

Comparative Study of Word Chunks in English Writing between Chinese Secondary School Students and Native Speakers

Qing Huang

China West Normal University, China

Abstract

By using corpuses of Chinese secondary school learners' and native speakers' writing, this paper compares and analyzes the top twenty three-word chunks in Chinese secondary school learners' and native speakers' English writing. The results show that: 1) the most frequently used word chunks in the English writing of both Chinese secondary school learners and native speakers are prepositional phrases; 2) the type of word chunks with the highest frequency in the English writing of Chinese learners is "other prepositional phrase" fragments, while the type of word chunks with the highest frequency of use in the English writing of native speakers is "noun phrases + of phrase" fragments. Based on the results of the analysis, this study gives suggestions for teaching English writing in terms of vocabulary and phrases, emphasizing that writing teaching should focus on the attention to writing context and the accumulation of writing vocabulary and writing phrases.

Keywords

Corpus; English Writing; Comparative Analysis; Word Chunks Analysis.

1. Introduction

1.1. Definition of Word Chunks

Word chunks, as defined by Wray (2002)^[13], are essentially word programs that frequently occur and are composed of consecutive or non-consecutive words or other elements. These linguistic units are stored or extracted as a whole, forming an integral part of our language processing. The proficiency with which language learners utilize word chunks is a telling indicator of their overall language knowledge and competence.

In the context of the new General Senior High School Curriculum Standards (English), as outlined by the Ministry of Education (2020)^[12], writing ability stands out as a crucial component of language ability. The adept use of word chunks in writing can significantly enhance the richness, fluency, and variety of the text, thereby elevating the quality of the written work. Consequently, as Zhang (2020) emphasizes, the learning and mastery of word chunks constitute an indispensable aspect of language teaching^[14].

Word chunks represent the midpoint between vocabulary and grammar, existing as fixed or semi-fixed structures. They are linguistic units that can be readily extracted and employed as a whole, bypassing the need for extensive grammatical processing. This characteristic makes them highly efficient for learners. As Esfandiari & Barbary (2017: 24) point out, the use of word chunks can notably increase the speed of comprehension, articulation, and fluency in produced discourse^[6].

Moreover, there exists a strong correlation between the ability to effectively use word chunks and the level of second language writing proficiency, as established by Ding & Qi (2005: 52)^[5]. This ability serves as a key measure of learners' second language proficiency, as Cortes (2004: 398) asserts^[4]. In tandem with this, fostering learners' awareness of word chunks and refining

their ability to use them accurately are central tenets of both teaching and learning, as Lewis (1993: 122) underscores^[7].

Given these factors, it becomes evident that the study of the characteristics of chunking is not just important, but one of the most critical areas of research within the domain of second language writing and teaching. Understanding and leveraging the power of word chunks can lead to more effective language acquisition and enhanced writing skills, making it a focal point for educators and researchers alike in their pursuit of improving second language education.

1.2. Classification of Word Chunks

In the realm of word chunk categorization, various scholars have proposed different classification systems to better understand and analyze these linguistic units. Nattinger & De Carrico, for instance, divided word chunks into four distinct categories: multi-word chunks, customary discourse chunks, phrase-restricted chunks, and sentence-constructed chunks (Anderson et al. 1994)⁰. This classification provides a broad framework for identifying and studying different types of word chunks based on their structural and functional characteristics. Biber and his colleagues took a slightly different approach, classifying multi-word discourse into four categories: idioms, collocations, lexical-grammatical relations, and chunks. They further refined their classification by dividing academic word chunks into twelve specific categories, which are detailed in Table 1 (Biber & Barbieri, 2007)^[2]. This detailed categorization allows for a more nuanced analysis of word chunks in academic contexts, highlighting the diversity and complexity of these linguistic elements.

Michel Lewis also contributed to the field by categorizing word chunks into four types based on their framing and functioning: aggregates, collocations, idiomatic discourse, sentence frames, and quotations (Grove & Lewis, 1994)^[7]. His classification emphasizes the functional aspects of word chunks, providing insights into how they are used in different communicative situations. Despite the existence of these multiple classification systems, each offering valuable perspectives on word chunks, this paper opts for the classification proposed by Biber et al. for its operational advantages (Ma, 2011)^[10]. Biber's classification system, which categorizes word chunks into twelve detailed categories as shown in Table 1, provides a comprehensive and systematic approach to analyzing word chunks. Biber et al. define word chunks as extended collocations that occur with a frequency higher than a certain threshold. These chunks possess either complete or partial meanings and typically consist of two or more words combined in a specific way (Biber & Barbieri, 2007)^[2].

