*Original Paper*

# Shell Nouns as Resources for Cohesion and Stance in Master's Thesis Abstracts: A Corpus-based Contrastive Study

Yinghui Li[1]

[1] School of Foreign Languages, Zhejiang Gongshang University, Hangzhou, China

***Abstract***

*Academic writing relies heavily on abstract nouns to structure information and signal stance. This study adopts a corpus-based, contrastive approach to examine the use of shell nouns in English abstracts of Chinese and Western master's theses in education. High-frequency shell nouns were analyzed in terms of their frequency, lexico-grammatical patterns, and discourse functions. Results show that Chinese learners use shell nouns more extensively and across a broader lexical range, while native speakers rely on a smaller, more conventional set. Both groups favor clause-based constructions, but Chinese students prefer more explicit clause expansion, whereas native speakers use demonstrative constructions to enhance cohesion. Functionally, shell nouns serve semantic, cohesive, and evaluative purposes. Epistemic meanings dominate both corpora, reflecting the objective style of abstracts. Chinese learners tend to use more attitudinal nouns, emphasizing procedural clarity, whereas native speakers favor deontic nouns, highlighting reliability and a balanced stance. The study contributes to research on shell nouns and L2 academic writing, offering pedagogical implications for postgraduate writing instruction in education.*

***Keywords***

*shell nouns, academic writing, lexico-grammatical patterns, corpus analysis*

## 1. Introduction

Academic writing is widely recognized for its dense, information-rich style and its strong reliance on nouns. Corpus-based evidence shows that academic prose contains approximately 300,000 nouns per million words, compared with about 150,000 in conversation (Biber et al., 1999). This preference for nouns is one way in which "scholarly communication distinguishes itself from other domains" (Jiang, 2022). Abstract nouns, in particular, play a central role in packaging complex meanings, among which shell nouns have attracted increasing attention. A shell noun is a semantically general noun that

functions as a "*conceptual shell*" (Schmid, 2000), with its meaning specified by the surrounding context.

As readers' first encounter with a text, abstracts contain rich interactional features that reveal how writers position themselves within disciplinary communities (Hyland, 2004). While shell nouns have been widely studied in expert academic writing, far less attention has been paid to their use by novice writers in thesis abstracts, particularly in the discipline of education. Master's thesis abstracts, produced under strict pedagogical and institutional constraints, require students to concisely convey research objectives, methods, and findings. The choice and grammatical realization of shell nouns can illuminate how novice writers organize information and signal stance, making English abstracts a key site for observing the development of academic writing skills in education. Using a corpus-based and contrastive approach, this study analyzes the frequency, lexico-grammatical patterns, and discourse functions of shell nouns. By revealing cross-cultural similarities and differences, the study extends current knowledge of academic writing practices and offers pedagogical insights for academic writing instruction in education.

*1.1 Shell Nouns*

Various definitions have been proposed to capture the nature of shell nouns, which are also referred to in the literature as *container nouns* (Vendler, 1968, 2019), *anaphoric nouns* (Francis, 1986), *carrier nouns* (Ivanicˇ, 1991), *signaling nouns* (Flowerdew & Forest, 2015), and *metadiscursive nouns* (Jiang & Hyland, 2018). Despite this terminological variation, scholars largely agree that these items constitute a distinctive class of abstract nouns that plays a crucial role in discourse organization.

Schmid (2000, p.4) defines shell nouns as an open, functionally defined class of abstract nouns that can encapsulate complex, proposition-like content and temporarily function as conceptual "containers" in discourse. From a similar perspective, Flowerdew(2003)emphasizes their context-dependent nature, noting that their specific meanings are not inherent but become fully interpretable through the surrounding co-text. Previous studies converge on the view that shell nouns retain a relatively general semantic core while those requiring contextual specification for interpretation (Flowerdew, 2003; Flowerdew & Forest, 2015; Ivanič, 1991; Schmid, 2000). As Winter (1977, p. 87; 2020) observes, their meanings remain incomplete until specified by lexical realizations in the immediate context (Winter, 1977, p. 87; 2020, p.10). On this basis, Schmid (2000) further proposes a semantic classification of shell nouns (see Table1).

**Table 1. Semantic Classification of Shell Nouns (Schmid, 2000)**

| Class | Examples | | | | | |
|---|---|---|---|---|---|---|
| Factual | *fact* | *thing* | *point* | *problem* | *reason* | *difference* |
| Linguistic | *news* | *message* | *report* | *order* | *proposal* | *rumour* |
| Mental | *idea* | *notion* | *belief* | *aim* | *plan* | *assumption* |

| Modal | *truth* | *possibility* | *need* | *ability* | *reality* | *obligation* |
| Eventive | *act* | *move* | *measure* | *reaction* | *tradition* | *attempt* |
| Circumstantial | *situation* | *context* | *place* | *area* | *way* | *approach* |

This semantic flexibility enables shell nouns to perform a metadiscursive function in academic discourse. At the textual level, shell nouns enhance cohesion by linking clauses and larger discourse units. They operate either cataphorically to project forthcoming information or anaphorically to summarize preceding propositions (Flowerdew, 2003; Flowerdew & Forest, 2015; Jiang, 2022; Jiang & Hyland, 2021; Schmid, 2000). At the interpersonal level, they also serve as stance markers, allowing writers to express evaluative positions toward the propositional content they package (Charles, 2007).

