

A Corpus-based Study on Acquisition Mode of Continuation Task's Collocational Constructions of Circumstantial Shell Nouns

Yu Qiang, Geraldine S. Wakat

[**Abstract**] This research is built upon a custom-developed continuation task mini-corpus, incorporating New Concept English and English novel corpora, the native-speaker LOCNESS corpus, and the COCA corpus. Utilizing Contrastive Interlanguage Approach, it thoroughly investigates the collocational constructions frequency, proportion, and Mutual Information (MI) values of five circumstantial shell nouns (environment, place, background, situation, position). The study reveals that students' exhibit varied proficiency levels in specific collocational constructions, underscoring a disparity between instructional content and real-world language usage. English instruction should thus focus on students' language application in specific contexts, employ a diverse range of teaching materials, and enhance the teaching of key lexical collocations. Educators should merge classroom teaching with practical language use to more effectively guide students' language acquisition. Moreover, the significance of knowledge about collocational constructions of shell nouns in English learning cannot be overlooked. Teaching strategies should be flexibly adjusted to meet students' specific needs, enhancing their deep understanding and effective application of English vocabulary collocations.

[**Key words**] continuation task; shell nouns; collocational constructions; corpus

[**About the author**] Yu Qiang (1988—), male, from Yibin, Sichuan, China, Ph. D. student in Saint Louis University. Research interests: basic education, corpus linguistics. Geraldine S. Wakat, coordinator in Languages and International Studies Unit, Languages and Communication Department, School of Teacher Education and Liberal Arts, Saint Louis University.

[**DOI**] <https://doi.org/10.62662/kxwxy0107009>

[**Website**] www.oacj.net

The concept of collocation in corpus linguistics can be traced back to Firth's (1957 Firth) theory of word pairing. He pointed out that words used in conjunction with each other form collocations, thus constituting a particular expression of meaning. Based on Firth's theory, Sinclair further deepened the study of collocation. At the operational level, Sinclair proposed a more specific definition of collocation in 1991, which refers to two or more words co-occurring at a neighboring distance in a text. This definition mainly emphasizes the importance of collocation in text and context. Stefanowitsch and Gries introduced the notion of constructive collocation, which refers to the expression of node words in grammatical patterns with specific meanings.

Schmid (2000) states that the basic functions of shell nouns can be summarized in three areas: the descriptive function, the concept-building function and the linking function. The descriptive function is used to depict the chunks of information contained in clauses, the concept-building function refers to the embedding of chunks of information into noun concepts, and the linking function is mainly concerned with discourse articulation. According to the semantic division, shell nouns can be categorized into six types: factual, linguistic, mental, modal, eventive, and circumstantial (e. g., environment, place, background, situation, position). In view of the frequent occurrence of circumstantial description in the process of continuation task, this study chooses the above five shell nouns as the research object, adopts the Contrastive Interlanguage Approach advocated by Granger (Granger 1998), analyzes their comparative use in the self-constructed mini-reading and writing corpus and the New Concept English & literature corpus, and the target corpus of the compositions of the vernacular language, and

then compares them with the COCA specific contexts and distributions to focus on the characteristics of students' acquisition of collocational constructions knowledge and related pedagogical insights.

1 Literature review

Since Schmid defined the connotation of shell nouns and deeply studied their constructive framework, semantic features and functional classification, scholars in China have carried out extensive and in-depth research and discussion on this category of abstract nouns that have a broad meaning of their own and at the same time are capable of carrying other complex and concrete information. In the CNKI database, by searching the papers related to shell nouns, a total of 92 research documents were covered from 2004 to 2023 publication years. Further narrowing down the search scope to SCI, EI, Beida core, CSSCI, CSCD and AMI academic journals, we found 20 research papers related to shell nouns. In terms of the yearly distribution of the papers published, there was 1 paper in 2004, 1 paper in 2007, 1 paper in 2013, 3 papers in 2016, 2 papers in 2017, 1 paper in 2018, 1 paper in 2019, 1 paper in 2020, 4 papers in 2021, 2 papers in 2022, and 3 papers in 2023. According to the classification of the journals in which the papers were published, they include Foreign Language Education in China, Foreign Languages and Literature, Foreign Language Learning Theory and Practice, Foreign Languages Research, Foreign Languages and Their Teaching, Technology Enhanced Foreign Language Education, Journal of Northwest Normal University, Contemporary Rhetoric, Chinese Teaching in the World, Modern Foreign Languages, and so on. These studies provide us with strong theoretical support for a deep understanding of the linguistic nature of shell nouns and their pragmatic functions in different contexts.