This choice of classification system is driven by the need for clear operationalization in research. Biber's detailed categorization allows for precise identification and analysis of word chunks, facilitating a more accurate understanding of their usage patterns and linguistic functions. By adopting this classification, the paper aims to provide a robust analytical framework for studying word chunks, contributing to the broader understanding of how these linguistic units operate in various contexts and how they can be effectively utilized in language teaching and learning.

"Noun phrase + of phrase" fragment refers to the noun as the center of the phrase, used to describe the subsidiary components of something, such as "the center of a building a building". "Noun phrase + other after the modifier" fragment is a noun as the center of the phrase, and the modifier fragment is used to supplement the description of the center noun. Such as "the dog is beautiful". "Prepositional phrase + embedded with of phrase" fragment is a preposition-centered phrase, usually used to express the direction, location, such as "on the top of the desk". "Other prepositional phrases" are also preposition-centered phrases, which can be used in time, space, etc. The specific meaning of which needs to be decided according to the specific phrase and context. "The preposition it + verb phrase / adjective phrase" fragment is a pronoun-centered phrase, and such phrases are usually used to describe what it refers to, i.e., the state

of someone or something, such as "it went out", "it, elegant and mannered". "Passive verb + prepositional phrase" fragment is a combination of passive voice and prepositional phrase, used to describe the passive state of things, such as "the dog was kept in the house". "The tethered verb be + noun phrase / adjective phrase" fragment is a phrase centered on a noun or an adjective, which is usually used to explain the state of things, such as "he is a doctor "; "he is tall". "Verb phrase + that clause" fragment is a verb-centered phrase used to describe the action of something. The "that" clause fragment in the fragment is used to modify the state of the action, such as "he runs so fast that no one can catch him". "Verb/adjective + to clause" fragment is a verb or adjective-centered phrase used to explain the action state of things, such as "the girl is too smart to be in a common school", "the cat runs fast to chase a mouse". "Adverbial clause" fragments are phrases centered on adverbs, which have a variety of meanings, such as adverbs indicating time, adverbs indicating the state of things in action and so on. Therefore, the meaning of adverbial clause fragment should be decided according to the specific phrase and context. "Pronoun/noun phrase + be +..." fragment is a phrase centered on a pronoun or noun, which is used to modify the state of the pronoun or noun, such as "it is cute," and "Linda is shy". Other forms are phrases are those mentioned above.

Table 1. Types of Word Chunks in English Written

1. Noun phrase + of phrase	7. "The tethered verb be + noun phrase / adjective phrase"
2. Noun phrase + other after the modifier	8. Verb phrase + that clause
3. Prepositional phrase + embedded with of phrase	9. Verb/adjective + to clause
4. Other prepositional phrases	10. Adverbial clause
5. The preposition it + verb phrase / adjective phrase	11. Pronoun/noun phrase + be +...
6. Passive verb + prepositional phrase	12. Other forms

1.3. Study Related to Word Blocks

Searching for the keyword "chunks" on Web of Science and analyzing the data, it can be found that chunks research has been fruitful in different aspects and fields, such as language teaching, computers, medicine, psychology, and so on. In recent years, empirical research on word chunks in learners' writing mainly involves two aspects: one is to explore the influence of external elements such as genre and discipline on the distribution of word chunks. Such studies are often based on the corpus of academic essays, exploring the similarities and differences in the use of chunks in different types of essays (Cortes, 2004; Paquot, 2017)^[4]. The second is to study the influence of internal factors such as English proficiency and native language background on the chunks of learners' writing. The first is cross-sectional or longitudinal study of learners with the same native language background and different English proficiency levels to explore the developmental patterns of writing chunks in the structural and functional dimensions (Huang, 2015; Huang, and Zhou, 2016)^[8-9], or to explore the developmental features of writing corpora (Chen & Baker, 2016)^[3]. The second type is to examine the influence of native language background on the use of word chunks by comparing the characteristics of word chunk use between second language learners and native speakers. Most of such studies adopt a macro perspective to provide a general description of the structure and functional distribution of learners' writing chunks (Esfandiari & Barbary, 2017)^[6]; there are also studies that adopt a micro perspective to observe specific chunks, such as stance marking chunks, and examine the similarities and differences in the usage characteristics of learners and native speakers.