Schmid (2000) further notes that shell nouns tend to recur in a limited set of lexico-grammatical patterns, within which their meanings are jointly constructed with the surrounding discourse. In academic texts, these patterns play a key role in maintaining textual cohesion (Francis, 2002; Halliday & Hasan, 1976). Specifically, cataphoric reference is commonly realized through the N + clause (N-*cl*) and N + *be* + clause (N-*be-cl*) constructions, whereas anaphoric reference typically takes the form of demonstrative-led patterns such as *th*-N and *th-be*-N (Schmid, 2000).

Adopting a functional perspective, Jiang and Hyland (2016) analyzed N-cl patterns across six disciplinary corpora and classified shell nouns into three categories, as shown in Table 2.

**Table 2. Interactional Function of Shell Nouns (Jiang &Hyland, 2016)**

| Functional Category | | Description | Examples |
| --- | --- | --- | --- |
| **Entity** | Text | Concrete meta-text | report, paper, extract |
| | Event | events, processes, and evidential cases | change, observation |
| | Discourse | verbal propositions and speech acts | argument, claim |
| | Cognition | cognitive beliefs and attitudes | decision, idea, doubt |
| **Attribute** | Quality | traits that are admired or criticized | advantage, difficulty |
| | Manner | circumstance of actions and state of affairs | time, method, way, extent |
| | Status | epistemic, deontic, and modality | ability, capacity, potential |
| **Relation** | Cause-effect, difference | logical or rhetorical relationships between propositions | reason, result, difference |

*1.2 Shell Nouns in Academic Writing*

Shell nouns have been identified as a key linguistic resource in academic writing, enabling writers to package propositions and express epistemic and evaluative stances (Schmid, 2000; Hyland, 2008).

Earlier studies on anaphoric and signalling nouns (Francis, 1986; Flowerdew, 2003) laid the foundation for understanding how such nouns contribute to discourse cohesion and information organization.

Building on this work, Schmid (2000) systematically categorized shell nouns and demonstrated their multifunctional roles in academic discourse, including semantic characterization, temporary concept formation, and textual linking. Subsequent research has shown that shell nouns function as important stance resources, allowing writers to express evaluation, causal relations, and degrees of certainty while maintaining an impersonal academic style (Charles, 2003, 2007; Jiang & Hyland, 2015, 2018). Given their discourse-organizing function, shell nouns have attracted increasing attention in studies of abstracts. Corpus-based research on article abstracts indicates that shell nouns are closely associated with major rhetorical moves, including stating research purpose, outlining methodology, and presenting findings (Jiang & Hyland, 2017; Alyousef, 2023). These studies highlight the role of shell nouns in achieving cohesion and coherence in highly compressed academic genres. However, existing work has focused largely on abstracts written by expert scholars, leaving the use of shell nouns by novice and second-language writers comparatively limited.

A parallel strand of research has examined the use of shell noun in student and L2 academic writing. Comparative studies suggest that novice and non-native writers tend to rely heavily on nominal structures, including shell nouns, to organize discourse and signal logical relations (Aktas & Cortes, 2008; Benitez-Castro, 2021). Despite these contributions, postgraduate thesis writing has received relatively limited attention, particularly at the level of abstracts.

Two gaps, therefore, remain in the literature. First, master's thesis abstracts, as a transitional genre produced by novice researchers, remain underexplored. Second, few studies have examined shell nouns by integrating semantic classification, lexico-grammatical realization, and stance-related functions within a single analytical framework, especially from a cross-cultural perspective.

## 2. Method

### 2.1 Corpora used in the Study

The study constructed two corpora of English master's thesis abstracts in education: a Learner Corpus (LC) and a Native Corpus (NC), in order to examine cross-cultural variation in shell noun use at the postgraduate level. The LC was compiled from English abstracts of master's theses retrieved from the China National Knowledge Infrastructure (CNKI), the largest and most comprehensive academic database in China. The NC was built from ProQuest Dissertations and Theses Global.

Abstracts were selected based on the following criteria: (1) all theses belonged to the discipline of education; (2) the abstracts were written in English and available in electronic form; (3) theses were completed between 2018 and 2025 ; (4) abstracts were drawn from multiple universities to minimize potential institutional bias; (5) abstracts that were excessively short, incomplete, or highly formulaic were excluded;(6)abstracts authored by students with overseas study experience were excluded to ensure greater sample homogeneity.

As abstracts in the NC were shorter on average, a strict one-to-one matching of texts would have resulted in a much smaller total token count for the NC, potentially affecting frequency-based comparisons. To balance corpus size, 200 abstracts were selected for the LC and 350 abstracts for the NC, resulting in comparable overall word counts (LC: 110,885; NC: 113,389; see Table 3).

