object is established by the question. Written clauses place a greater burden on the sentence itself. When a verb's complement is needed to complete the intended relation or prevent referential uncertainty, recoverability alone does not guarantee an acceptable omission.

This distinction matters for learner-corpus annotation. A transitive verb does not require an overt object in every use, and an absent noun phrase is not automatically an error. The analyst must consider the verb's intended sense, the local construction, and the surrounding discourse. Functional accounts of Mandarin likewise show that grammatical choices respond to information status and discourse organization (Li & Thompson, 1981). The present study therefore uses contextual acceptability, not transitivity by itself, to distinguish licensed ellipsis from erroneous omission.

2.2. Learner Errors and Crosslinguistic Influence

Japanese and Korean permit unexpressed arguments in contexts where their reference can be recovered. Analyses of Japanese null objects and Korean nominal ellipsis show that apparently similar surface gaps can arise from different structures and interpretations (Hoji, 1998; Kim, 1999). Learners may carry over expectations about when a referent can remain implicit. Even so, a corpus instance cannot reveal its own psychological source. Evidence for transfer should include patterns

that are consistent within an L1 group and that differ in predicted ways across groups or target-language baselines (Jarvis, 2000).

The same caution applies to word order. Japanese and Korean are commonly described as SOV languages, whereas Mandarin has a canonical SVO order. That contrast concerns the position of expressed constituents; it does not directly explain whether an object is pronounced. Studies of null arguments in second-language acquisition show that learners can have persistent difficulty with the distribution of overt and null arguments even after acquiring basic word order (Yuan, 1997).

2.3. Verb Knowledge and Recurrent Word Combinations

Knowing a verb involves more than recalling a translation. Learners also need its argument pattern, the semantic types of its complements, and the combinations that speakers regularly use. Learner-corpus research has shown that collocational restrictions remain difficult even at advanced levels (Nesselhauf, 2005), and formulaic sequences can reduce the processing required during production (Wray, 2002). Vocabulary instruction is therefore more useful when word meaning is taught together with typical grammatical and collocational behavior (Nation, 2013). This perspective is relevant to errors such as *overcome without difficulties* or *concentrate without attention*.

3. Data and Method

3.1. Corpus and Sample

Learner corpora make it possible to examine recurrent forms across many texts while retaining evidence from their original contexts (Granger et al., 2015). The data come from the HSK Dynamic Composition Corpus 3.0 maintained by Beijing Language and Culture University (n.d.). An initial query for object omission returned 751 annotated instances. Each instance was reviewed in its sentence and broader composition context. Cases that did not meet the study's criteria, including acceptable ellipsis and records that could not be reconstructed with sufficient confidence, were excluded. The final sample contains 541 instances: 277 from Japanese learners and 264 from Korean learners.

3.2. Identification and Reconstruction

Three questions guided the decision to retain a case. First, in its intended sense and construction, does the verb require or strongly favor an overt object? Second, if the referent can be inferred, is the resulting clause still structurally and idiomatically acceptable without the object? Third, does the omission meet the expectations of written discourse for referential clarity and syntactic completeness? The omitted constituent was then reconstructed from the verb, nearby clauses, and the topic of the composition.

For example, one learner wrote that many places suffer from hunger because they lack. The following phrase makes food the most plausible object, and the coordination is more natural when it is stated. In another case, a learner wrote that some people are not afraid of setbacks and can overcome. *Difficulties* is recoverable from setbacks, yet the written clause strongly favors an overt

complement. These examples were retained because the missing element was both contextually identifiable and grammatically relevant.

3.3. Coding and Analysis

The reconstructed objects were coded by grammatical category. Nominal objects were divided into personal, demonstrative, and interrogative pronouns; nouns and noun phrases; and quantitative phrases. Predicative objects were divided into verbs or verb phrases and adjectives. Frequencies and percentages were calculated for the full sample. The analysis then grouped examples by recurring lexical environments, discourse recoverability, and syntactic complexity. The counts describe the screened records; no inferential test of differences between Japanese and Korean learners was conducted.

4. Results

4.1. Grammatical Distribution

Table 1 shows that the omitted elements are overwhelmingly nominal. Nominal objects account for 487 cases (90.02%), and predicative objects account for 54 (9.98%). Nouns and noun phrases form the largest single category, with 325 cases (60.07%), followed by personal pronouns, with 140 (25.88%). The remaining categories each account for fewer than 10% of the sample.

Table 1 Distribution of object omission errors by grammatical category

Major category	Reconstructed object	Illustrative forms	n	%
Nominal	Personal pronouns	nin 'you polite' and ninmen 'you plural polite'	140	25.88
Nominal	Demonstrative pronouns	zhe 'this' and nali 'there'	13	2.40
Nominal	Interrogative pronouns	shei 'who' and shenme 'what'	7	1.29
Nominal	Nouns and noun phrases	bieren 'others' and duzi 'stomach'	325	60.07
Nominal	Quantitative phrases	yidian 'a little' and yi sheng 'one call'	2	0.37
Predicative	Verbs and verb phrases	ting 'stop' and jianghua 'speak'	51	9.43
Predicative	Adjectives	henduo 'many' and yiyang 'the same'	3	0.55
Total			541	100.00

Note. Percentages may not total 100 because of rounding.

