

A Comparative Study of Shell Nouns in the Abstracts of Chinese Master's Theses and Research Articles by English Expert Writers

Zihan Wu

Northwest Normal University, College of Foreign Languages and Literature, Gansu, Lanzhou, 730070, China

Abstract

This study compares the use of shell nouns in Chinese master's thesis abstracts and English expert research article abstracts. Based on Schmid's (2000) definition of shell nouns and Jiang and Hyland's (2015) functional classification, two small-scale corpora were built: a Master's Corpus and an Expert Corpus. The findings reveal three major differences. First, experts use shell nouns significantly more frequently than master's students. Second, experts employ a balanced distribution of lexico-grammatical constructions, with N-that accounting for 18.3%, whereas master's students heavily rely on N-of and underuse N-that. Third, master's students overuse event and manner nouns, while experts excel in using cognition, quality, and state nouns to express stance. These findings suggest pedagogical implications for improving Chinese master's students' academic writing by enhancing N-that construction training and shifting focus from process description to research evaluation.

Keywords

Shell nouns; Abstract; Academic writing.

1. Introduction

In academic English writing, the use of abstract nouns has always been an important topic in discourse analysis and second language acquisition research. As a special category of abstract nouns, shell nouns play a significant role in discourse cohesion and stance construction in academic discourse. Schmid first systematically proposed the concept of shell nouns, defining them as a class of abstract nouns used to convey and represent propositional information, possessing the potential to serve as "conceptual shells" for complex propositional information [1]. The primary function of shell nouns is that they can summarize, package, and convey complex propositional information, thereby helping authors achieve discourse coherence while expressing their attitudes and stances toward the content being presented.

In recent years, with the rapid development of corpus linguistics, scholars at home and abroad have conducted an increasing number of studies on shell nouns. Jiang and Hyland, based on a large-scale academic corpus, systematically classified the stance-expressing functions of shell nouns, dividing them into three major categories: nature, attribute, and relation, further subdivided into eight subcategories: thing, event, discourse, cognition, quality, manner, state, and relation [2]. This classification framework provides an important theoretical foundation for cross-linguistic and cross-proficiency comparative studies of shell nouns. However, existing research has largely focused on the use of shell nouns in journal articles or doctoral dissertations, with relatively limited attention given to the specific genre of master's thesis abstracts. Master's thesis abstracts serve as an important window for novice academics to demonstrate their research abilities and academic literacy, and the standardization and maturity of their language use directly reflect the level of graduate students' academic English writing. Meanwhile, abstracts from SSCI journal articles written by experts represent the

academic writing norms and highest standards of the discipline. Comparing the two can effectively reveal the gap between Chinese master's students and expert levels in academic English writing.

Based on Schmid's definition of shell nouns and Jiang and Hyland's functional classification of stance expression, this study builds two small-scale corpora to compare and analyze the use characteristics of shell nouns in Chinese master's thesis abstracts and English expert paper abstracts. Specifically, this study aims to explore the similarities and differences between the two groups of authors in three aspects: the frequency of shell noun use, lexico-grammatical constructions, and stance-expressing functions, with the goal of providing reference for the teaching of academic writing to Chinese English major students^[1,2].

2. Research Design

2.1. Analytical Framework

The theoretical analytical framework of this study is based on two core theories: Schmid's definition of shell nouns and his theory of lexico-grammatical constructions, and Jiang and Hyland's functional classification of shell nouns for stance expression.

At the definitional level, Schmid proposed that shell nouns are defined not by their inherent properties but by their function. When an abstract noun is used in a "shell-content" complex to summarize, package, and convey complex propositional information, it can be regarded as a shell noun. Based on this definition, this study establishes the identification criteria for shell nouns as follows: the noun must appear in constructions such as "N-that," "N-to," "N-of," or "N-wh" and serve to summarize or explain the content of the following clause or phrase.