**Table 3. Description of the Corpora**

| Corpus | No. of texts | Types | Tokens | Average text length |
|--------|--------------|-------|---------|---------------------|
| LC | 200 | 5865 | 110,885 | 554 |
| NC | 350 | 8085 | 113,389 | 324 |

*2.2 Focus of the Study*

This study focuses on the frequency, lexico-grammatical patterns, and discourse functions of selected shell nouns. By comparing the two corpora, it aims to identify cross-cultural similarities and differences in how novice writers organize information and construct stance in thesis abstracts.

Accordingly, the study addresses the following research questions:

(1) What are the overall frequencies and distributional patterns of selected shell nouns in English abstracts of Chinese and Western master's theses in education?

(2) What lexico-grammatical patterns are associated with these shell nouns in the two corpora?

(3) What discourse functions do these lexico-grammatical patterns perform in thesis abstracts, and how do they differ between the two corpora?

*2.3 Data Analysis*

The analysis adopted a multi-step, corpus-based procedure.

Step 1: All abstracts were POS-tagged using TreeTagger (Schmid, 1994), a tool with an accuracy exceeding 96% (Schmid, 1999). The tagged texts were then imported into AntConc 4.3.1 (Anthony, 2023) to retrieve instances of the lexico-grammatical patterns associated with shell nouns proposed by Schmid (2000). Frequency lists of abstract head nouns were generated for each pattern, from which nouns with a high likelihood of shell noun function were selected as candidates. These lists were consolidated in Excel, normalized frequencies were calculated, and the top 35 potential shell nouns were identified for each corpus.

Step 2: All occurrences of the selected candidate shell nouns were manually coded according to Schmid's (2000) four pattern types. Additional constructions not fully captured by the original framework were identified and incorporated, forming a revised set of patterns (see Table 4). Frequencies of each pattern were then calculated and compared across the two corpora, producing frequency lists of attested shell nouns.

**Table 4. Revised Lexico-Grammatical Patterns**

| No | Construction | Abbreviation | Regular Expressions |
|----|--------------|--------------|---------------------|
| 1 | **SN + postnominal clause** (N-cl)<br>Variants:<br>*to* infinitive-clause | N-*to-do* | \S+_NN(S)\s(\S+_IN\s\S+_DT\s\S+_NN\s)*\S+_TO\s(\S+_RB\s)*\S+_VV\s |
| | *that/wh*-clause | N-*that/where/wh*<br>*en/how* | \S+_NN(S)\s(\S+_IN\s\S+_DT\s\S+_NN\s)*that/That/where/Where/when/When/how/How |
| | of prepositional clause | N-*of-doing* | \S+_NN(S)\sof_IN\s(\S+_RB\s)*(\S+_VBG\s\|\S+_VVG\s\|\S+_VHG\s) |
| 2 | **SN + be/verb +complementing clause**(N-be-cl)<br>Variants:<br>*that*-clause | N-*be/v-that* | \S+_NN(S)\s(\S+_IN\s\S+_DT\s\S+_NN\s)*(\S+_MD\s)*\S+_VB\w*\s(\S+_RB \s)*that |
| | *to* infinitive-clause | N-*be/v-to/doing* | \S+_NN(S)\s(\S+_IN\s\S+_DT\s\S+_NN\s)*(\S+_MD\s)*\S+_VB\w*\s(\S+_RB \s)*\S+_TO\s |
| 3 | **Demonstrative +SN**<br>(*this/that/these/those/the/*<br>*such*, etc-N) | *th*-N | this/This/that/That/these/These/those/Those/the/The_DT\s(\S+_RB\w*\s)*(\S+_JJ\w*\s)*\S+_NN(S)\s<br>such/Such_JJ\s(\S+_RB\w*\s)*(\S+_JJ\w*\s)*\S+_NN(S)\s<br>such/Such_PDT\s\S+_DT\s(\S+_JJ\s)*\S+_NN(S)\s |
| 4 | **Demonstrative + be + SN**<br>(*this/that/these/those/the/*<br>*such*,etc-*be*-N) | *th-be*-N | this/This/that/That/these/These/those/Those_DT\s(\S+_MD\s)*\S+_VB\w*\s(\S+_DT\s)*(\S+_JJ\w*\s)*\S+_NN(S)\s<br>such/Such_JJ\s(\S+_MD\s)*\S+_VB\w*\s(\S+_DT\s)*(\S+_JJ\w*\s)*\S+_NN\s<br>such/Such_PDT\s(\S+_MD\s)*\S+_VB\w*\s(\S+_DT\s)*(\S+_JJ\w*\s)*\S+_NN \s |
| 5 | ***There be***+SN | *there be*-N | there/There_EX\s(\S+_MD\s)*\S+_VB\w*\s(\S+_DT\s)*(\S+_JJ\w*\s)*\S+_NN(S)\s<br>there/There_EX\s(\S+_MD\s)*\S+_VB\w*\sthe/DT\s(\S+_JJ\w*\s)*\S+_NN(S)\s |

| | | | |
|---|---|---|---|
| 6 | **SN +Enumerative phrases** (*like*, *such as*,etc.) | N-EP | \S+_NN(S)\s+(like/IN\|such/JJ\s+as/IN)\s+(\S+_D T\s+)*(\S+_JJ\w*\s+)*\S+_NN(S) <br> \S+_NN(S)\s+for/IN\s+example/NN <br> \S+_NN(S)\s+for/IN\s+instance/NN |
| 7 | **Fixed phrases** | *in fact, as a result*, etc. | as/As/IN\s+a/DT\s+(result/NN\|consequence/NN) <br> in/IN\s+(fact/NN\|general/JJ\|particular/JJ) |

Step 3: Candidate nouns were ranked by frequency, and the top 20 shell nouns in each corpus were selected for further analysis. High-frequency shell nouns capture recurrent patterns of use and represent central resources for discourse organization as well as the realization of semantic, cohesive, and evaluative functions, in line with established practice in corpus-based research on abstract nouns and stance (Charles, 2007; Jiang & Hyland, 2017).