In Zhang Gaoyuan and Yang Xiaojun's paper "New Perspectives on English Abstract Nouns", it is mentioned that whether a noun is categorized as a shell noun or not does not depend on its intrinsic characteristics, but on its function. In other words, once a speaker decides to use a noun in a "shell-contain" complex for a specific purpose, it is considered a shell noun. There is no class of words in the language that specializes in conceptual shells, but rather some nouns that also function as conceptual shells. According to their actual functions, they should be accurately called "nouns used as shells". Therefore, it is necessary to give rich meaning to nouns used as shells through collocation and specific contexts to make them more vivid. Qiu Xixin, Dong Min and Liu Weiwei established a cognitive semantic analysis framework based on Schmid's (2000) classification system for shell nouns. They utilized a corpus of English academic papers in physics and aerospace engineering containing 2 million words to investigate different sub-domains of knowledge construction in academic discourse in science and engineering. In their study, Chen Yingfang and Ma Xiaolei employed cluster analysis and multinomial specific co-occurrence lexical position analysis to statistically analyze the combinations of "verb+shell noun" in 497 English academic papers. At the same time, with the help of the Event Domain Cognitive Model, the collocation mechanism of these combinations was studied in depth. Based on the shell nouns proposed by Schmid, Zhang Xuemei and Liu Ping selected 656 shell noun paraphrases from the Oxford Advanced Learner's English Dictionary (8th edition) and constructed a small corpus. Subsequently, they conducted an in-depth analysis of these shell noun paraphrases using the path analysis method of local grammar, aiming to reveal the paraphrasing paradigm of abstract nouns. Niu Baoyi and Shen Shaoshuai utilized the statistical data of the British National Corpus to study the two constructions "try to V" and "try V-ing" in English using the Collocational Constructions Analysis. Their study aims to explore the collocational tendencies, preferences and exclusions of "try" with "to V" and "V-ing" in these two constructions, and to analyze the semantic clustering features of "V" verbs in these two constructions, as well as their relationships with the meanings of the constructions. Wang Huan and Lin Zhengjun used the Collocational Constructions Analysis and applied the ΔP method to assess the unidirectional relationship between the constructions and the verbs in order to explore the dominance of the verbs in collocations with the "way" constructions. The aim of their study is to investigate the directional collocation between verbs and constructions.

The results of the study show that there are differences in the strength of collocations and differences in dominance in unidirectional relations. Using learner corpus data and Collocational Constructions Analysis as a research framework, Sun Haiyan investigated the use of narrative real class shell nouns by students at different learning stages, aiming at exploring the developmental characteristics of collocational knowledge among Chinese students. Liu Shan used the Collocational Constructions Analysis as an analytical study to investigate the differences between learners at different levels in the use of mental category shell nouns and with native speakers, and explored the learners' patterns of second-language construal acquisition. Wang Min and Cai Ning used a mixed pre- and post-test design to investigate the effects of linguistic complexity on language fluency, accuracy and complexity of reading and subsequent writing of input texts in Chinese learners' English writing. Wang Chuming (2016, 2017) proposed the Continuation Theory, which asserts that language acquisition is realized through "continuation", and the efficiency of language learning also depends on the practical use of "continuation". In this context, "continuation" refers to the fact that in language communication, speakers express their own thoughts by understanding and continuing the expressions of others, realizing the articulation of previous and subsequent content, thus facilitating effective communication. In addition, "continuation" can also promote language learning by stimulating learners' intrinsic motivation to express themselves and by utilizing previous correct input. Therefore, when practicing reading followed by writing, students need to deeply understand the specific language forms in the original text (e.g., environmental shell nouns, etc.) in order to express their personal views and thoughts more clearly and accurately. These studies not only deepen our understanding of the role of shell nouns in language learning and application, but also provide important insights for English teaching. In the process of teaching, teachers should pay attention to developing students' understanding and mastery of shell nouns and their collocational constructions, while emphasizing the importance of context in vocabulary use, in order to promote students' linguistic expression and writing literacy and competence.

In this study, according to Schmid's categorization of shell nouns, the Contrastive Interlanguage Approach is used to analyze the frequency of occurrence of five words of the circumstantial category (environment, place, background, situation, and position) in different corpus, the intensity of node-word collocation (MI value), and explore the students' usage characteristics and collocation distribution in their specific writing, and then dig out the reasons behind and the teaching inspiration. The Contrastive Interlanguage Approach has been a highly respected classic model in second language corpus research. Army pointed out that Contrastive Interlanguage Approach adopts a comparative analysis approach, focusing on typical language phenomena. The method is capable of comparing learners' languages with the target native language as well as comparing the languages of different groups of learners, thus revealing the general trend of learners' linguistic phenomena and providing an effective way of discovering typical second language features. This study focuses on analyzing five different shell noun constructions, including SN + that-clause, SN + wh-clause, SN + prep., SN + be + that-clause, and prep. +SN.