The distribution suggests that many learners can identify the intended participant or entity but do not consistently express it where the clause requires it. The pattern is therefore compatible with incomplete argument realization, not a general absence of the relevant nouns from the learner's vocabulary. The category counts alone cannot determine why an object was omitted, so the following sections examine the environments in which the omissions occur.

4.2. Common Verbs and Fixed Patterns

Many examples involve frequent verbs and familiar verb-object combinations. Learners wrote forms equivalent to answer in Chinese, overcome, face and solve, and clean, where the intended objects were the question, the difficulties, the difficulties, and the room. The verbs themselves were semantically appropriate. What failed was the realization of the complement selected or strongly favored by the construction.

Frequency of exposure does not ensure accurate use. If a verb is learned mainly through a first-language gloss, its collocations and complement patterns may remain weak. Under timed writing conditions, the learner may retrieve the verb before activating the noun phrase that completes it. This account fits the concentration of errors around ordinary verbs that learners are otherwise able to use.

4.3. Contextually Recoverable Objects

A second group contains objects that are easy to infer. In the clause many places suffer from hunger because they lack, the phrase suffer from hunger points to food. In a retelling about monks who carry water, the noun water may appear elsewhere in the passage and remain salient when a later verb is produced. Such examples suggest that some learners treat discourse recoverability as a sufficient condition for omission.

Chinese does allow this strategy in suitable contexts, but acceptable ellipsis depends on more than whether a reader can guess the missing referent. The verb's construction and the informational work of the clause also matter. The learner's rule may therefore be systematic but too broad: a recoverable object is omitted even when its expression would make the written clause more complete and idiomatic.

4.4. Complex Syntactic Environments

Omissions also occur when learners must coordinate several relations at once. In a clause equivalent to the measures taken by the relevant departments enabled to, the missing noun phrase is the object of enable and the understood subject of the following predicate. A form equivalent to some people remind that smoking affects others requires the writer to represent the person who reminds, the addressee, and the content of the reminder. These constructions place heavier demands on planning than a simple verb-object clause.

Parallel predicates can create a similar problem. A learner who writes cannot concentrate and be calm may treat both predicates as structurally alike, although concentrate in the intended Chinese expression calls for an object such as attention or energy. The adjacent intransitive or adjectival predicate provides a misleading model and increases the chance that the complement will be left unrealized.

5. Discussion

5.1. Incomplete Argument Structure Knowledge

The largest pattern in the corpus is a separation between lexical meaning and syntactic use. Learners choose verbs such as answer, overcome, lack, concentrate, and remind appropriately, but they do not always encode the participants required by those verbs. This finding agrees with work on collocation and formulaic language: lexical knowledge becomes usable in fluent production when words are linked to recurrent combinations and constructional frames (Nesselhauf, 2005; Wray, 2002).

A simple vocabulary deficit would predict semantically unsuitable word choice or failed reconstruction. Many of the present examples show the opposite. The correct object is supplied by the context, and the learner has chosen a compatible verb. The remaining difficulty lies in mapping this meaning onto an overt noun phrase in the current construction.

5.2. Overextension of Ellipsis

The errors also reflect an overextended target-language rule. Learners encounter natural exchanges in which a known object is omitted, such as Have you read this book followed by Yes, I have. They may generalize from such examples that any recoverable object can be left implicit. Interlanguage theory predicts this kind of regular learner-created rule, especially when the resulting sentence remains comprehensible and therefore attracts little corrective attention (Selinker, 1972).

This explanation does not require learners to be unaware of Chinese ellipsis. On the contrary, they may be using a genuine property of Chinese without having acquired its restrictions. The contrast between successful communication and target-like written form helps explain why these omissions can persist.

5.3. Possible Crosslinguistic Influence

First-language habits may reinforce the use of implicit arguments. Japanese and Korean both allow unexpressed objects under discourse and structural conditions, and their null-object patterns have received detailed analyses (Hoji, 1998; Kim, 1999). A learner who routinely relies on contextual recovery may carry that processing preference into Chinese writing.

The present data support this account only as a hypothesis. A defensible transfer claim would require a Chinese baseline, comparable learner groups with different first languages, and evidence that the predicted pattern is stronger within the relevant L1 groups (Jarvis, 2000). The current sample also combines Japanese and Korean instances. It can identify shared environments of omission but cannot establish group-specific transfer effects.

5.4. Task and Instructional Factors

Instruction that emphasizes word meanings without complement patterns may leave gaps in productive knowledge. A student may know the meaning of concentrate without knowing the common expressions concentrate attention or concentrate energy. The same student may know remind but not the frame remind someone to do something. Presenting such information as part of

vocabulary knowledge follows established principles for learning words in context (Nation, 2013) and for teaching Chinese as a second language (Liu, 2000).

Timed composition can amplify the problem. Writers often prioritize ideas and lexical retrieval, and an omitted object may not block understanding. If feedback concentrates on content or conspicuous word-choice errors, the writer may receive little evidence that the clause is incomplete. Repeated omissions can then become stable.