## 3. Result

### 3.1 High-frequency shell nouns in LC and NC

A total of 11,726 shell nouns were identified across the 550 abstracts, with 6,673 occurrences in the LC and 5,053 in the NC. After normalization, the LC yielded 602 shell nouns per 10,000 words (mean = 33 tokens per abstract), whereas the NC yielded 444 per 10,000 words (mean = 14 tokens per abstract). Overall, shell nouns occurred more frequently in LC abstracts than in NC abstracts.

The top 20 shell nouns in each corpus accounted for the majority of all shell noun occurrences, as shown in Table 5.

**Table 5. The Top Twenty Shell Nouns in LC and NC**

| Rank | LC | | | | NC | | | |
|---|---|---|---|---|---|---|---|---|
| | Word | Raw freq | % of total | Functional Category | Word | Raw freq | % of total | Functional Category |
| 1 | **research** | 693 | 10.38 | Entity | **study** | 1057 | 20.92 | Entity |
| 2 | **study** | 523 | 7.83 | Entity | **research** | 566 | 11.21 | Entity |
| 3 | quality | 382 | 5.72 | Attribute | **analysis** | 240 | 4.75 | Entity |
| 4 | ability | 363 | 5.44 | Attribute | **use** | 195 | 3.86 | Relation |
| 5 | **analysis** | 318 | 4.76 | Entity | **design** | 173 | 3.43 | Entity |
| 6 | **theory** | 251 | 3.76 | Entity | approach | 139 | 2.75 | Entity |
| 7 | system | 240 | 3.60 | Relation | **theory** | 137 | 2.71 | Entity |
| 8 | method | 238 | 3.57 | Entity | **results** | 135 | 2.67 | Entity |
| 9 | **design** | 226 | 3.39 | Entity | impact | 127 | 2.52 | Attribute |

| 10 | **process** | 214 | 3.21 | Entity | case | 126 | 2.49 | Entity |
| 11 | **results** | 185 | 2.77 | Entity | context | 121 | 2.40 | Entity |
| 12 | **model** | 172 | 2.58 | Entity | technology | 111 | 2.20 | Entity |
| 13 | situation | 167 | 2.50 | Entity | **process** | 118 | 2.34 | Entity |
| 14 | **use** | 147 | 2.20 | Relation | focus | 108 | 2.14 | Relation |
| 15 | characteristics | 139 | 2.08 | Attribute | purpose | 105 | 2.08 | Entity |
| 16 | effect | 138 | 2.07 | Attribute | need | 94 | 1.86 | Relation |
| 17 | **time** | 130 | 1.95 | Entity | **time** | 90 | 1.78 | Entity |
| 18 | application | 119 | 1.78 | Relation | **model** | 82 | 1.62 | Entity |
| 19 | value | 105 | 1.57 | Attribute | change | 74 | 1.46 | Relation |
| 20 | objectives | 104 | 1.56 | Entity | potential | 56 | 1.11 | Attribute |

*Note. Bold items indicate co-occurring words in both corpora.*

Table 5 presents the top 20 most frequent shell nouns in the LC and NC, with ten items shared by both corpora. *Study* and *research* rank as the most frequent nouns in each corpus, with *study* occurring more frequently in the NC (20.92%) than in the LC (7.83%). Overall, the LC displays a wider range of high-frequency nouns, including *ability*, *quality*, and *method.* By contrast, the NC shows a sharper drop in frequency after the top two nouns, suggesting a more concentrated lexical distribution and stronger usage preferences among native English writers.

The functional classification of shell nouns reveals systematic differences in stance realization across the two corpora. In the LC, entity nouns such as *research*, *study*, and *analysis* dominate, whereas attribute nouns (*ability*, *quality*, *value*) and relation nouns (*system*, *use*, *application*) occur less frequently.  By contrast, the NC shows a more balanced distribution across entity, attribute, and relation categories. Alongside high-frequency entity nouns (*study*, *research*, *analysis*), items such as *case*, *context*, and *process* co-occur with evaluative and relational nouns, including *impact*, *purpose*, and *change*. These patterns indicate that Chinese postgraduate writers tend to draw on a wider set of nouns to encode research content, whereas native writers rely on a more compact set of high-frequency nouns while achieving a relatively even distribution of evaluative and relational meanings.

Table 6 presents the raw and normalized frequencies of the top 20 shell nouns in each corpus, together with the log-likelihood test results.