The data for the study are mainly derived from a small self-constructed corpus (hereafter denoted as C1), integrated with a corpus of New Conceptual English and screened for some literature (hereafter denoted as C2), and referenced to corpus excerpts from the LOCNESS onomatopoeia corpus (hereafter denoted as C3) and the COCA corpus (hereafter denoted as C4). C1 is derived from 154,487 words of a small continuation task corpus of students with different English proficiency levels from the freshmen and sophomore classes of a Sichuan middle school; C2 has a capacity of 152,271 words; C3 has a capacity of 156,626 words; C3 has a capacity of 156,626 words. The main research tools are AntConc 4.2.0, Wordless 3.4.0 and so on. In this study, we extracted data from each corpus and observed the frequency of occurrence and MI values of indexed line collocation constructions for five circumstantial shell nouns (environment, place, background, situation, and position). AntConc software provides Mutual Information Score (MI score), which is calculated in this study as an example to emphasize that

this statistic shows how much one of the two words co-occurring with each other influences the other, or the probabilistic information that the frequency of occurrence of one word in the corpus can provide the occurrence of the other word. The higher the MI value is, the stronger the collocation of the two words is. This study focuses on answering the following two questions.

- 1. What are the specific types of collocational constructions for English circumstantial shell nouns? What is their specific distribution?
- 2. What are the implications of the circumstantial shell nouns for the teaching of collocational constructions use in English vocabulary?

2 Constructions characteristics of the circumstantial shell nouns

2.1 Collocational constructions of the shell noun environment

The purpose of this section is to analyze the frequency, proportion and mutual information (MI) value of the circumstantial core noun “environment” in each corpus. The number of occurrences of the word “environment” is 19 in the self-constructed continuation task corpus (C1), 17 in the corpus combining new concepts textbooks and novel texts (C2), and 24 in the native language corpus LOCNESS (C3). Referring to the Corpus of Contemporary American English (COCA), the number of occurrences of “environment” is 88,301. This study focuses on the following four main grammatical collocations: “prep. (of)+environment” with the preposition “of”, “prep. (for)+environment” with the preposition “for”, “environment + that” with the conjunction, and the clause - like “environment+where”.

Table 1. Frequency, proportion and MI value of the collocational constructions of node word environment

| Corpus | Total Frequency | 1 of+environment | MI | 2 environment | MI | 3 environment+that | MI | 4 environment+where | MI |
|--------|-----------------|------------------|---------|---------------|---------|--------------------|---------|---------------------|---------|
| C1 | 19 | 3(15.8%) | 3.85157 | 2(10.5%) | 3.61525 | 2(10.5%) | 4.07034 | 0(0%) | 0 |
| C2 | 17 | 4(23.5%) | 2.55255 | 0(0%) | 0 | 7(41.2%) | 4.60688 | 4(23.5%) | 6.79487 |
| C3 | 24 | 1(4.1%) | 0.50563 | 1(4.1%) | 2.41245 | 2(8.3%) | 4.29479 | 0(0%) | 0 |
| C4 | 88301 | 1140(1.3%) | | 3012(3.4%) | | 3027(3.4%) | | 1736(1.9%) | |

The data presented in Table 1 show us the collocation patterns and distributional properties of the word “environment” in four different corpora. The collocation “of + environment” is presented three times in C1, accounting for 15.8%, and four times in C2, accounting for 23.5%. This finding suggests that high school students have a strong preference for and familiarity with this collocation, reflecting the fact that teachers may have paid sustained attention to and reinforced this collocation in their daily teaching, so that students can use it more easily in their writing practice. However, the collocation “environment + for”, which appeared 3012 times in COCA, is not adopted in C2, and only appears 2 times and 1 time in C1 and C3, respectively. The comparison with COCA suggests that the sparse frequency may be related to the small size of the corpus. On the other hand, “environment + that” is the most frequent construction in COCA with 3,027 occurrences, which is the same as that of LOCNESS (C3) in C1, and 7 occurrences in C2, accounting for 41.2% of the total number of occurrences, which shows that both Chinese and native English speakers have a good mastery of this usage. This shows that both Chinese and native English speakers have a good grasp of this usage. As for “environment + where”, it is not recorded in C1 and C3, but it occurs 4 times in C2 and 1736 times in COCA. This distribution may be related to the small number of words in the selected corpus. It is obvious that Chinese students’ use of “environment + where” is still insufficient, which needs to be strengthened in future teaching, and teachers should emphasize more on contextual use and deeper meaning.

Further analysis shows that for native speakers, a mutual information (MI) value greater than 3 is usually

considered statistically significant. From the data, it can be observed that the MI value of “of + environment” in C1 is 3.85157, which shows strong statistical correlation; the MI value of “environment + for” in C1 is 3.61525; while the MI value of “environment + that” in C1, C2 and C3 is more than 4, which shows its greater strength of collocation; in particular, the MI value of “environment + where” is as high as 6.79487, showing high statistical correlation. These high MI collocation patterns should be emphasized in future English teaching and writing instruction to raise students’ awareness of their importance and practical application.