At the level of lexico-grammatical constructions, Schmid pointed out that shell nouns frequently occur in four constructions: N-that, N-to, N-of, and N-wh. Among these, the N-that and N-of constructions are particularly common in academic discourse because they effectively package and explain complex information, helping readers understand the author's intentions and argumentative logic.

At the level of stance-expressing functions, Jiang and Hyland classified the functions of shell nouns into three major categories and eight subcategories. Nature shell nouns are used to indicate the essence of things or events, subdivided into thing (e.g., study, paper, issue), event (e.g., process, change, development), discourse (e.g., claim, argument, conclusion), and cognition (e.g., idea, view, understanding). Attribute shell nouns are used to evaluate the characteristics of things or events, subdivided into quality (e.g., advantage, difficulty, importance), manner (e.g., way, method, approach), and state (e.g., possibility, ability, need). Relation shell nouns are used to express causal or similarity relations between things or events, such as reason, result, difference, and relationship. This classification framework provides a systematic tool for analyzing the differences in stance construction between the two groups of authors in this study.

2.2. Research Questions

Based on the above theoretical framework, this study proposes the following three research questions:

- 1) What are the similarities and differences in the frequency of shell noun use between Chinese master's thesis abstracts and English expert paper abstracts? What are the differences in high-frequency shell nouns used by the two groups of authors?
- 2) What are the similarities and differences in the distribution of lexico-grammatical constructions of shell nouns between Chinese master's thesis abstracts and English expert paper abstracts? Are there significant differences between the two groups in the frequency of use of the four constructions: N-that, N-to, N-of, and N-wh?

3) What are the similarities and differences in the choice of shell nouns based on stance-expressing functions between Chinese master's thesis abstracts and English expert paper abstracts? What characteristics do the two groups of authors show in the use of nature, attribute, and relation shell nouns?

2.3. Research Methods

This study adopts a corpus-based comparative analysis approach. The specific research methods and procedures are as follows.

This study built two small-scale corpora. The Master's Thesis Abstract Corpus (hereinafter referred to as the "Master's Corpus") contains 10 English abstracts, all selected from master's theses in the Department of Applied Linguistics of five normal universities in Northeast, Northwest, Southeast, Central, and Southwest China, with theses published between 2023 and 2025. To ensure the representativeness of the data, the selected abstracts are all from linguistics research, covering subfields such as second language acquisition, discourse analysis, and foreign language teaching. Statistics show that the Master's Corpus has a total of 38,044 characters. The Expert Paper Abstract Corpus (hereinafter referred to as the "Expert Corpus") contains 10 English abstracts, all selected from three SSCI Q1 journals: *Language Learning*, *Applied Linguistics*, and *The Modern Language Journal*. These journals represent the highest international standards in the fields of applied linguistics and foreign language teaching, and their abstracts can serve as a reference standard for native speaker or high-level academic writing. The Expert Corpus has a total of 13,828 characters.

This study used TreeTagger 3.0 software to perform part-of-speech tagging on the texts in the two corpora to identify the part-of-speech category of each word. The tagged texts were then imported into the AntConc 3.5.7 corpus retrieval tool, and regular expressions were used to extract shell nouns in the four constructions: N-that, N-to, N-of, and N-wh. Since the automatically extracted results may contain items that do not conform to the definition of shell nouns, all extraction results were manually screened to remove non-conforming concordance lines, retaining only those shell nouns that conform to Schmid's definition of the "shell-content" relation. Subsequently, according to Jiang and Hyland's functional classification framework, each shell noun was tagged with its functional category, and the frequency and normalized frequency of each category were calculated. Since the total number of characters in the two corpora differs, all frequencies were normalized per 10,000 words to ensure data comparability.