**Table 6. Frequencies and Log-Likelihood of High-Frequency Shell Nouns**

| Parameter | LC | | NC | | LL | p-value |
|---|---|---|---|---|---|---|
| | Raw freq | Norm. freq | Raw freq | Norm. freq | (G²) | |
| Token count | 6673 | 602 | 5053 | 446 | 269.62 | <0.001 |

Notes.

LC = Learner Corpus; NC = Native-speaker Corpus

Raw freq = raw frequency of shell nouns

Norm. freq = normalized frequency per 10,000 tokens

LL (G²) = log-likelihood value

p = significance level.

As shown in Table 6, shell nouns occur more frequently in the LC than in the NC, with the former registering 6,673 tokens (601.8 per 10,000 tokens) and the latter 5,053 tokens (445.7 per 10,000 tokens). The log-likelihood test ($G^2 = 269.62$, $p < 0.001$) confirms that this difference is statistically significant, indicating that there is a markedly higher overall use of shell nouns in learner-produced abstracts than in native-speaker texts. This pattern may reflect systematic differences in the deployment of shell nouns as discourse-organizing and stance-marking resources across the two corpora. Given the concise and information-dense nature of educational abstracts, shell nouns appear to function as an efficient linguistic device for condensing complex propositional content while maintaining textual cohesion.

*3.2 Lexico-grammatical Patterns of High-frequency Shell Nouns*

Drawing on the lexico-grammatical classification in Table 5, the frequency distribution of the ten high-frequency co-occurring shell nouns was examined across constructions (Figure1). All relevant instances were retrieved using AntConc, and concordance lines were manually analyzed to identify their typical lexico-grammatical realizations.
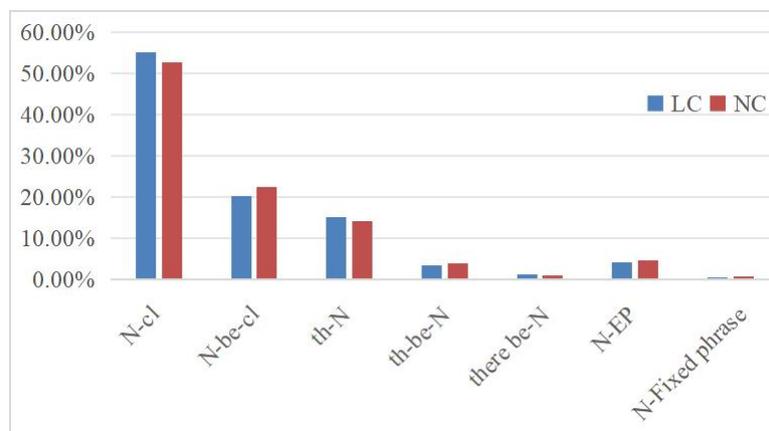
**Figure 1. Frequency of Lexico-Grammatical Patterns of High-Frequency Shell Nouns in LC and NC**

As shown in Figure 1, the overall distribution of lexico-grammatical patterns is highly similar across the two corpora. In both corpora, the N-*cl* is the most frequently used construction, accounting for over half of all occurrences of the ten shell nouns. The N-*be-cl* ranks second, around 20%, followed by *th*-N constructions, which are slightly more frequent among native English writers. Other patterns, including *th-be*-N, *there be*-N, N-EP, and N-Fixed phrase, are marginal, each constituting less than 5% of the total.

With regard to the N-*be/v*-cl construction, both groups of writers show a strong preference for the shell noun *study*. This tendency can be largely attributed to the recurrent use of formulaic structures such as *the study found/suggests/indicates that*, which are commonly employed in thesis abstracts to present a generalized characterization of the research aims or its findings.

The N–EP pattern is an additional pattern identified in the present study. Fixed expressions such as *like*, which introduce subsequent noun phrases for exemplification, are classified here as enumerative phrases. Within the abstracts examined, these phrases function to specify or elaborate the referential content of the preceding shell noun, thereby enhancing textual cohesion. Among the ten high-frequency shell nouns, *model* occurs most frequently in this construction, typically co-occurring with enumerative markers such as *such as* to introduce specific categories of educational models.

Table 7 further shows that the ten overlapped shell nouns exhibit differentiated distributional preferences across lexico-grammatical patterns in the LC and NC.

**Table 7. Frequency of Overlapped Shell Nouns in Lexico-grammatical Patterns (Per 10,000)**

| SN | N-*cl* | | N-*be-cl* | | *th*-N | | *th-be*-N | | *there be*-N | | N-EP | | Fixed phrases | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | LC | NC | LC | NC | LC | NC | LC | NC | LC | NC | LC | NC | LC | NC |
| research | 65.40 | 62.70 | 18.50 | 20.10 | 12.30 | 11.50 | 0.00 | 0.00 | 0.00 | 0.00 | 3.50 | 4.00 | 0.30 | 0.70 |
| study | 70.25 | 68.10 | 15.30 | 18.20 | 10.15 | 9.80 | 0.00 | 0.00 | 0.00 | 0.00 | 3.80 | 4.20 | 0.50 | 0.60 |