2.2 Collocational constructions of the shell noun place

In the study of the use of the word “place” in the corpus, its frequency, percentage and mutual information (MI) values are analyzed. The total frequency of “place” in the self-constructed corpus C1 (the continuation task corpus) is 50, in C2 (the corpus combining new concepts textbooks and novels) 100, in C3 (the native language corpus LOCNESS) 68, and in the big Corpus of Contemporary American English (COCA) 475,406. In order to better analyze the use of the enclitic noun place, combined with the observation of specific contexts in the corpus, this study focuses on four different collocational constructions: “place+of”, “place+where”, “place+is that”, and “in+place”.

Table 2. Frequency, proportion and MI value of the collocational constructions of the node word place

| Corpus | Total Frequency | 1 place+of | MI | 2 place+where | MI | 3 place+is that | MI | 4 in+place | MI |
|--------|-----------------|-------------|---------|---------------|---------|-----------------|---------|-------------|---------|
| C1 | 50 | 0(0%) | 0 | 1(2.0%) | 3.95868 | 0(0%) | 0 | 1(2.0%) | 2.40295 |
| C2 | 100 | 23(23.0%) | 3.19684 | 17(17.0%) | 6.93892 | 1(1.0%) | 0.48181 | 8(8.0%) | 3.33819 |
| C3 | 68 | 4(5.9%) | 0.74009 | 2(2.9%) | 4.85662 | 2(2.9%) | 0.47994 | 18(26.5%) | 6.17288 |
| C4 | 475406 | 16491(3.4%) | | 19301(4.1%) | | 130(0.02%) | | 36230(7.6%) | |

Specifically, C2 has the most frequent use of the construction 1 “place+of”, with 23 times, accounting for 23.0% , and the MI value reaches 3.19684, showing a strong degree of collocation, while the number of times this construction occurs in C1 is 0. In fact, in the real reading and subsequent writing of the original text, there are a lot of concrete descriptions and narratives of place’s specific descriptions and narratives, our students can increase their understanding of the original text to really master the usage and apply it to their post-reading and writing practice.

Construction 2 “place+where” has the highest frequency and percentage of use in C2, 17 times and 17.0% respectively, and the MI value is 6.93892, indicating that this collocation has significant statistical relevance in C2. In order to verify whether the collocation place where is the most widely distributed in the works of fiction, it is analyzed by searching the chart of place+where in COCA, as shown in Table 3:

Table 3. Distribution of the node word place+where in COCA



According to Table 3, the number of occurrences of the construction “place+where” in COCA’s novel texts

reaches 3,345, and the significance of this number not only emphasizes its frequent use in novel writing, but also highlights the importance of this phrase in depicting the scene and developing the storyline. This statistic provides solid data support for our study and suggests that our analysis of lexical collocation patterns in literary texts is sound. In addition, these data reflect the tendency of fiction authors to use specific language collocations in specific contexts, a phenomenon that is particularly significant in the learning and teaching of English and deserves educators’ attention. Therefore, increased attention should be paid to these typical collocations in instructional design to guide learners to understand and utilize these kinds of expressions more deeply in order to enhance their language comprehension and application of literary and narrative texts.

Construction 3 “place+is that” has 0 occurrences in C1, and of course it has a total of only 130 occurrences in COCA. Whereas, construction 4 “in +place” shows an increasing trend in frequency from C1, C2, and C3 respectively (from 1, 8 to 18), and occupies a high frequency count of 36,230 in the COCA corpus. In addition, in+place accounted for 26.5% in C3, with an MI value of 6.17288, which reflects that the preference of this construction in C3 may be related to the contextual use in the native language environment. The above data provide insights into the usage habits and preferences of the word “place” in different contexts, while the analysis of the mutual information (MI) values reveals the collocational strength of the grammatical structures in different corpora, which is of some significance to the understanding of language acquisition and pedagogical research.

2.3 Collocational constructions of the shell noun situation

The purpose of this section is to investigate the frequency and proportion of the distribution of the shell noun “situation” and its Mutual Information (MI) index in the four types of corpus selected. In the self-constructed continuation task corpus (C1), the word “situation” occurs 39 times; in C2, it occurs 17 times; in the native language corpus LOCNESS (C3), it occurs 55 times; and compared to the Corpus of Contemporary American English (COCA), it occurs as frequently as 136,921 times. The next main focus was to assess the following four main collocational constructions: “situation + in”, “situation + is that”, “situation + where”, and “situation+of”.