2.4. Research Procedures

The specific implementation steps of this study are as follows: Step 1: Collect English abstracts that meet the requirements from master's theses of five normal universities and three SSCI journals, and build two text corpora; Step 2: Use TreeTagger to perform part-of-speech tagging on the corpora; Step 3: Use AntConc to extract candidate shell nouns in the four lexico-grammatical constructions; Step 4: Perform manual screening to remove items that do not conform to the definition of shell nouns; Step 5: Calculate the total frequency, normalized frequency, and high-frequency words of shell nouns; Step 6: Tag each shell noun with its function according to Jiang and Hyland's classification framework; Step 7: Compare and analyze the similarities and differences between the two corpora in the three aspects of frequency of shell noun use, distribution of lexico-grammatical constructions, and stance-expressing functions; Step 8: Discuss and draw conclusions based on the comparative results.

3. Results and Discussion

3.1. Comparison of Shell Noun Frequency

By counting the shell nouns in the four lexico-grammatical constructions in the two corpora, this study found significant differences in the frequency of shell noun use between the two

groups of authors. In the Master's Corpus, the normalized frequency of shell noun use was 87.36 per 10,000 words, while in the Expert Corpus, this figure was 112.47 per 10,000 words. The frequency of shell noun use in the Expert Corpus was significantly higher than that in the Master's Corpus, by about 25 cases per 10,000 words.

In terms of high-frequency shell nouns, the top five most frequent shell nouns in the Master's Corpus were: study, process, development, relationship, and use. The top five most frequent shell nouns in the Expert Corpus were: effect, relationship, role, finding, and implication. The only high-frequency shell noun shared by the two corpora was relationship. This comparison reveals that Chinese master's students tend to use broader, more general shell nouns (e.g., study, process, development) in their academic writing, whereas experts use more semantically specific and analytically deep shell nouns (e.g., effect, finding, implication). This difference may reflect a gap in academic writing maturity between the two groups: master's students are accustomed to using "the paper itself" or "the research process" as topic-guiding words, while experts focus more on substantive content such as "research findings" and "research effects."

3.2. Comparison of Lexico-Grammatical Construction Distribution

In terms of lexico-grammatical constructions, the two groups showed significant differences. In the Expert Corpus, the distribution of the four constructions was relatively balanced: the N-of construction accounted for the highest proportion, approximately 65.8%; the N-that construction ranked second, accounting for approximately 18.3%; the N-to construction accounted for approximately 12.1%; and the N-wh construction accounted for the lowest proportion, approximately 3.8%. In contrast, the distribution of constructions in the Master's Corpus was extremely unbalanced: the N-of construction accounted for as high as 85.2%, the N-that construction accounted for only 5.6%, the N-to construction accounted for 7.2%, and the N-wh construction accounted for 2.0%. This data indicates that Chinese master's students rely heavily on the N-of construction to express the relationship between shell nouns and subsequent content, while their use of the N-that construction is severely insufficient.

The N-that construction has an important function in academic writing because it can clearly mark the explanatory relationship between the shell noun and the following clause. For example, in structures such as "the fact that ..." or "the idea that ...," the that-clause directly specifies the specific content of the shell noun, making the author's stance expression more explicit and direct. The reasons why Chinese master's students seldom use this construction may include: first, the syntactic structure of the N-that construction is relatively complex, requiring the author to have a clear grasp of the logical relationship between the main clause and the subordinate clause; second, Chinese English teaching has long emphasized the N-of structure, while providing insufficient training in the complement function of nominal clauses; third, the N-that construction has no completely corresponding structure in Chinese, and Chinese students tend to avoid this construction under the influence of native language transfer. It is worth noting that although Chinese master's students use the N-that construction less frequently, their use of the N-to and N-wh constructions is relatively closer to that of experts. For example, the difference between the two groups in the use of N-to constructions such as "ability to ..." or "need to ..." was not significant. This may indicate that Chinese master's students have relatively good command of infinitive structures, but their use of complex clause structures still needs improvement.

3.3. Comparison of Stance-Expressing Functions

Based on Jiang and Hyland's functional classification framework, this study tagged and counted the stance functions of shell nouns in the two corpora. The results showed differences between the two groups in the use of nature, attribute, and relation shell nouns.

