

Original Paper

A Corpus-based Contrastive Study of Shell Nouns Complement Constructions in English and Chinese Academic Writing

Wanting Zhang^{1*}

¹ Henan Normal University, Henan, China

* Wanting Zhang, Henan Normal University, Henan, China

Received: June 30, 2026

Accepted: July 20, 2026

Online Published: August 7, 2026

doi:10.22158/sll.v10n3p71

URL: <http://dx.doi.org/10.22158/sll.v10n3p71>

Abstract

Shell Nouns Complement Constructions (SNCCs) play a central role in packaging complex information and encoding authorial stance. Previous studies have mainly focused on their classification, stylistics distribution, and cross-linguistic difference in SNCCs use among different groups of writers. Less attention, however, has paid to how the same group of L1 Chinese scholars redistribute constructional resources across different publication contexts. Through the lens of usage-based construction grammar, this study investigates whether SNCCs preferences reflect stable authorial tendencies or context adaptation by comparing Chinese and English research articles written by the same group of Chinese applied linguistics scholars. This research is conducted on the self-built corpus, consisting of research articles published by 13 Chinese applied linguistics scholars in prestigious Chinese and English journals.

The findings indicated that: 1) although clausal complement was the dominant construction family in both contexts, Chinese publications showed a stronger concentration on this construction, whereas English publications demonstrated a significantly higher proportion of verbal complements. No significant contextual difference was found for eventive complements. 2) SNCC preferences should not be understood solely as stable individual writing styles. Instead, constructional choices represent context-sensitive resource allocation shaped by the interaction between authors' habitual tendencies and publication environments. This study aims to provide strategies for using shell noun complement constructions in English and Chinese.

Keywords

shell noun complement constructions, shell noun, construction grammar, academic discourse

1. Introduction

Shell Nouns Complement Constructions (SNCC) are important resources in academic discourse, as they package information, express stance (Jiang, 2016; Biber et al., 1999; Charles, 2007) and contribute to logical cohesion. Schmid (2005) specified five types of SNCC: “SN + that clause”, “SN + to infinitive”, “SN + of + -ing clause”, “SN + be + to-infinitive”, and “SN + be that-clause”. These are markers for readers to find core information or conclusions.

Because of globalization, academic exchanges have no longer been limited to within the country. The most obvious example is that numerous Chinese scholars can be seen in journals such as CSCI and SCI. According to the scientific research articles report of the China National Knowledge Infrastructure (2020), Chinese scholars publish over 400,000 English academic papers every year, and SCI papers in hard sciences account for more than 20% of the world’s share. From the reports of each year, the data has been continuously increasing. This inevitably requires consideration of the rhetorical strategies involved in shifts between English and Chinese, especially the same author’s preferences in the selection of complement constructions.

Extensive research has focused on ontological classification, pedagogy, grammar and discourse functions, as well as their distribution across disciplines (Flowerdew & Forest., 2015; Jiang & Hyland, 2015, 2016; Wang & Du, 2023; Huang & Deng, 2025). Studies in Chinese also mainly investigated these aspects (Flowerdew, 2003; Lou, 2013; Jiang, 2016; Dong, 2017; Qiu et al., 2021). Studies that combine two languages also existed and examined pieces of research articles such as abstracts (Hu & Huang, 2021) and objects like L1 authors and L2 learners, or L1 authors and L1 authors. Few studies have explored the same writers’ preferences of shell noun complement constructions in English and Chinese. This research is therefore to systematically focus on the frequency, range, different distribution, discourse function and writers’ preference.

2. Literature review

2.1 Construction Grammar and Academic Discourse Studies

Construction Grammar has broken through the traditional dichotomy of lexis and syntax in language studies, arguing that the meaning of certain language units cannot be fully derived from their components, but needs to be explained at the overall construction level (Fillmore & Kay & O’Conor, 1988). Goldberg (1955, 2006) further developed the construction theory, pointing out the conventional pairing relationship between form and meaning/function. Under this theoretical framework, language knowledge is not innate rules, but a structured list of conventionality abstract constructions that are solidified by repeated use (Langacker, 1987). Constructions include not only syntactic structures in the traditional sense, but also lexico-grammatical patterns with specific textual functions (Schmid, 2000). Therefore, the high-frequency expression patterns (such as shell nouns + complement forms) in academic discourse can also be regarded as construction resources with specific forms, functions and

contextual conventions.