Nevertheless, there are still a lot of word chunking issues that deserve further exploration. An overview of domestic and international studies on word chunks in writing reveals the following: 1) Most of the previous studies focused on intermediate and advanced learners, but few of them explored beginners in English. Beginners are an important part of English learners, and exploring their word chunks acquisition characteristics can help reveal the systematic pattern of second language word block acquisition. In addition, beginners are at the critical stage of laying a foundation, and their ability to use chunks will affect their future development of English proficiency, which is of great practical significance for teaching and learning. 2) There is no study that compares L2 beginners with different native language backgrounds, and thus fails to find out the unique features of learners' chunks use. In this regard, it is important to explore the learners' preferences and unique misuse of words, so as to guide the teaching and learning of word chunks in a more targeted way. In addition, there are not many studies that have examined Chinese secondary school students' output chunks in L2 time-limited writing with reference to the chunks commonly used by native speakers of English (Ma, 2009)^[10]. In view of this, this study explores Chinese beginners' preferences and unique misuse of chunks based on a corpus of Chinese beginners' and native speakers' writing.

Based on previous studies, this paper attempts to answer the following questions by comparing the characteristics of word blocks in the writing of native speakers and Chinese secondary school students: 1) What are the characteristics of word chunks in the writing of native speakers? 2) What are the characteristics of word chunks in the writing of Chinese secondary school students? 3) What are the similarities and differences between the two? Ultimately, it provides feasible suggestions for teaching English writing in Chinese secondary schools.

2. Research Design

2.1. Research Tools

For the purpose of this study, BFSU Power Conc 1.0 was meticulously selected as the corpus analysis tool. This sophisticated software is a product of Beijing Foreign Studies University, specifically designed to facilitate in-depth corpus analysis with high precision and efficiency. It boasts a user-friendly interface and a suite of powerful features that enable researchers to extract valuable linguistic data from large text corpora.

In the context of this research paper, BFSU Power Conc 1.0 will serve as an instrumental asset in conducting a comprehensive comparison and analysis of word chunks. The focus will be on contrasting the writing corpus of Chinese secondary school students with that of native speakers. By leveraging the capabilities of this tool, we aim to uncover patterns, frequencies, and distributions of word chunks in both corpora.

The analysis will involve segmenting the text into manageable units, identifying recurring word sequences, and examining their syntactic and semantic properties. This will allow us to draw meaningful conclusions about the linguistic habits and preferences of the two groups of writers. Furthermore, the tool's ability to handle large volumes of text with ease ensures that our analysis will be both thorough and representative of the overall writing characteristics of the target populations.

Through this meticulous comparison and analysis, we hope to shed light on the similarities and differences in word chunk usage between Chinese secondary school students and native speakers. The insights gained from this study will not only contribute to the field of corpus linguistics but also have practical implications for language teaching and learning, particularly in the context of English as a second language.

2.2. Subjects of the Study

The focal points of this research paper are the "NESSI" corpus and the "TECCL" corpus, both of which have been meticulously compiled by the faculty and students of Beijing Foreign Studies University. These corpora serve as invaluable resources for linguistic analysis, each with a distinct focus and purpose.

The "NESSI" corpus is a carefully curated collection of written compositions produced by native speakers of English and American English. It encompasses a substantial volume of text, totaling 174,429 words. This corpus provides a rich and authentic representation of the language usage patterns of native speakers, capturing the nuances and complexities of their written expression. It includes a diverse range of genres and topics, ensuring that the linguistic features analyzed are comprehensive and reflective of real-world writing scenarios.