| analysis | 0.94 | 1.69 | 2.82 | 3.37 | 11.27 | 10.11 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| theory | 45.09 | 20.23 | 0.00 | 2.53 | 2.82 | 3.37 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| design | 50.15 | 48.30 | 12.20 | 14.50 | 8.75 | 9.20 | 0.00 | 0.00 | 0.00 | 0.00 | 2.80 | 3.10 | 0.40 | 0.50 |
| process | 63.88 | 19.39 | 0.94 | 0.00 | 14.09 | 5.90 | 0.00 | 0.00 | 0.00 | 0.00 | 0.94 | 0.00 | 0.00 | 0.00 |
| results | 0.94 | 0.00 | 17.85 | 27.81 | 2.82 | 15.18 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 4.70 | 7.59 |
| model | 35.60 | 33.80 | 8.40 | 10.10 | 6.20 | 7.30 | 0.00 | 0.00 | 0.00 | 0.00 | 4.00 | 4.50 | 0.20 | 0.60 |
| use | 28.50 | 26.70 | 5.20 | 6.80 | 10.80 | 12.10 | 0.00 | 0.00 | 0.00 | 0.00 | 3.20 | 3.70 | 0.30 | 0.50 |
| time | 2.82 | 5.06 | 0.94 | 0.00 | 16.91 | 6.74 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.84 | 0.00 | 0.00 |

*3.3 Functional Analysis of High-frequency Shell Nouns*

3.3.1 Semantic Function in Lexico-grammatical Patterns

Schmid(2000) argues that shell nouns play a central role in semantic characterization. The semantic function of characterization is to define writers' opinions and attitudes towards shell contents, and there are two ways to realize this function. One is the choice of shell nouns, and the other is the selection of the premodifiers of shell nouns.

On the basis of semantic meanings, Schmid (2000) classifies shell nouns into six categories: factual, linguistic, mental, modal, eventive and circumstantial. Nouns construing objective states of affairs were categorized as factual; those related to written or spoken language as linguistic; those expressing beliefs, viewpoints, expectations, or cognitive processes as mental; and modal shell nouns express meanings associated with possibility, ability, or necessity. In addition, eventive shell nouns denote actions, processes, or changes of state, while those associated with location, setting, or contextual circumstances were categorized as circumstantial.

Figure 2 presents the semantic classification of the 20 high-frequency shell nouns (see Table 6) identified in the two corpora.
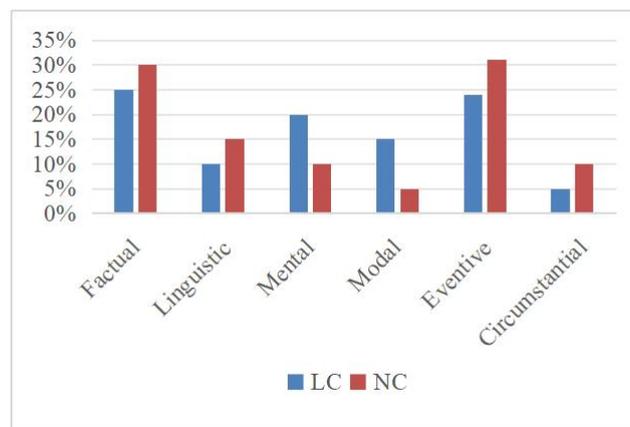


**Figure 2. Comparison of the Semantic Distribution of Top 20 High-frequency Shell Nouns in LC and NC**

In both corpora, factual and eventive shell nouns constitute the largest proportions, jointly accounting for about half of all instances. Specifically, factual nouns represent roughly 25% in the LC and 30% in the NC, while eventive nouns account for around 24% in the LC and 31% in the NC, emphasizing the reporting of findings and description of research processes, which are central to educational research. Compared with natives, Chinese students use more mental and modal nouns, indicating greater focus on evaluation and expressing certainty, whereas natives favor linguistic and circumstantial nouns, highlighting attention to discourse framing and contextual details. These differences may be associated with differences in academic writing conventions, instructional emphases, and rhetorical traditions, with Chinese EFL writers placing greater emphasis on explicit evaluation and result-oriented reporting. Both groups' low use of linguistic, modal, and circumstantial nouns reflects the concise, objective style of abstracts in education. This distribution aligns with the applied and practice-oriented nature of educational research, where research procedures and outcomes are foregrounded over theoretical abstraction.

3.3.2 Cohesive Function in Lexico-grammatical Patterns

Schmid (2000) argues that shell nouns, as an open and functional word class, play a key role in conveying complex propositional information in discourse. He identifies three main functions: semantic characterization, temporary concept formation, and linking(Schmid2000:14). The linking function, also called the cohesive function, refers to the ability of shell nouns to act as discourse connectors. Through either anaphoric or cataphoric reference, shell nouns establish explicit links between clauses and larger discourse units, thereby facilitating information flow and enhancing textual coherence.

Several cases of the cohesive function performed by selected shell nouns in patterns are presented below.