In recent years, construction grammar has gradually applied to the study of academic discourse, and researchers have begun to pay attention to the form-function pairings which is frequently used in academic writing. Different from the traditional research based on lexical frequency, the core tenet of construction is the language form is related to discourse function. From this perspective, the fixed or semi-fixed pattern in academic writing cannot be simply regarded as syntactic rules, but should be understood as influenced by genre conventions, communication purposes, and the long term practices of the academic community. Biber and Barbieri (2007) found the lexical bundles with stable functions are widely used in academic English. These frequencies show obvious genre characteristics and participate in the construction of information organization and authorial stance expression in academic discourse. Hyland (2008) further pointed out that language selection in academic writing not only serves information transmission, but also plays an important role in the interaction between the author and the reader and the establishment of argumentation relationship. The choice of construction in academic discourse reflects the author's adaptation to the norms of academic community.

According to the usage-based construction grammar, language choice is probabilistic, and the frequency and distribution of input use of constructions are influenced by experience, context and communicative needs (Bybee, 2010). Therefore, the high frequency of a certain construction in a specific context. Gries (2010) pointed out that the choice of linguistic forms can be described by probability distribution, and different users and different contexts may show different degrees of constructions of construction bias. In cross-language academic writing, the author needs to adjust the allocation of language resources according to the target language and publishing environment. However, the existing research mainly focused on the constructional resources in different author groups, and pays insufficient attention to how the same group of core scholars with both Chinese and English top publication records reconfigure constructional resources in different publication contexts. Therefore, it is necessary to investigate the dynamic change mechanism of SNCC from the perspective of cross-publishing context.

2.2 SNCC in Cross-Linguistic Academic Writing

Shell nouns are important abstract noun resources in academic discourse, which refer to a class of nouns with relatively general semantics and need to be concretized by context or subsequent complements, such as fact, idea, possibility, problem, reason, claim, evidence, etc. Early research (Jiang, 2016; Flowerdew, 2003; Halliday & Hasan, 1976; Schmid, 2000; Wei & Zhang, 2016) focuses on the functions of shell nouns, which rely on context to obtain specific meaning, organize information, textual cohesion, express evaluation and position expression. Later researchers gradually put it under the framework of corpus linguistics, cognitive linguistics, pattern grammar and applied linguistics. The discourse functions of different nouns are not exactly the same: words, such as problem, key, core, etc., are more commonly used to organize discussions and lead to discussion focus, while words like

meaning, advantage, deficiency, rationality, etc. are more clearly reflected in evaluation and position expression, reason, basis, premise, possibility, etc. are more used to construct causal, conditional and inferential relationships.

Schmid (2000) pointed out that the meaning of shell nouns is usually incomplete and needs to be concretized by context or specific syntactic structure. Specific syntactic structures include SN + clause, SN + be + complement clause, demonstrative + SN, and demonstrative + be + SN. These structures are not only abstract syntactic forms, but also some aspects cannot be fully predicted from their components, which belong to the constructions of form-meaning pairings (Goldberg, 2006). The above four constructions have the functions of stance markers, encapsulation of complex propositions and textual cohesion, which are the conventional results in the academic discourse community.

Around these constructions, scholars have carried out detailed research: Jiang and Hyland (2015) studied noun complement construction, focusing on how this structure serves as a stance resource; Jiang and Hyland (2015) focused on the position construction and interpersonal function of shell nouns, focusing on disciplinary differences. Hu and Huang (2021) studied the move differences of interpersonal function of shell nouns in English abstracts of Chinese and foreign scientific papers, focusing on the differences between Chinese writers and native writers. Previous studies have fully proved that: 1) clausal complement, especially that-clause, is most frequently used; 2) related functions of shell nouns; 3) distribution differences in different disciplines. However, these studies mainly answer which constructions are used, without answering in detail why the author prefers to use certain constructions. Therefore, this study will focus on the author's construction selection mechanism. Secondly, for cross-group comparison, the research objects of existing research are mainly Chinese authors VS native English authors or L1 VS L2 authors, but the reasons for the differences between the two are difficult to determine, such as mother tongue, educational experience, writing training or publishing environment. This paper adopts the method of intra-author comparison. There is another point to be considered. Cross-linguistic studies usually interpret differences as mother tongue transfer or language system differences, but scholars will reconfigure construction resources when publishing papers in different languages. Based on the above points, this study attempts to solve the following problems:

- 1) Do SNCC differ in Chinese and English research articles written by L1 Chinese scholars in applied linguistics? If so, in what ways?
- 2) How do these scholars redistribute different types of SNCC across the two publication contexts?
- 3) To what extent do these differences reflect context-sensitive restructuring rather than stable authorial preference?

3. Methods

3.1 Self-built Corpus

The research scope is limited to applied linguistics. Based on the criteria of JCR partition and whether it is included in SSCI or CSSCI, 50 English and 50 Chinese papers are selected according to the publication records, academic positions and total citations of Chinese scholars' core journals. We selected research papers published by 13 scholars as single, second or third works, and ensured that each selected English paper and Chinese paper were 2 to 5 each, and the years were as close as possible. Table 1 is the basic information of the corpus.