6. Pedagogical Implications

6.1. Teach Verbs with Their Objects and Constructions

Common transitive verbs should be introduced with typical objects and complete frames. Useful sets include answer a question, overcome a difficulty, lack experience or food, concentrate attention, clean a room, and remind someone to be careful. Teachers can ask learners to sort possible objects by semantic type and then use each verb in a complete clause. This practice connects lexical meaning with selectional restrictions and sentence structure.

6.2. Compare Licensed and Unlicensed Omission

An absolute rule that every transitive verb must be followed by an object would create new errors. Classroom activities should place acceptable ellipsis beside incomplete written clauses. Learners can compare Yes, I have after an explicit question with The teacher asked me a question, but I cannot answer in Chinese. They should identify the antecedent, judge whether the construction permits omission, and explain what changes when the object is restored.

6.3. Correct Objects in Connected Discourse

Correction should extend beyond isolated sentences. In a paragraph about carrying water, learners can mark each verb, restore any possible object, and decide which repetitions Chinese can omit naturally. For causative frames such as make or let plus noun phrase plus verb phrase, a bracketed analysis can show that the noun phrase is the object of the first verb and the semantic subject of the second. Similar role mapping helps with remind by making the speaker, addressee, and message explicit.

A short revision routine can support transfer to composition. After drafting, the learner identifies each main verb, asks whether its intended use selects an object, and checks whether any absent object is licensed by the context and register. Teachers can keep an individual record of recurring verbs so that feedback targets stable patterns instead of treating each omission as an unrelated mistake.

7. Limitations

The study relies on an existing corpus annotation category and a manually screened sample. Reconstruction of an omitted object depends on contextual judgment, and the available records do

not include an independent second coding or interrater reliability measure. The counts should therefore be read as a descriptive analysis of the retained cases.

The sample also does not control proficiency, task type, or the number of words written by each L1 group. The 277 Japanese and 264 Korean instances are raw error records, not normalized rates. They cannot support claims about which group omits objects more often. Future research should use double coding, report agreement, normalize frequencies by corpus size, and compare learner groups under matched writing conditions.

8. Conclusion

Among the 541 screened omissions, 90.02% involve nominal objects and 60.07% involve nouns or noun phrases. The errors recur with common verbs, contextually recoverable referents, and constructions that distribute several roles across a clause. The distribution is consistent with incomplete knowledge of verb argument patterns and an overly broad use of Chinese ellipsis. First-language influence remains possible, but the current corpus does not isolate it from proficiency, task demands, or instruction. Teaching should connect verbs with their typical objects, make the conditions on ellipsis explicit, and give learners repeated practice in revising connected written discourse.

Author Contributions:

All authors have read and agreed to the published version of the manuscript.

Funding:

This research received no external funding.

Institutional Review Board Statement:

Not applicable.

Informed Consent Statement:

Not applicable.

Data Availability Statement:

The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflict of Interest:

The author declares no conflict of interest.

References

Beijing Language and Culture University. (n.d.). HSK Dynamic Composition Corpus 3.0.
<https://yuyanzyuan.blcu.edu.cn/>

- Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics in Language Teaching*, 5(1-4), 161-170.
- Granger, S., Gilquin, G., & Meunier, F. (Eds.). (2015). *The Cambridge handbook of learner corpus research*. Cambridge University Press.
- Hoji, H. (1998). Null object and sloppy identity in Japanese. *Linguistic Inquiry*, 29(1), 127-152.
- Huang, B., & Liao, X. (2017). *Xiandai Hanyu [Modern Chinese] (6th ed., Vol. 2)*. Higher Education Press.
- Huang, C.-T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 15(4), 531-574.
- Huang, C.-T. J., Li, Y.-H. A., & Li, Y. (2009). *The syntax of Chinese*. Cambridge University Press.
<https://doi.org/10.1017/CBO9781139166935>
- Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning*, 50(2), 245-309. <https://doi.org/10.1111/0023-8333.00118>
- Kim, S. (1999). Sloppy/strict identity, empty objects, and NP ellipsis. *Journal of East Asian Linguistics*, 8(4), 255-284. <https://doi.org/10.1023/A:1008354600813>
- Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar*. University of California Press.
- Liu, X. (2000). *Duiwai Hanyu jiaoyuxue yinlun [Introduction to teaching Chinese as a foreign language]*. Beijing Language and Culture University Press.
- Nation, I. S. P. (2013). *Learning vocabulary in another language (2nd ed.)*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. John Benjamins.
<https://doi.org/10.1075/scl.14>
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524537>
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10(1-4), 209-232. <https://doi.org/10.1515/iral.1972.10.1-4.209>
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511519772>
- Yuan, B. (1997). Asymmetry of null subjects and null objects in Chinese speakers' L2 English. *Studies in Second Language Acquisition*, 19(4), 467-497.

License: Copyright (c) 2026 Author.

All articles published in this journal are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and source are properly credited. Authors retain copyright of their work, and readers are free to copy, share, adapt, and build upon the material for any purpose, including commercial use, as long as appropriate attribution is given.