On the other hand, the "TECCL" corpus is dedicated to the written work of Chinese secondary school students. With a total word count of 96,520 words, this corpus offers a unique insight into the language proficiency and writing styles of young learners of English as a foreign language. It reflects the challenges and achievements of these students as they navigate the complexities of English writing, providing a window into their linguistic development and educational experiences.

In the context of this study, the "NESSI" corpus assumes the role of the reference corpus. It serves as the benchmark against which comparisons are made, offering a standard of native-like language use for reference and contrast. Its data is utilized to establish norms and expectations for language performance, allowing researchers to gauge the proficiency levels and identify areas of strength and weakness in the writing of non-native speakers.

Conversely, the "TECCL" corpus functions as the observational corpus. It is the primary source of data for observation and in-depth analysis. Through the examination of this corpus, researchers can identify patterns, trends, and characteristics specific to the writing of Chinese secondary school students. This analysis is crucial for understanding the unique linguistic landscape of these learners and for developing targeted pedagogical strategies to enhance their language skills.

By juxtaposing these two corpora, this study aims to draw meaningful comparisons and generate valuable insights into the differences and similarities in the use of language between native speakers and Chinese secondary school students. The findings will not only contribute to the broader field of corpus linguistics but also have practical implications for language education, curriculum development, and teaching methodologies.

2.3. Research Steps

In the realm of corpus-based empirical research, the concept of a word chunk is pivotal. It is precisely defined as a sequence of N words that recurs F times per million words (mw), exhibiting a particular distribution pattern within the corpus. This definition underscores the significance of both frequency and distribution in identifying meaningful linguistic units.

Typically, in the study of word chunks, researchers focus on extracting and analyzing combinations of three or more words. These multi-word sequences are of particular interest as they often carry specific meanings or functions within the language. The threshold at which these word combinations are deemed noteworthy is known as the extraction frequency cut-off. Any word combination that meets or exceeds this cut-off, given a specified number of text distributions, is classified as a target word chunk. As Ma (2009) aptly noted, the research value and general significance of word chunks are directly proportional to their frequency and the breadth of their text distribution. In other words, the more frequently and widely a word chunk appears, the more likely it is to hold valuable insights for linguistic research.

To operationalize this research, the BFSU Power Conc 1.0 software was employed for the crucial task of frequency counting. The methodology was systematic and straightforward. Initially, the "NESSI" and "TECCL" corpora were loaded into the BFSU Power Conc 1.0 software. This step was foundational, as it provided the raw data necessary for the subsequent analysis.

Following this, the software's capabilities were harnessed to extract three-word chunks from both corpora. The extraction was based on word frequency, ensuring that only those chunks that occurred with a notable frequency were selected. This process was meticulous, as it aimed to filter out less significant word combinations and focus on those that held the most potential for meaningful analysis.

The final step in this process was to distill the findings further by retaining only the top 20 three-word chunks from each corpus for comparative analysis. This selection was strategic, as it allowed for a focused examination of the most prevalent word chunks, which are likely to be the most representative of the linguistic patterns within each corpus.

The culmination of this analytical journey is presented in Fig. 1 and Fig. 2, which visually depict the search results. These figures are not merely illustrative; they are the tangible outcomes of the research process, offering a clear and concise overview of the most frequently used word chunks in each corpus. By comparing these top 20 three-word chunks, researchers can glean valuable insights into the linguistic similarities and differences between the writing of native speakers and Chinese secondary school students, thereby enriching our understanding of language use and acquisition.

BFSU PowerConc 1.0 beta 25

ettings | Concordance | N-gram List | Keyness | About |

Data Type: Word | Length: 3 | Count

No.	Term(s)	Freq.
	Size	111799
	Tokens	68392
	Types	52048
1	a lot of	103
2	In my opinion	81
3	more and more	67
4	the Spring Festival	67
5	I want to	62
6	in the future	55
7	for us to	52
8	to be a	47
9	it is a	42
10	one of the	41
11	in order to	40
12	the other hand	37
13	I think it	36
14	The most important	36
15	There are many	36
16	you want to	36
17	I have a	35
18	in a word	35
19	and so on	34
20	is the most	34

Figure 1. Frequency table of three-word word chunks in Chinese secondary school students' English writing