Table 1. Basic Information of the Corpus

Subject	Language	Tokens	Types	Numbers of articles	Time Period
Linguistic	Chinese	525923	13283	50	2002-2024
	English	348623	13625	50	

3.2 Tools and Procedures

3.2.1 Identification and Annotations of SNCC

In the process of identifying the shell noun complement construction, we use different tools to process English and Chinese materials: Treetagger performs word segmentation and part-of-speech tagging on English text; Chinese corpus uses CorpusWordParser; AntConc is used for Chinese and English corpus retrieval. The above tools provide preliminary results, and then we manually judge the instances that meet the requirements according to Schmid's classification.

3.2.2 Classification

In order to make a cross-linguistic and inter-author comparison, we divide the English and Chinese construction families into three categories: clausal complement, eventive complement and verbal complement. After the identification is completed, we classify all instances into three types of constructions.

3.2.3 Analysis of Constructional Distribution and Intra-author Constructional Variation

We mark the instances that meet the requirements in R Studio, including Text_ID, Language, Sentence ID, Shell Noun, Construction Type, Complemental Type, Construction family and Function. After annotation, descriptive statistics, contingency table and regression modeling are used to analyze the influence of language on the choice of complement constructions. We use R Studio to present the frequency and in-line proportion of three types of sub-constructions (clausal complement, verbal complement, eventive complement) to analyze the distribution and preference characteristics of the three types of constructions. In order to further test the correlation between language and construction, we use the contingency table. The overall distribution difference was evaluated by Fisher's exact test and chi-square test, and the difference intensity was observed by Cramér's V report effect size. On this

basis, we conduct 2x2 Fisher's exact test on the three types of constructions to confirm the overall source of constructional changes within the author. We also use the occurrence of the three types of constructions as a binary dependent variable to establish a regression model: clausal and eventive use mixed effect logistic regression, and set the author as a random intercept, which is mainly to deal with the difference in the baseline tendency of different authors in the use of constructions, so as to estimate the independent influence of publishing context on the choice of constructions. Due to the zero-frequency problem of verbal complement in subsequent data processing, we use the Firth penalty logistic regression model to complete the analysis.

4. Results and Discussion

4.1 Cross-Linguistic Differences in SNCC

In order to analyze the constructional preferences used by similar authors in publishing journals in both English and Chinese, we also summarize three subcategories that can correspond one-to-one and cover English and Chinese, namely, clausal complement, verbal complement and eventive complement. Clausal complement represents shell nouns followed by propositional clauses, including SN-that, SN-be that, and SN + 是 + clauses. Verbal complement refers to shell nouns followed by non-finite verb structures or verb phrases, including SN + be to, SN-to, SN + of doing, SN + 是 + VP, and SN + 是. Eventive complement is a complement with event content as the core, but does not appear as a typical that clause or infinitive structure, including SN + 的 + clause.

There are 284 shell noun complement constructions in this corpus, of which 238 cases (pmw = 4.52) appear in English corpus and 46 cases (pmw = 1.53) appear in Chinese corpus. From Table 2, we can see that the distribution of complement constructions in English and Chinese is not balanced. From the perspective of in-line proportion, the highest proportion of the two languages is clausal complement, which appears 182 times and 43 times in English and Chinese corpus respectively. This shows that Chinese applied linguistics scholars have important resources for information packaging and proposition development when complementing clauses. Corresponding to the shell noun itself as a conceptual container, it is necessary to complement the language to provide specific content. The concentration of Chinese in this type of construction is more prominent, accounting for 93%; verbal complement takes the second place, appearing 48 times and 0 times respectively in English and Chinese corpus, and English shows a higher proportion of verbal complement. The proportion of eventive complement is the least, which appears 7 times and 3 times in English and Chinese corpus respectively.