Table 4. Frequency, proportion and MI value of the collocational constructions of the node word situation

| Corpus | Total Frequency | 1 situation+in | MI | 2 situation+is that | MI | 3 situation+where | MI | 4 situation+of | MI |
|--------|-----------------|----------------|---------|---------------------|---------|-------------------|---------|----------------|---------|
| C1 | 39 | 1(2.0%) | 0.64111 | 2(5.1%) | 2.80707 | 0(0%) | 0 | 3(7.7%) | 2.81409 |
| C2 | 17 | 1(5.9%) | 1.89459 | 7(41.2%) | 5.84555 | 4(23.5%) | 7.40785 | 0(0%) | 0 |
| C3 | 55 | 5(9.1%) | 2.17091 | 5(9.1%) | 2.10797 | 1(1.8%) | 4.16272 | 4(7.3%) | 1.04622 |
| C4 | 136921 | 8723(6.4%) | | 370(0.3%) | | 5154(3.8%) | | 2910(2.1%) | |

As can be seen from the table, the construction “situation+is that” occurs 7 times in C2, accounting for 41.2% of the total number of occurrences, and has the highest MI value of 5.84555, which shows that this construction has a high dependency strength in the corpus of C2. In C3, “situation+where” occurs only once, but its MI value reaches 4.16272, which means that although it is not common, when it occurs, it has its specific use. In addition, the MI values of “situation + in” in C1, C2, and C3 are 0.64111, 1.89459, and 2.17091, respectively, which shows a gradual increase in the strength of the collocation. What is more obvious is that the frequency of the construction situation+where in C1 is zero, which may be related to the size of the reservoir, but of course it is enough to show that our students have not mastered this collocation and its usage, and need to be strengthened in the future study. The above comparative analysis on the level of MI values provides us with a statistical analysis on the use of “situation” collocations in different corpus, which is of guiding significance for the formulation of English reading and writing strategies, and for the refinement of teaching content.

2.4 Collocational constructions of the shell noun background

The purpose of this section is to discuss the frequency, percentage and distribution of mutual information (MI)

values of the shell noun “background” in the four corpora. In the self-constructed continuation task corpus (C1), “background” occurs only 5 times; in C2, 23 times; in the LOCNESS corpus (C3), 33 times; in contrast, in the Corpus of Contemporary American English (COCA), 53,175 times. The following is a further data analysis for the four collocational constructions “background+of”, “background+to”, “background+in” and “background+that”.

Table 5. Frequency, proportion and MI value of the collocational constructions of the node word background

| Corpus | Total Frequency | 1 background+of | MI | 2 background+to | MI | 3 background+in | MI | 4 background+that | MI |
|--------|-----------------|-----------------|---------|-----------------|---------|-----------------|---------|-------------------|---------|
| C1 | 5 | 1(20.0%) | 3.77757 | 0(0%) | 0 | 0(0%) | 0 | 0(0%) | 0 |
| C2 | 23 | 4(17.4%) | 2.73213 | 3(13.0%) | 2.2516 | 2(8.7%) | 2.45849 | 4(17.4%) | 3.97832 |
| C3 | 33 | 4(12.1%) | 0.74009 | 2(6.0%) | 1.85662 | 15(45.5%) | 3.7128 | 2(6.0%) | 0.47994 |
| C4 | 53175 | 2952(5.6%) | | 1050(1.8%) | | 1984(3.7%) | | 661(1.2%) | |

As can be seen from the table horizontally, of the four constructional collocations of background in C1, only construction 1 background+of occurs 1 time, and the other three are all 0 times. Comparing vertically, it is clear that the construction 3 background+in appears 1984 times in COCA and has a frequency of 0, 2 and 15 times in C1, C2 and C3. The percentage of this construction in C3 reaches 45.5% and the MI value of 3.7128 reflects its prominence in this context.

The in-depth comparison of frequencies, percentages and MI values can suggest that we teachers should focus on developing students' sensitivity and adaptability to lexical collocations in different teaching stages and contexts.

2.5 Collocational constructions of the shell noun position

This section analyzes the frequency, percentage, and distribution of mutual information (MI) values of the shell noun “position” in the four categories of the selected corpus. In the self-constructed reading and writing corpus (C1), “position” occurs 32 times, while in C2 it occurs 32 times, in the LOCNESS corpus (C3) it occurs 47 times, and in the Corpus of Contemporary American English (COCA), the frequency reaches 141,770 times. The next main analyses the collocational constructions of “position + prep. (in)”, “position + conj. (where)”, “position+prep. (of)”, and “position+conj. (that)”.