BFSU PowerConc 1.0 beta 25

ettings | Concordance | N-gram List | About |

Data Type: Word Length: 3 Count

No.	Term(s)	Freq.
	Size	199924
	Tokens	133458
	Types	112240
1	in order to	89
2	the number of	66
3	as well as	63
4	there is a	63
5	One of the	61
6	the United S...	55
7	more and m...	49
8	in the US	43
9	in terms of	42
10	to be a	40
11	of the world	39
12	Some of the	39
13	in the United	38
14	more likely to	38
15	the amount of	38
16	As a result	37
17	due to the	37
18	part of the	37
19	the use of	35
20	the fact that	33

Figure 2. Frequency table of three-word word chunks for native speakers' English writing

3. Data Analysis and Discussion

3.1. The Most Common Word Chunks in English Writing for Both Chinese Secondary School Learners and Native Speakers are Prepositional Phrases

Upon examining Table 2 and Table 3, a noteworthy pattern emerges regarding the usage of three-word phrases in the English compositions of Chinese secondary school students. The phrase that takes the top spot in terms of frequency is "a lot of", a common expression employed to denote quantity. This phrase appears a staggering 103 times, indicating its widespread use among students. Following closely behind is the phrase "In my opinion", which is utilized to convey personal viewpoints. It occurs 81 times, further highlighting its prevalence in students' writing. Interestingly, both of these frequently used phrases are classified as prepositional phrases.

This observation suggests that Chinese secondary school students have a tendency to incorporate prepositional phrases into their written work. This inclination is not unique to non-native speakers, as a similar trend can be observed in native language writing. When it comes to native speakers' English compositions, the most frequently occurring three-word chunk is "in order to", a phrase used to express purpose. It appears 89 times and, like the aforementioned phrases in the students' writing, is also a prepositional phrase. This parallel between the writing habits of Chinese secondary school students and native speakers highlights the significance of prepositional phrases in English composition, regardless of the writer's linguistic background.

Table 2. Statistical Table of Word Frequency of Three-Word Chunks in Chinese Secondary School Students' English Writing (Case-Sensitive)

Word frequency	Categorization of word chunks	Specific word chunks	Frequency
1	Prepositional phrases + embedded of phrase	a lot of	103
2	Other prepositional phrase	In my	81
3	Other forms	more and more	67
4	Other forms	the Spring Festival	67
5	Verb/Adjective + to Clause	I want to	62
6	Other prepositional phrase	in the future	55
7	Other prepositional phrase	for us to	52
8	Tethered verb be + noun phrase/adjective phrase	to be a	47
9	Tethered verb be + noun phrase/adjective phrase	It is a	42
10	Noun phrase + of phrase	one of the	41
11	Other prepositional phrase	in order to	40
12	Other forms	the other hand	37
13	(Verb Phrase +) that Clause	I think it	36
14	Tethered verb be + noun phrase/adjective phrase	The most important	36
15	Tethered verb be + noun phrase/adjective phrase	There are many	36
16	(Verb/Adjective +) to Clause	You want to	36
17	Noun phrases + other post-modifier	I have a	35
18	Other prepositional phrase fragments	in a word	35
19	Other forms	and so on	34
20	Tethered verb be + noun phrase/adjective phrase	is the most	34

Table 3. Statistics on word frequency of three-word chunks in native speakers' English writing (case-sensitive)

word frequency	categorization of lexical items	specific word block	frequency
1	Other prepositional phrase	in order to	89
2	Noun phrase + of phrase	the number of	66
3	Other forms	as well as	63
4	Tethered verb be + noun phrase/adjective phrase	there is a	63
5	Noun phrase + of phrase	One of the	61
6	Other forms	the United States	55
7	Other forms	more and more	49
8	Noun phrases + other post-modifier	in the US	43
9	Other prepositional phrase	in terms of	42
10	Tethered verb be + noun phrase/adjective phrase	to be a	40
11	Noun phrase + of phrase	of the world	39
12	Noun phrase + of phrase	Some of the	39
13	Other prepositional phrase	in the United	38
14	Verb/Adjective + to Clause	more likely to	38
15	Noun phrase + of phrase	the amount of	38
16	Other forms	As a result	37
17	Other forms	due to the	37
18	Noun phrase + of phrase	part of the	37
19	Noun phrase + of phrase	the use of	35
20	Noun phrases + other post-modifier	the fact that	33