Table 2. Further Category of Complements

	clausal_complement	eventive_complement	verbal_complement
	/prop	/prop	/prop
	/ Fisher	/ Fisher	/ Fisher
Chinese	43/0.935/0.231	3/0.065	0
English	182/0.768	7/0.030	48/0.203

From Table 2, we can see that the construction of “SN + that-clause” appears frequently in the English corpus, accounting for 37% of the total frequency of the five constructions. As the most typical construction, the high frequency of this kind of construction is not unexpected. Moreover, that-clause perfectly carries the propositional content and satisfies the author’s abstraction of proposition, explanation and conclusion, so it is very in line with the need of proposition encapsulation in academic discourse (Jiang & Hyland, 2015). This construction is also very consistent with the expression habit that English tends to construct logical hierarchy through explicit grammatical means. The source SN governs the whole theme of the discussion, and the meaning is relatively abstract. With the help of the “that clause”, it provides specific details or argumentation basis, forming a total-division cognitive structure. This stratification conforms to the information presentation mode of frame before detail in academic writing, which reduces the reader’s cognitive load (Swales, 1990). The second most frequently used construction is “SN + be + that” (34%), probably because it plays a very important role in the organization of the text. Compared with “SN + that”, this kind of construction is more expressive and reminds readers of the next key content, which is very suitable for explaining the core arguments, research significance, methods and conclusions, which is also the high-frequency text position of this kind of construction.

Compared with more complex and marginal structures, these two types of constructions have low processing cost, stable form and strong replicability. In addition, the construction structure belongs to lexical bundles to some extent, which is helpful for the author to call it repeatedly in academic writing and is easier to form standardized expression. Many combinations of high-frequency shell nouns with that-clause in academic English can also be regarded as semi-fixed collocations, such as “the fact that, the idea that, the possibility that, the argument that”. This high-frequency pairing will further increase the overall frequency of “SN + that” and “SN + be that”. Charles’s (2006) study of “N + that” in academic writing also shows that this pattern is very common in disciplinary argumentation because it can efficiently embed different types of “stance nouns” into argumentation structures.

The structure of “SN + of -ing” has the lowest frequency. The possible reasons are as follows: its semantics involve action behavior, and its applicability and universality are low. A few abstract nouns

belong to shell nouns, and the nominalization of events or activities is further reduced. It is not concise enough to elicit key points or conclusions, as “that-clause”.

Table 3. Frequency of each Type

Construction_Type	English_Frequency	Construction_Type	Chinese_Frequency
SN-that	87	SN+的+小句	3
SN-to	48	SN+在于+VP/小句	17
SN-be that	81		
SN-of doing	7	SN+是+VP/从句	26
SN-be to	15		
Total/p	238/4.53	Total/p	46/1.32

The five shell noun complement constructions in English cannot be directly translated into Chinese. Referring to Schmid’s definition (2000), Kou (2023) summarized three syntactic structures that can judge shell nouns in Chinese: SN + 是 + clause; clause + 这一(/这+ quantifier) + SN; clause + 的 + SN. This paper also adopts the first construction, which basically corresponds to the English construction “SN + be that”, for example:

(1)我们的目标是建立学科分布均衡的大型学术汉语口语语料库，建库原则参照英语学术口语语料库 MICASE 的设计，分学科地对语料进行收集，建成独立的学科子库。

(2)制作词表的前提是基于超大规模的真实语料，当前 SCAC 的规模尚显不足，但在后续扩增其他学科子库后，该库可为研制汉语学术口语词表提供丰富的语料基础。

The second syntactic structure is essentially “demonstrative pronoun + shell noun” or “demonstrative pronoun + copula + shell noun”, corresponding to “this / that + SN”, “th-SN” and “demonstrative + SN” mentioned by Schmid (2000). But strictly speaking, it is not a complement construction, because the complement construction emphasizes that shell nouns need to realize lexicalization in the subsequent complement clauses (Winter, 1982). For example:

(3)其次，是陷阱就有可能较快地跨越或避免，核心的问题是看处理陷阱的方式是否有效。

(4)其原因在于，动词是组织信息的中心，动词短语也一直被视为英语词汇习得的难点。

In Example (3) and Example (4), “问题是” and “原因在于” are shell noun complement constructions. Both of them are SN in the front and complement in the back, but the structure of “clause + 这一(/ 这 + quantifier) + N” is that the proposition is in the front and SN is in the back, which does not satisfy that there is a structure to supplement the content of this shell noun. Therefore, this paper does not consider this type of anaphoric pretext construction. Similarly, the third category does not consider inclusion. Correspondingly, this paper adopts this type of construction-SN + de + clause / proposition content, for example, “这一做法的关键在于控制变量”, “这种解释的前提是语境一致” or “该模型

的优势在于能够解释历时变化”。The core of this category is that shell nouns connect a more complete propositional content by following the “的”. It should be noted that this paper only includes shell noun structures that are connected by “的” and further carry more complete propositional content with the help of structures such as “是/ 在于/ 取决于 /体现在”, but does not include ordinary affiliation, object relationship or general modifier-head structure. Based on the semantic relationship of shell nouns and the actual distribution of corpus, we add “N + 在 + VP / clause”. This kind of construction belongs to the “noun-related” relationship, and its corresponding event cannot be simply said to be “content = clause”, but to establish a connection with other relational components (Kou, 2023). For example, the problem lies in the operational definition of topic chain “and” the significance lies in providing a new research path for language description and analysis. Therefore, this paper uses the following three types of Chinese constructions for frequency analysis and research:

- 1) N+是+VP/小句
- 2) N+的+小句/命题内容
- 3) N+在于+VP/小句

The first frequency in Chinese constructions is “SN + 是 + VP / clause” (56 %), which may be due to the insufficient semantics of shell nouns themselves. It is necessary for the copula “是” to introduce the following VP or clause to supplement the specific content of shell nouns (Schmid, 2000; Jiang, 2016). Moreover, the Chinese organization view tends to be topic-comment, which serves a specific discourse function (Shi, 2015). The core prototypical structure “SN + that-clause” and the secondary structure “SN + be to” in English belong to “un + copula + propositional content”, which conforms to the above framework and appears frequently in sub-corpora. The two perfectly correspond to the Chinese construction “SN + 是 + VP / clause”. Correspondingly, we can also understand that this kind of Chinese construction appears frequently in the Chinese sub-corpus. The shell nouns in front of the copula “是” are mainly “原因, 关键, 目标, 前提, 因素, 问题”, etc., indicating that its position is the core of the article, which plays a role in improving the clarity of proposition expression and promoting the discussion. This structure can also organize information and highlight the focus, because from the specific corpus data : shell nouns are commonly used as topics in Chinese corpus, and then “是 + VP / clause” is used to specify.

The frequency of “SN + 在 + VP / clause” is slightly lower than that of “SN + 是 + VP / clause”. We continue to analyze from several possible reasons. One of the most basic and core grammatical structures in Chinese syntax is copular construction, including the “是” construction (Shi, 2023). “SN + 在 + VP / clause ” is not included, and the semantics of “在于” is more restricted than that of “是”. In this corpus, it is also often matched with a few shell nouns such as “原因, 区别, 意义, 目标, 关键”, and there is no wide range of use of “是” construction. For example:

(5)我们认为, 问题的关键不在于创造新的术语, 而是要准确阐释“语义韵”这一概念, 才能免去纷争。

(6)语言规划重点在于推行语言文字改革，简化和整理汉字以及制定和推行汉语拼音方案。

From the above examples, we can see that in addition to the role of connecting content, this type of construction is in an interpretation-positioning framework, which is not like the neutral and broad framework of the “是” construction (Shi, 2023). Academic discourse needs to be defined and explained frequently, and the “is” construction is more in line with the requirements.

From table 3, we note that “SN + de + clause” accounts for the smallest proportion, only 7 %, which may be because the preference in academic discourse can directly lead to a complete proposition, organize argumentation and show the structure of position (Jiang & Hyland, 2015). In Chinese grammar, the most common and basic function of “的” is to connect premodifiers with central nouns, which are used to modify attributes (such as 语言的特点, 词汇的意义) and give nouns properties, features or qualifications (Zhu & Huang, 2023). Therefore, in most cases, “的” does not play a role in specifically marking the proposition to complement the relationship. Only a small number of examples that really carry the proposition content and complete events will be included in the subject analysis. For example:

(7)以上研究结果的差异可能是学习者母语背景不同和研究者的 T 单位操作定义不同造成的。

(8)其中一个意义的描述是 if you consider a person or thing to be something, you have the opinion that this is what they are, ...

The clauses after “差异” and “描述” in Example (7) and Example (8) both explain the specific content of pre-shell nouns, so they belong to the complement construction.

4.2 Redistribution of Construction Types

There is a significant correlation between the language and the family of complement constructions. The results of Fisher’s exact test showed that the association was significant ($p = 0.0001929$), and the results of Chi-square test were also significant ($\chi^2 = 12.053$, $df = 2$, $p = 0.002413$). Cramér’s V was 0.206, indicating that there was a small to moderate association between language and construction family selection. We still need to further analyze which category leads to the most obvious difference, so this study makes 2x2 Fisher test for these three subcategories. Further Fisher’s exact test shows that the language difference of clausal complement is significant ($p = 0.008865$), and Chinese scholars are more inclined to use clausal complement. The language difference of verbal complement is the most prominent ($p = 0.0001434$). The main reason is that the construction has no examples in Chinese corpus, showing cross-language asymmetry. In contrast, the language difference of eventive complement is not significant ($p = 0.2107$), indicating that this type of construction is not the main source of the overall distribution difference.