Table 6. Frequency, proportion and MI value of the collocational constructions of the node word position+where

| Corpus | Total Frequency | 1 position+in | MI | 2 position+where | MI | 3 position+of | MI | 4 position+that | MI |
|--------|-----------------|---------------|---------|------------------|----|---------------|---------|-----------------|---------|
| C1 | 32 | 0(0%) | 0 | 0(0%) | 0 | 0(0%) | 0 | 1(3.1%) | 1.73332 |
| C2 | 32 | 3(9.4%) | 2.56701 | 0(0%) | 0 | 11(34.3%) | 3.71512 | 5(15.6%) | 3.82381 |
| C3 | 47 | 8(17.0%) | 3.07575 | 0(0%) | 0 | 13(27.7%) | 2.97341 | 0(0%) | 0 |
| C4 | 141770 | 7854(5.5%) | | 1808(1.3%) | | 12605(8.9%) | | 4696(3.3%) | |

Table 7. Distribution of the node word position+where in COCA



According to the data in Table 6, a horizontal comparison shows that all four collocations (except position+that which has only 1 occurrence) have 0 occurrences in C1. Comparing vertically, both the frequency and percentage of construction 1 position+in and construction 3 position+of in the C1\C2\C3 corpus show an increasing trend respectively. In terms of MI values, in C2, the MI value of “position+of” reaches 3. 71512, indicating the specificity and strength of this collocation in this corpus, while the MI value of “position+in” in C3 is 3. 07575, indicating the specificity and strength of this collocation. The MI value of “position+in” in C3 is 3. 07575, indicating that it is also significant in this context. Table 7 is designed to examine the reason behind the fact that the construction position+where is found 0 times in C1\C2\C3. By comparing the distribution in the COCA corpus, it is clear that this construction occurs 1808 times, but only 91 times in written or academic language, and it is mainly used in spoken (629 times) or communicative language. Of course, the fact that the construction position+where is zero in C1\C2\C3 is also related to the size of the three corpora.

The construction 3 position+of occurs 0 times in C1, which shows that our students have not mastered the collocational constructions of position. Because there is an original sentence in the article of Module 5 of the New Foreign Studies Edition of the New Senior High School English Teaching Materials, “They have found out that the monarch is able to tell the time of day. It uses its eyes to measure the position of the sun”. This collocation position +prep. (of) occurs 12605 times inside the COCA corpus, accounting for 8.9% ; 11 times inside the New Concepts +Fiction Texts corpus, occupying 34. 3% ; and a whopping 13 times inside the Native Corpus, accounting for 27. 7% . It can be seen that the collocational constructions of position of are more and more widely used, but in the self-constructed corpus, the number of times this collocation occurs is 0, which means that students have not mastered the collocational constructions and usage, and in the future teaching, our students need to strengthen the learning of the collocational constructions, and produce more outputs of this structure and its practical use in reading and writing.

Table 8. SN+wh-clause collocation

| Corpus | 1 environment+where | 2 situation+where | 3 place+where | 4 postion+where |
|--------|---------------------|-------------------|---------------|-----------------|
| C1 | 0(0%) | 0(0%) | 1(2.0%) | 0(0%) |
| C2 | 4(23.5%) | 4(23.5%) | 17(17.0%) | 0(0%) |
| C3 | 0(0%) | 1(1.8%) | 2(2.9%) | 0(0%) |
| C4 | 1736(1.9%) | 5154(3.8%) | 19301(4.1%) | 1808(1.3%) |

By sorting out the five shell nouns discussed in this paper, it can be found that the frequency of SN+where construction in C1 is 0. This shows that our students have not mastered the type of subordinate clauses. It shows that our students have not mastered the type of subordinate clauses following shell nouns, which is of course a key and difficult point for secondary school students in their study. SN+where construction is used most frequently in New Concept English and novel texts (C2), in which place+where occurs 17 times. The fact that place+where appears 17 times shows that it is widely used in novels and other subjects. In the table, the construction 4 position +where appears 1808 times in the COCA corpus, but the number of times it appears in C1, C2 and C3 is 0. Teachers can reproduce this collocational construction in various forms in the future classroom to deepen the students’ understanding of this collocation and its use.

3 Discussion

The present study begins by noting that in the C1 corpus, students’ frequency of use and MI values for the word “environment” and its related collocations reveal mastery of specific phrases. For example, the high frequency and MI value of “of + environment” reflect the prevalence of this collocation in students’ writing, which

may stem from the emphasis on this structure in teaching. In contrast, “environment + where” was less frequently used, in contrast to its common usage in large corpora such as COCA, suggesting a lack of awareness of this collocation.