1) N-*cl*

This study which focuses on the middle and high grades English classes studies the current **situation** *of learning opportunities* in primary English classes. (LC)

The conceptual framework was grounded in Engeström's activity **theory** *to allow* an exploration of individual agency within a specific community setting. (NC)

2) N-*be/verb-cl*

The experimental **results** *show that* transforming relevant vocabulary and sentence patterns under different topics in the textbooks into written records can significantly deepen students' ...(LC)

The purpose of this qualitative **case study** *was to explore* how veteran teachers of adult English as a second language met the challenge to transfer to the virtual learning environment.(NC)

3) *th*-N

Domestic studies on critical thinking ability are limited in number and often overlook the evaluation of *these* **ability** among high school students. (LC)

*This* **study** was created to collect and analyze the experiences of first-and second-year English as a Second Language (ESL) college students who attended ESL programs. (NC)

4)     *th-be*-N

The effective way of spiritual connotation is an inevitable requirement for higher vocational colleges to meet the development needs of the times and to train high-quality skilled talents. *This is* also the **significance** and **value** of this topic. (LC)

*This is of* **importance** to school leaders as teachers' perspectives on school climate and self-efficacy as a result of improvement approaches are fundamental in teacher satisfaction...(NC)

5)     *there be*-N

Among the student factors, *there is* a significant positive **impact** between the personal qualities of the student and the student's learning and learning results.(LC)

*There is* a **need** to better understand how recently qualified teachers are integrating technologies into their teaching practice.

6)     N-EP

The higher vocational colleges of most majors have various **problems**, *such as* the results of English language learning are not ideal, they aren't interested in English, and so on.(LC)

In addition, various contextual factors like issues in classroom management, students' lack of motivation, dichotomic and discriminatory relationship...(NC)

7)     N-Fixed phrase

***In fact***, projecting the field into Chinese excellent traditional culture education, especially for English (education) students, is of great significance to strengthening their cultural literacy.(LC)

...it also reduces English exposure for students. ***As a result***, teachers struggled in explaining the lesson, assessing students' tasks and observing students' behaviour ...(NC)

Across patterns such as N-cl, N-*be/verb-cl*, *th*-N, *th-be*-N, *there be*-N, N-EP, and N-fixed phrases, they help introduce concepts, summarize findings, and maintain discourse coherence. In Chinese learners' abstracts, shell nouns mainly highlight the research focus and outcomes, guiding the reader through what was done or observed. In native abstracts, the same nouns are used more flexibly, embedding research goals, theoretical context, and evaluative or causal nuance. Overall, shell nouns function as discourse connectors in both corpora. Native authors tend to use them to integrate contextual and interpretive links, while Chinese learners more frequently employ them to foreground research focus and outcomes, reflecting different rhetorical orientations in abstract writing.

3.3.3 Evaluative Function in Lexico-grammatical Patterns

In academic writing, authors do more than present empirical facts; they also position their claims and guide readers' interpretations through stance construction. Long and Xu (2010) distinguish four types of stance markers: epistemic markers, which express degrees of certainty or evidential support (e.g., *fact*, *evidence*); deontic markers, which indicate necessity, obligation, or causal relations (e.g., *necessity*, *influence*); attitudinal markers, which convey evaluation or judgment (e.g., *preference*, *importance*);

109

and textual markers, which primarily primarily organize discourse. As shell nouns inherently perform a discourse-organizing role, textual markers were excluded here. The analysis, therefore, focuses on the remaining three categories, calculating the proportion of the top 20 high-frequency shell nouns functioning as epistemic, deontic, and attitudinal stance markers (see Figure 3).
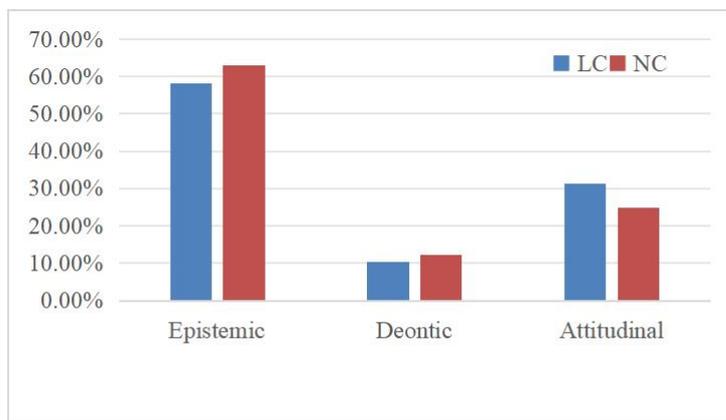


**Figure 3. Comparison of the Evaluative Function Distribution of High-frequency Shell Nouns in LC and NC**

Figure 3 shows that epistemic shell nouns are the most frequently used stance resources in both corpora, whereas attitudinal shell nouns occur at relatively low frequencies. This distribution aligns with the genre conventions of educational research abstracts, which favor a neutral and formal register and limit explicit evaluative expression. Epistemic nouns thus serve to emphasize the certainty and reliability of research findings, which is particularly important in reporting educational interventions or classroom studies. Chinese students make more frequent use of deontic nouns, especially those indicating procedural necessity and methodological steps(*method*, *system*, *application*). This pattern suggests a greater emphasis on explicit procedural structuring in LC abstracts, whereas NC abstracts display a more balanced distribution of epistemic, evaluative, and procedural meanings. More flexible use of epistemic shell nouns—particularly in clauses characterizing findings and implications—may enable student writers in education to balance procedural clarity with evaluative positioning in thesis abstracts.