3.2. The Most Frequent Word Chunk Type in Chinese Learners' English Writing was Other Prepositional Phrase Fragments, While the Most Frequently Used Word Chunk Type in Native Speakers' English Writing Was Noun Phrases + of Phrase Fragments

As indicated by Table 4 and Table 5, a clear distinction can be observed between the word chunk classification and word frequency patterns of Chinese secondary school students and native speakers in their English writing. Among Chinese secondary school students, the predominant type of word chunk is “the other prepositional phrase”, which occurs 263 times and makes up 28% of the total. This suggests that students frequently utilize such phrases to construct their sentences.

On the other hand, native speakers' English writing shows a different preference. The most commonly used word chunk type is “noun phrase + of phrase”, occurring 315 times and representing 33% of the total. This indicates that native speakers tend to rely more heavily on this particular structure to convey complex ideas and relationships between entities.

When comparing the two groups, a stark contrast emerges. In the writing of Chinese secondary school students, the frequency of “noun phrase + of phrase” is significantly lower, at only 41 occurrences, which is a mere 4% of the total. Conversely, in native speakers' writing, the frequency of other prepositional phrases is 169, accounting for 18%.

From this analysis, it is evident that while prepositional phrases are a staple in the writing of native speakers, Chinese secondary school students seem to overlook the significance of incorporating noun phrases, particularly those coupled with of phrases, into their English compositions. This disparity may point to a gap in the students' understanding or application of certain linguistic structures that are more prevalent in native speaker writing.

Table 4. Categorization of word chunks and frequency of word chunks in Chinese speakers' English writing