The collocational construction of the shell noun “place” is presented in different corpora, and we can find that the use of “place” in various texts has obvious characteristics and patterns. First of all, the collocation “place + of” is frequently used in the New Concepts textbook and the novel text (C2). This suggests that in teaching, teachers should emphasize the use of this collocation to help students master its usage and meaning. For example, students’ understanding of the collocation “place of” can be deepened by providing and showing a lot of examples and exercises in COCA, such as a place of respect and love; a place of refreshment and everlasting blessedness; a place of exploration; a place of invulnerability or undefeatability, and so on. In addition, teachers can design some continuing writing tasks to encourage students to use this collocation when describing a scene or narrating a story, so as to improve their writing ability. Second, the collocation “place + where” is widely distributed in novel texts. This shows that in teaching, teachers should pay attention to cultivating students’ ability to perceive and understand literary language. By analyzing the scenes and contexts in which “place where” is used in the novel, students can better understand the role of this collocation in the narrative, such as COCA corpus example sentences: Because you are studying in a place where food is literally laid out for them; are you looking for a place where you can unwind and relax yourself from your busy day; museum organizers wanted to use the space as a place where community members could engage with art in a new way; we have comments turned on our blog so we can create a place where people can gather and converse around the subject matter we write about. Teachers can guide students to observe and record the usage of “place where” when reading novels, and then discuss and analyze it in class to help students understand its meaning and application. In addition, teachers can also use multimedia resources, such as videos and pictures, to show the specific application of “place + of” and “place + where”, so that students can visualize the context and meaning of these collocations. Through this interactive and visual teaching method, students’ interest and participation can be enhanced, so that they can master these important vocabulary collocations more effectively.

The comparative analysis of SN+wh-clause includes the following four groups of collocational constructions: environment+where; situation+where; place+where; position+where. First, horizontally, all four groups of collocational constructions have 0 occurrences in the C1 corpus (with place+where occurring only 1 time). This shows that our students have not mastered the grammatical structure of abstract nouns + conjunctions (leading to homonymous clauses and place clauses), so they cannot apply it to our reading and writing. Again, the vertical comparison shows that place+where appeared 17 times in the C2 corpus, which indicates that the structure is used more in the novel text context. In the future, we can guide our students to focus on the original novel text and acquire this collocational construction in order to improve their language application ability in reading and writing or in extended discourse. In addition, Position+where appears 1808 times in COCA, but not in C1, C2 and C3 corpora. Therefore, we teachers can also use the corpus to extract specific texts and encourage students to acquire this type of knowledge structure and improve their language output ability through a multimodal approach.

In addition, when discussing the words “situation”, “background” and “position”, the study emphasizes the differences in the use of collocations and structures in different corpora. The study emphasized the differences in the use of collocations in different corpora. These differences may be due to the range of corpora to which students are exposed, the pedagogical emphasis on specific collocations, and individual language acquisition habits. For example, the MI of “situation+in” increased gradually from C1 to C3, indicating developmental learning; the use of “position+in” and “position+of” increased. The use of “position+in” and “position+of” is on the rise; the high MI value of “background + in” is particularly significant in C3, reflecting the specific usage in the native language

context. However, it should be noted that the shortcoming of MI is its insensitivity to low – frequency co – occurrences. Overall, MI values, as an important indicator of the strength of lexical collocations, provide strong statistical support for understanding the use of collocations in different contexts. In teaching practice, teachers can refer to and utilize the results of the analysis of MI values to target and strengthen students’ mastery of high MI collocations, while paying attention to collocational constructions that may be neglected, in order to enhance students’ vocabulary use and language expression. In addition, teachers should encourage students to read and practice extensively to understand and master collocations in different contexts, so as to improve their comprehensive language use.

Based on these analyses, this study aims to provide some insights into the teaching of English lexical collocations through data from a specific corpus. Xu Jiajin points out that entering the post – classical era, corpus research is increasingly focusing on analyzing the matching mechanism of form and meaning from full – caliber contextual factors. With the help of analyzing tools, teachers can extract examples from the rich corpus and show students the laws and strategies of language acquisition through a multimodal and innovative approach. Educators should pay attention to students’ language use in specific contexts and combine classroom teaching with actual language use to support students’ language learning more effectively. For example, by analyzing the use of collocational constructions in large corpora such as COCA, teachers can find out which collocations are more common among native speakers and adjust their teaching accordingly to reflect authentic language use. In conclusion, this study emphasizes the importance of acquiring knowledge of collocational constructions in English language learning, and based on the characteristics of the use of shell noun collocational constructions in different corpora, it puts forward specific suggestions for teaching practice, including choosing diverse teaching materials, focusing on actual language use, reinforcing the teaching of key vocabulary collocations, and flexibly adapting the teaching strategies to suit the specific needs of students. These strategies aim to promote students’ in – depth understanding and effective use of English lexical collocations so as to enhance their linguistic expression and literacy in reading and writing.