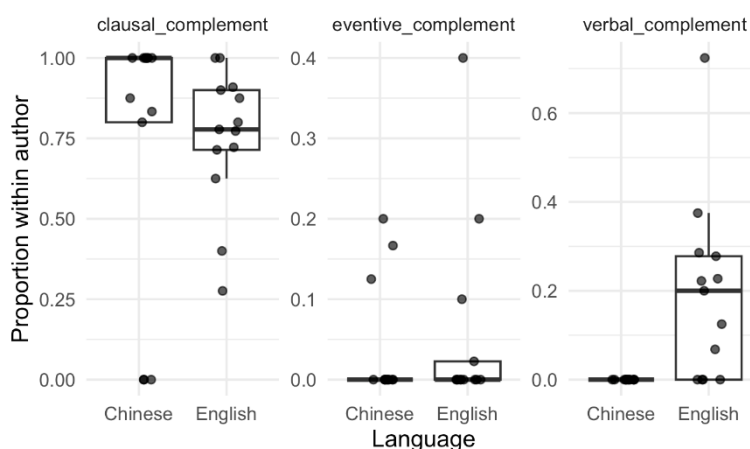


Figure 1. Intra-Author Proportional Distribution

According to the figure, the author's internal scale diagram shows the cross-context adjustment made by the author. Most of the Chinese values under the Clausal complement classification are 0.8-1.0, and the English values are reduced to 0.6-0.9. This shows that most writers will reduce the use of clausal complement in English context. The value of Verbal complement is more extreme, 0 in Chinese, and multiple authors in English are located in 0.1-0.3, which indicates that Verbal complement is more used in English context. Eventive complement changes little, which to some extent indicates that not all constructions are affected by context.

Compared with Chinese corpus, the probability of Chinese scholars using clausal complement in English corpus is significantly reduced. Because the value of clausal complement in English is negative, clausal complement is less likely to appear in English corpus than in the benchmark group Chinese. The advantage of clausal complement in English is about 0.157 times that in Chinese.

The reason behind these data may be that the classification resources of complement constructions in English are better than those in Chinese. Specifically, English has relatively mature non-finite verb complement resources, such as to do and of doing, which will compete with the clausal family. However, there are few alternative options in Chinese, and there is a lack of completely corresponding, universally recognized and grammaticalized complement categories such as English non-finite verbs, so the proportion of clausal is increased. In his review of the research history of Non-finiteness, Yang (2022) also mentioned that many scholars do not accept the existence of "on-finite" or infinitive distinctions in the strict sense of Chinese. For example, Hu, Pan and Lie jiong Xu (2001) objected that "there is a strict distinction between finite and non-finite in Chinese," which is also reflected in a relatively new review (Tang, 2023). In addition, in the register of academic writing, phrase complexity is more representative and important than clause complexity in academic English (Biber & Gray, 2016).

Since there is no instance of verbal complement in Chinese corpus, we use Firth penalty logistic regression because of the problem of complete separation in ordinary logistic regression. The results showed that the language effect was significant ($\beta = 3.170$, $SE = 1.431$, $p < 0.001$), and the 95% confidence interval did not cross 0 ([1.193, 8.015]). The advantage of verbal complement in English is about 23.8 times that in Chinese. Therefore, the author tends to use verbal complement in English corpus, which is also the most prominent source of cross-linguistic differences among the three types of constructions. The zero frequency phenomenon of Verbal complement does not mean that there is no behavioral content in Chinese, but for this kind of content, the Chinese author does not use the typical non-finite verb complement form in English, but diverts to other families, especially the clausal family. The main reason is that they do not form a stable verbal family in the corpus, classification criteria and authors of this study.

Compared with the other two categories, the influence of English and Chinese on eventive complement is not significant. Although the coefficient is negative, the author may use less eventive complement in English, but the value is less than clausal complement, so it is not significant. The main reason may be that it belongs to marginal resources in both English and Chinese, and the author has better use of superior resources.

When writing English and Chinese academic papers, Chinese scholars are not only changing the corresponding vocabulary according to different languages, but also reconstructing how to fill in the content of shell nouns. In order to put a complex proposition, process or research judgment behind shell nouns, it is easier for the author to present it as a more compact non-finite or phrasal structure in the English context.

Academic discourse will transform the content that originally appears in the form of process, action and clause into nominalization, phrasing and compression (Halliday, 2008). It also embodies the grammatical metaphor, especially the ideational grammatical metaphor. The original more natural process expression has been transformed into a more academic nominal or compressed expression. For example: common expression: Someone found that.../academic expression: the finding that.../we decided to... / academic expression: the decision to...

4.3 Context Adaptation and Author Variation in the Selection of Complement Constructions

There is a stable pattern of stylistic variation in language use (Biber, 1988), in which individual differences can be explained by a probabilistic model (Gries & Ellis, 2015).