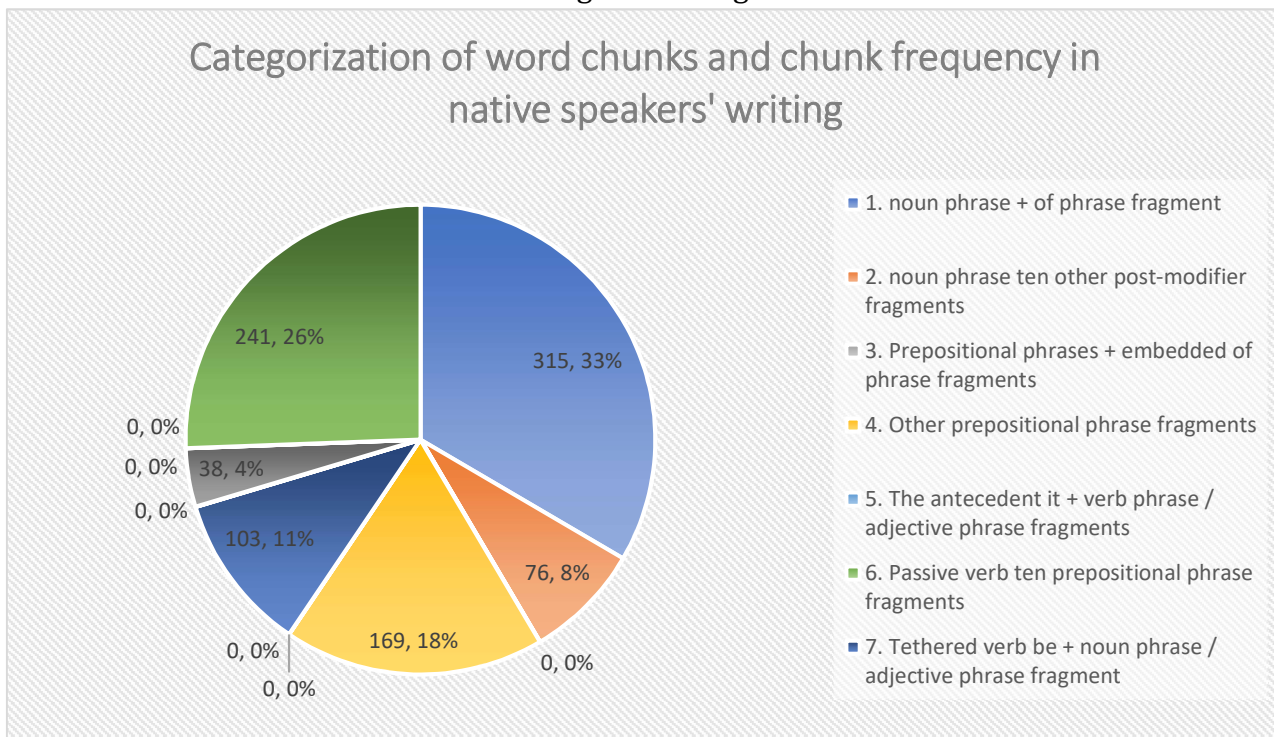
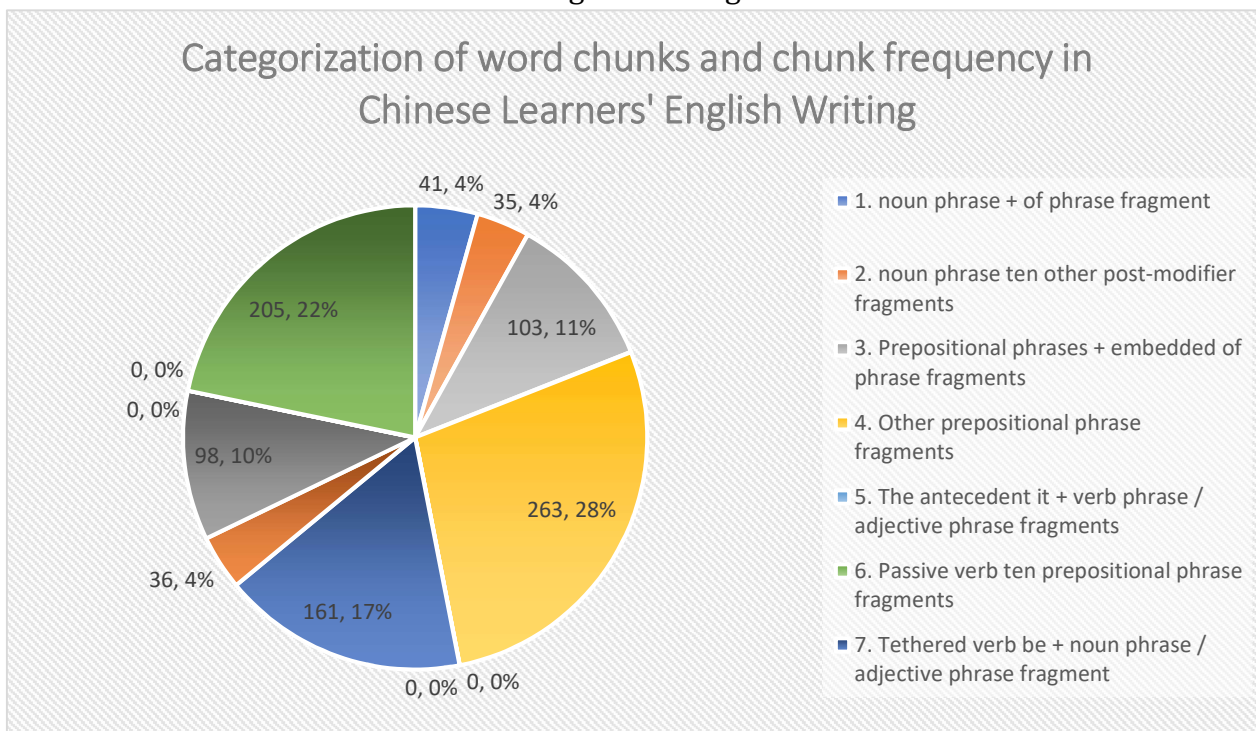


Table 5. Categorization of word chunks and frequency of word chunks in native speakers' English writing



4. Conclusion

Word chunks play an indispensable role in writing. They are the building blocks that add texture and coherence to a piece of writing, making it more engaging and accessible to readers. By incorporating a diverse array of word chunks, writers can elevate the quality of their work, ensuring that it is not only informative but also aesthetically pleasing.