## 4. Discussion

This study examines the use of shell nouns in English abstracts of Chinese and Western master's theses in education. Adopting a corpus-based approach, it compares their frequencies, lexico-grammatical patterns, and discourse functions to reveal similarities and differences in academic stance construction. The findings show a notable difference in overall frequency. Chinese EFL writers use shell nouns more frequently and across a wider lexical range, whereas native writers rely on a smaller and more conventionalized set. This contrast suggests that Chinese learners depend more heavily on shell nouns to organize content and signal stance. Both corpora show a strong preference for clause-based constructions, especially N-*cl* and N-*be/verb-cl*. This indicates a tendency toward explicit elaboration

110

rather than compressed discourse organization. Three core functions of shell nouns are identified: semantic, cohesive, and evaluative. Semantically, both corpora mainly rely on Factual and Eventive nouns, consistent with the abstract genre's emphasis on research content and procedures. The LC shows a higher proportion of Mental and Modal nouns, while the NC features more Linguistic and Circumstantial nouns. Cohesively, shell nouns serve as key discourse connectors across patterns. Chinese writers place greater emphasis on factual reporting and outcomes, and native writers integrate contextual information more flexibly. Evaluatively, epistemic shell nouns dominate in both corpora, reflecting the objective stance of academic abstracts; however, the LC exhibits relatively greater use of attitudinal nouns, whereas the NC favors deontic nouns, indicating distinct strategies for expressing evaluation and obligation.

Overall, this study offers a systematic account of shell noun use in English abstracts of educational master's theses and demonstrates how lexical and functional choices jointly shape academic stance. The findings suggest that effective abstract writing requires a careful balance between content packaging, procedural description, and evaluative positioning. Pedagogically, instruction can guide students to select appropriate shell nouns, construct clear anaphoric and cataphoric links, and use epistemic, deontic, and attitudinal expressions effectively, enhancing the clarity, coherence, and disciplinary alignment of their academic writing. Raising students' awareness of shell noun functions can help Chinese learners structure abstracts more effectively, balance certainty with procedural explanation, and develop more native-like academic writing skills in education.

Several limitations should be acknowledged. First, this analysis focused only on the top 20 high-frequency shell nouns; further studies could extend the scope to include a larger set of items to better capture broader usage tendencies. Second, while this study concentrates on lexico-grammatical patterns and three core discourse functions, other dimensions, such as pragmatics, cognitive processes, and disciplinary variation, remain to be explored. Future research adopting multi-dimensional approaches would contribute to a more comprehensive understanding of shell noun use in academic writing.

## References

Aktas, R. N., & Cortes, V. (2008). Shell nouns as cohesive devices in published and ESL student writing. *Journal of English for Academic Purposes*, *7*, 3-14. https://doi.org/10.1016/j.jeap.2008.02.002

Charles, M. (2003). 'This mystery': A corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines. *Journal of English for Academic Purposes*, *2*(4), 313-326.

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.

Dong, M., Fang, A. C., & Qiu, X. (2020). Shell nouns as grammatical metaphor in knowledge construal: Variation across science and engineering discourse. *Lingua*, *248*, Article 102946.

Dos Santos, M. B. (1996). The textual organization of research paper abstracts in applied linguistics. *Text & Talk*, *16*(4), 481-500.

Floerdew, J. (2003). Signalling nouns in discourse.*English for Specific Purposes*, *22*(4), 329-346.

Flowerdew, J., & Forest, R. (2015). *Signalling nouns in English: A corpus-based discourse approach*. Cambridge: Cambridge University Press.

Francis, G. (1986). *Anaphoric nouns*. Birmingham: English Language Research.

Gray, B. (2010). On the use of demonstrative pronouns and determiners as cohesive devices: A focus on sentenceinitial this/these in academic prose. *Journal of English for Academic Purposes*, *9*, 167-183. https://doi.org/ 10.1016/j.jeap.2009.11.003

Gray, B., & Cortes, V. (2011). Perception vs. evidence An analysis of this and these in academic prose. *Journal for Specific Purposes*, *30*(1), 31-43. https://doi.org/10.1016/j.esp.2010.06.004

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman Group Limited.

Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Hyland, K. (2005b). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, *7*(2), 173-192.

Ivaniˇc, R. (1991). Nouns in search of a context: A study of nouns with both open-system and closed-system characteristic. *IRAL – International Review of Applied Linguistics in Language Teaching*, *29*(2), 93-114.

Jiang, F. (2017). Stance and voice in academic writing: The "noun + that" construction and disciplinary variation [J] . *International Journal of Corpus Linguistics*, *22*(1), 85-106.

Jiang, F., & Hyland, K. (2015). "The fact that": Stance nouns in disciplinary writing. *Discourse Studies*, *17*(5), 529-550.

Jiang, F., & Hyland, K. (2017). Metadiscursive nouns: Interaction and cohesion in abstract moves. *English for Specific Purposes*, *46*, 1-14.

Schmid, H. (1994). TreeTagger. https://www.cis.lmu.de/~schmid/tools/TreeTagger/. February 18th, 2024.

Schmid, H.-J. (2000). *English abstract nouns as conceptual shells: From corpus to cognition*. Berlin; New York: Mouton de Gruyter.