There are differences in the choice of constructions between the authors. In the case of following the norms of academic writing, academic writing has a personal style. Some Chinese scholars have long preferred some shell noun structures, such as: However, this cannot be regarded as an individual language feature, but as a dynamic choice under the joint action of the author's habits and context requirements.

We summarize three types of authorship patterns: 1) stable; 2) Adaptive type; 3) Mixed type

Stable authors have a stable preference. Regardless of language changes, they are highly dependent on clausal complement, which may be related to long-term academic training, fixed writing habits, academic community norms, etc. Adaptive authors adjust resources according to the norms of English journals, which is mainly reflected in the obvious increase in the use of verbal complement after entering the context of English publication. Mixed authors account for the largest proportion, which is reflected in the flexible adjustment of more types of construction resources.

5. Limitations

This study focuses on comparison and controls the factors of author source, journal source and publishing language, but the condition of publishing ecology (such as reader expectation, review regulation, academic discourse tradition, etc.) cannot be separated from the study. However, this also shows that bilingual academic writing is the result of the author's multiple considerations, and it is also the author's discourse reconstruction in two different knowledge production fields. Future research can focus on the above aspects, combined with the author's interviews to further enhance the persuasiveness.

In view of the fact that this paper only selects 100 papers published in English and Chinese authoritative journals of applied linguistics by the same group of authors, future research can further expand the author scale, subject distribution and genre type to support the research findings of this paper.

In order to meet the needs of research, we refer to previous studies to classify English and Chinese shell noun complement constructions to a higher level. Although this paper enhances the operability of cross-language comparison through the merger of construction family level, the boundary of eventive family is relatively narrow, and there may still be different explanatory spaces for the attribution of some Chinese relational realizations. Future research can further test and revise the relevant classification framework in combination with more detailed syntactic and semantic standards. The zero frequency of verbal complement and the low frequency of eventive complement in Chinese corpus indicate that the distribution of different construction families in the data is not balanced. On the one hand, this imbalance itself constitutes the research findings of this paper. On the other hand, it also limits statistical modeling. For example, ordinary logistic regression will have a complete separation problem in the verbal category. This paper has dealt with this problem through Firth penalty logistic regression, but the relevant results are still the estimation results in the current sample. Future research can further improve the ability to test low-frequency constructions by expanding the scale of corpus or adding more marginal examples.

6. Conclusion

This study takes the authoritative English and Chinese journal articles published by the same group of Chinese authors as the research object, and examines the systematic characteristics of the authors' choice of shell noun complement constructions in different top-publication contexts. The research uses TreeTagger, CorpusWordParser, AntConc and other tools to complete corpus preprocessing and structure retrieval. Combined with manual verification and functional annotation, the shell noun complement constructions in English and Chinese materials are classified into three types of superordinate constructions: clausal complement, verbal complement and eventive complement. On this basis, this paper uses R Studio for statistical analysis to better describe the relationship between linguistic context and author's choice of construction. Specifically, it includes: 1) the frequency and proportion distribution of the three types of constructions in English and Chinese are presented by descriptive statistics; 2) Use contingency tables to analyze the overall distribution differences, including Fisher's exact test, chi-square test and effect size analysis; 3) Regression modeling was used to examine the language effects and author differences.

The results show that the distribution of shell noun complement constructions by Chinese authors in English and Chinese publishing contexts is not completely consistent. On the whole, clausal complement is the main construction type in both English and Chinese, which shows that clause complement is an important resource for Chinese writers to organize the semantic content of shell nouns and realize the concretization of propositions. However, the concentration of clausal complement in Chinese corpus is significantly higher, indicating that Chinese authors are more inclined to use the clause structure to directly expand the content of shell nouns in Chinese academic writing. In contrast, although the English corpus is also dominated by clausal complement, the proportion of verbal complement is significantly increased, indicating that the author will be more active in mobilizing verbal complement resources such as infinitives and gerunds after entering the context of English publication, so as to adapt to the more diversified ways of noun complement in English academic discourse.

Further statistical tests show that there are significant differences in the overall distribution of shell noun complement construction families in English and Chinese. Among them, clausal complement shows stronger concentration in Chinese corpus, verbal complement mainly appears in English corpus, while no such construction examples are observed in Chinese corpus. The difference of eventive complement between English and Chinese is not significant. The results of category-by-category modeling further show that the influence of linguistic context on clausal complement and verbal complement still exists after controlling the individual differences of authors, indicating that these differences cannot be simply attributed to the fixed writing style of individual authors, but are more likely to be related to the language system, genre norms and academic expression practices in different publishing contexts. It can be seen that the shell noun complement construction preference of Chinese

authors is not a completely stable personal style, but is shaped by the published context to a certain extent. The same group of authors use more clausal complement in Chinese papers, while in English papers, they will redistribute complement resources and increase the use of verbal complement. This shows that Chinese writers do not mechanically copy the mother tongue construction model in cross-language academic writing, but will re-schedule and adapt the available construction resources according to the requirements of the target language and the publishing context.