In terms of nature shell nouns, the Master's Corpus most frequently used event nouns (e.g., process, development, change), with a normalized frequency of 32.15 per 10,000 words; followed by thing nouns (e.g., study, paper, analysis), with a frequency of 28.43 per 10,000 words; cognition nouns (e.g., understanding, idea, view) and discourse nouns (e.g., claim, argument) were used less frequently, with frequencies of 12.67 and 6.28 per 10,000 words, respectively. In the Expert Corpus, the frequency of event nouns was 28.96 per 10,000 words, thing nouns 24.37 per 10,000 words, cognition nouns 18.92 per 10,000 words, and discourse nouns 10.45 per 10,000 words. The most significant difference between the two groups lies in the use of cognition and discourse nouns: experts used these two types of shell nouns significantly more frequently than master's students. This indicates that experts tend to express their cognitive stance and make meta-comments on discourse acts in their academic abstracts, while master's students focus more on describing the research object or the research process itself. This finding confirms Hyland's view that high-level academic writers are better at using metadiscourse resources to construct author identity and interact with readers^[3].

In terms of attribute shell nouns, the Master's Corpus most frequently used manner nouns (e.g., method, way, approach), with a frequency of 14.22 per 10,000 words; quality nouns (e.g., importance, difficulty, advantage) had a frequency of 8.45 per 10,000 words; and state nouns (e.g., possibility, ability, need) had a frequency of 6.83 per 10,000 words. In the Expert Corpus, manner nouns had a frequency of 10.68 per 10,000 words, quality nouns 12.34 per 10,000 words, and state nouns 14.56 per 10,000 words. The two groups showed opposite trends in the use of attribute nouns: master's students overused manner nouns, while experts used quality and state nouns more evenly. This suggests that master's students are accustomed to describing "how the research was done" in their abstracts, while experts pay more attention to "the significance and value of the research" and "the possibilities and limitations of the research." This difference may reflect different stages of academic writing maturity: beginners focus more on describing the research process, while experts focus more on evaluating and positioning research findings.

In terms of relation shell nouns, the frequency of use was 15.38 per 10,000 words in the Master's Corpus and 12.95 per 10,000 words in the Expert Corpus. Master's students used relation nouns (e.g., relationship, effect, reason) slightly more frequently than experts, but the difference was not significant.

4. Conclusion

Based on Schmid's theory of shell nouns and Jiang and Hyland's functional classification of stance expression, this study compared and analyzed the characteristics of shell noun use in Chinese master's thesis abstracts and English expert paper abstracts. The main findings are threefold: First, the normalized frequency of shell nouns in the Expert Corpus was 112.47 per 10,000 words, significantly higher than that in the Master's Corpus (87.36 per 10,000 words); master's students tend to use broad, general nouns, while experts use more semantically specific nouns. Second, the distribution of constructions in the Expert Corpus was balanced, with N-that constructions accounting for 18.3%, whereas the Master's Corpus relied heavily on the N-of construction (85.2%), with N-that accounting for only 5.6%. Third, master's students overused event and manner nouns, while experts were more adept at using cognition, quality, and state nouns to express stance. In terms of pedagogical implications, teachers should guide students to increase the frequency of shell noun use, strengthen training in the N-that construction, and help students shift from "describing the research process" to "evaluating research findings." The limitations of this study include the small corpus size, the lack of analysis of full texts, and the absence of proficiency-level stratification analysis. Future research

could explore the phenomenon of shell nouns in Chinese academic discourse and the mechanism of native language transfer on learners' use of shell nouns.

References

- [1] Schmid, H. (2000). English abstract nouns as conceptual shells: From corpus to cognition. Walter de Gruyter.
- [2] Jiang, F., & Hyland, K. (2015). "The fact that": Stance nouns in disciplinary writing. *Discourse Studies*, 17(5), 529–550.
- [3] Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173–192.