4 Conclusion

Based on the self – constructed reading and writing corpus, this paper integrates and combines the New Concepts English and some novels text corpus, the LOCNESS corpus and the COCA corpus, and analyzes the circumstantial shell nouns “environment”, “place”, “situation”, “background”, and “position” both horizontally and vertically for their different collocational constructions, percentage and MI values. The results of this paper provide an in – depth understanding of the collocation patterns and usage habits of these nouns in different contexts. The findings of this paper have certain significance for the collocational construction teaching of shell nouns, which can guide teachers to design more scientific and effective teaching contents and teaching methods to help students improve their language application ability and reading and subsequent writing literacy skills. The shortcoming of this paper is that the corpus is not large enough, and in the future research, the corpus can be further enlarged to increase the corpus of different genres and sources in order to obtain more comprehensive and accurate data analysis results. At the same time, the cognitive mechanism and language acquisition law behind the collocational constructions of shell nouns can be explored in depth by combining the research results of linguistic theory and cognitive psychology and other related fields. Further research and excavation can also be carried out from the perspective of semantic rhyme, stylistic grammar or local grammar, so as to better understand and apply the collocational constructions of shell nouns and promote the development of English teaching.

References:

- [1] Firth, J. Papers in Linguistics[M]. Oxford: Oxford University Press, 1957.

- [2] Sinclair, J. *Corpus, Concordance, Collocation* [M]. Oxford: Oxford University Press, 1991.
- [3] STEFANOWITSCH A, GRIES S T. Collocations: Investigating the Interaction between Words and Constructions [J]. *International Journal of Corpus Linguistics*, 2003, 8(2): 209-243.
- [4] SCHMID H. *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition* [M]. Berlin: Mouton de Gruyter, 2000.
- [5] Granger, S. *The Computer Learner Corpus: A Versatile New Source of Data for SLA Research* [A]. *Learner English on Computer* [M]. London and New York: Longman, 1998a: 3-18.
- [6] Zhang Gaoyuan, Yang Xiaojun. New Perspectives in the Study of English Abstract Nouns—A Review of “English Abstract Nouns as Conceptual Shells: From Corpus to Cognitive System” [J]. *Foreign Language Teaching and Research*, 2004(6): 467-473.
- [7] Qiu Xixin, Dong Min, Liu Weiwei. Cognitive Functions of Shell Nouns and the Construction of Knowledge in Academic Discourse in Science and Engineering [J]. *Foreign Languages and Their Teaching*, 2021(6): 23-33, 72, 147-148.
- [8] Chen Yinfang, Ma Xiaolei. A Study of Verb Collocation of Shell Nouns in English Academic Discourse: An Event-domain Cognitive Modeling Perspective [J]. *Foreign Languages and Their Teaching*, 2020(1): 91-100, 149.
- [9] Zhang Xuemei, Liu Ping. A Localized Grammatical Study of Abstract Noun Interpretation in English Extraverted Learning Dictionaries—An Example from the Shell Noun Corpus of the Oxford Advanced Learner's Dictionary of English (8th Edition) [J]. *Technology Enhanced Foreign Language Education*, 2019(4): 99-105.
- [10] Niu Baoyi, Shen Shaoshuai. A Collocational Study of English “try to V” and “try V-ing” Constructions [J]. *Foreign Language Teaching and Research*, 2022, 54(5): 656-667, 798.
- [11] Wang Huan, Lin Zhengjun. A Study of Verb-construction Directional Collocation Relationships in English Way Constructions [J]. *Foreign Languages Bimonthly*, 2022, 45(4): 26-34, 87, 160.
- [12] Sun Haiyan. Characterizing the Development of Collocational Constructions in Learners' Narrative Shell Nouns [J]. *Foreign Languages and Their Teaching*, 2017(2): 81-89, 149.
- [13] Liu Shan. A Corpus-based Study on the Acquisition of Mental Category Shell Noun Constructions by Chinese English Learners [J]. *Journal of Chengdu Normal University*, 2019, 35(11): 66-72.
- [14] Wang Min, Cai Ning. The Effect of Reading and Subsequent Writing of Input Text Linguistic Complexity on Language Development in Second Language Writing [J]. *Foreign Language Teaching and Research*, 2022, 54(6): 890-901, 960.
- [15] Wang Chuming. Promoting Learning through “Subsequent Writing” [J]. *Modern Foreign Languages*, 2016, 39(6): 784-793, 873.
- [16] Wang Chuming. From “Writing Promoting Learning” to “Subsequent Writing Promoting Learning” [J]. *Foreign Language Teaching and Research*, 2017, 49(4): 547-556, 639-640.
- [17] Lu Jun. *L2 Lexical Co-selection Patterns and Teaching Design: A Corpus-based Study* [M]. Shanghai: Shanghai Jiao Tong University Press, 2022.
- [18] Xu Jiajin. Corpus Research Methods in the Post-classical Era and Their Theoretical Implications [J]. *Foreign Language Teaching and Research*, 2023, 55(3): 442-454, 481.