The significance of this study is that it supplements the deficiency of previous shell noun studies that mainly focus on single language, single genre or comparison of different author groups from the perspective of “cross-context reorganization of the same author”. The results show that the shell noun complement construction is not only a syntactic resource for semantic complement and information packaging, but also an important window to observe the adaptation mechanism of Chinese authors’ academic writing. By comparing the distribution of constructions in English and Chinese authoritative journal articles, this paper further reveals the interactive relationship among author preference, language structure and publishing context.

References

- Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press.
- Bybee, J. (2010). *Language, Usage and Cognition*. Cambridge University Press.
- Biber, D. (1999). A register perspective on grammar and discourse: Variability in the form and use of English complement clauses. *Discourse studies*, 12, 131-150.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Longman.
- Biber, D. (1988). *Variation Across Speech and Writing*. Cambridge University Press.
- Charles, M. (2007). Argument or evidence? Disciplinary variation in the use of the noun that pattern in stance construction. *English for Specific Purposes*, 26(2), 203-218.
- Flowerdew, J., & Forest, R. W. (2015). *Signalling nouns in English: A corpus-based discourse approach*. Cambridge: Cambridge University Press.
- Flowerdew, J. (2003). Signalling nouns in discourse. *English language Teaching Journal*, 57(3), 226-233.
- Fillmore, C. J., Kay, P., & O’Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language*, 64(3), 501-538.
- Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford University Press.
- Gries, S. T., & Ellis, N. C. (2015). Statistical measures for usage-based linguistics. *Language Learning*, 65(1), 228-255.

- Huang, L., & Deng, J. (2025). "This dissertation intricately explores...": ChatGPT's shell noun use in rephrasing dissertation abstracts. *System*, 129(2025).
- Hu, X., & Huang, Y. (2021). A study of move differences in the interpersonal functions of shell nouns in English abstracts of Chinese and international scientific papers. *Foreign Language Learning Theory and Practice*, (3), 56-63.
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purpose*, 27(1), 4-21.
- Hyland, K. (2005). *Metadiscourse: Exploring Interaction in Writing*. Continuum.
- Halliday, M. A. K., Matthiessen, C. M. I. M. (2004). *An Introduction to Functional Grammar* (3rd ed.). Arnold, London.
- Hu, J., Pan, H., & Xu, L. (2001). Is there a finite vs. nonfinite distinction in Chinese? *Linguistics*, 39, 1117-1148.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Jiang, F., & Hyland, K. (2016). Nouns and academic interactions: A neglected feature of metadiscourse. *Applied Linguistics*, 1-23.
- Jiang, F. (2016). Stance construction and interpersonal functions of shell nouns. *Modern Foreign Languages*, 39(4), 470-482+583.
- Jiang, F., & Hyland, K. (2015). The fact that: Stance nouns in academic writing. *English for Specific Purposes*, 40, 22-34.
- Kou, X. (2023). A study of presupposition types and pragmatic functions of shell nouns. *Chinese Teaching in the World*, 37(1), 57-70.
- Lou, B. (2013). A corpus-based analysis of shell noun use in postgraduate academic English discourse. *Foreign Language Education*, 34(3), 46-49+53.
- Langacker, R.W. (1987). *Foundations of Cognitive Grammar, Volume 1: Theoretical Prerequisites*. Stanford University Press.
- Qiu, X., Dong, M., & Liu, W. (2021). Cognitive functions of shell nouns and knowledge construction in science and engineering academic discourse. *Foreign Languages and Their Teaching*, (6), 23-33+72+147-148.
- Shi, Y. (2023). *The evolution of Chinese grammar*. Cambridge University Press.
- Shi, C., & Cai, S. (2022). An analysis of theoretical issues in construction grammar research. *Foreign Language Teaching and Research*, 54(5), 643-655.
- Schmid, H.-J. (2000). *English abstract nouns as conceptual shells: From corpus to cognition*. De Gruyter Mouton.
- Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.

- Tang, Mengmeng. (2023). The Distinction between Finite and Nonfinite Verbs in Mandarin Chinese. *The Educational Review*, 78, 1233-1243.
- Yang, B. (2022). Non-finiteness in the literature. In *Non-finiteness: A process-relation perspective* (pp. 9-67). Cambridge University Press.
- Zhu, Y., & Huang, C.-R. (2023). *A student grammar of Chinese*. Cambridge University Press.