In the present study, the advanced corpus analysis tool BFSU Power Conc 1.0 was employed to delve into the intricacies of word chunk usage. The "NESSI" and "TECCL" corpora served as the foundation for a comparative analysis, focusing on the frequency of three-word chunks in the English compositions of Chinese secondary school learners and native speakers. The findings revealed two distinct patterns:

Prepositional phrases emerged as the most commonly used three-word chunks in the writings of both groups. This suggests a universal reliance on such phrases to convey relationships and additional information within sentences.

Among Chinese learners, the category of other prepositional phrases was predominant, indicating a tendency to use a specific subset of these linguistic structures. In contrast, native speakers' writing featured a higher frequency of noun phrases combined with of phrases, highlighting a different approach to expressing complex ideas.

These insights have significant implications for the teaching and learning of word chunks:

Writing instruction should place a greater emphasis on the use of noun phrases, as they appear to be underutilized by Chinese learners but are a key feature of native speaker writing.

The importance of fixed expressions in writing cannot be overstated. Teachers should ensure that students are aware of and practice using these expressions to enhance the authenticity and richness of their writing.

Context is king in writing. Teachers should guide students in selecting appropriate sentence patterns based on the specific context of their writing, ensuring that the chosen structures effectively convey the intended meaning.

Accumulating a wide range of vocabulary and phrases is crucial for students. This can be achieved through regular exposure to diverse reading materials and active engagement in writing exercises.

In the realm of English writing pedagogy, teachers must be attuned to the writing context and adept at selecting expressions that resonate with the subject matter and purpose of the writing. The curriculum should prioritize the learning and accumulation of vocabulary and phrases, equipping students with the linguistic tools they need to express themselves effectively.

Moreover, regular writing practice, coupled with constructive evaluation and peer learning, should be an integral part of daily teaching. This holistic approach will not only improve students' writing skills but also foster a deeper appreciation for the art of writing.

Looking ahead, future research can build upon these findings by further refining the classification of word chunks. This will enhance the precision and applicability of the research. Given that word chunks can consist of three or more words, expanding the scope of analysis to include chunks of varying lengths will provide a more comprehensive understanding of language use. Additionally, comparing the English writing chunks of secondary school students across different grades, in addition to the comparison with native speakers, will offer valuable insights into the developmental trajectory of students' writing abilities. Such comparative analyses will shed light on the progress and challenges faced by students as they advance through their educational journey, ultimately contributing to more effective teaching strategies and improved learning outcomes.

References

- [1] ANDERSON N J, NATTINGER J R, DECARRICO J S. Lexical Phrases and Language Teaching[J/OL]. *The Modern Language Journal*, 1994, 78(2): 242.
- [2] BIBER D, BARBIERI F. Lexical bundles in university spoken and written registers[J/OL]. *English for Specific Purposes*, 2007, 26(3): 263-286.
- [3] Chen, Y. H. & P. Baker. Investigating critical discourse features across second language development: lexical bundles in rated learner essays, CEFR B1, B2 and C1 [J]. *Applied Linguistics*, 2016, 37 (6): 849-880.
- [4] Cortes, V. Lexical bundles in published and student disciplinary writing: Examples from history and biology[J]. *English for Specific Purposes*, 2004, 23 (4): 397-423.
- [5] Ding Yanren, Qi Yan. Correlation Study of Word Block Utilization in English Speaking and Writing Levels [J]. *Journal of PLA Foreign Language College*, 2005, (3): 49 -53.
- [6] Esfandiari, R. & F. Barbary. a contrastive corpus-driven study of lexical bundles between English writers and Persian writers in psychology research articles[J]. *Journal of English for Academic Purposes*, 2017, 29 (5): 21-42.
- [7] GROVE C, LEWIS M. The Lexical Approach: the State of ELT and a Way Forward [J/OL]. *TESOL Quarterly*, 1994, 28(4): 828.
- [8] Huang, K. More does not mean better: frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing [J]. *System*, 2015, 53 (3): 13-23.
- [9] Huang Kaisheng, Zhou Xinping. corpus-based study on the trend of Chinese English learners' word block output ability [J]. *Foreign Language Community*, 2016, (4): 27 -34.
- [10] Ma Guanghui. A Study of Word Blocks in English Majors' Second Language Limited Writing[J]. *Foreign Language Teaching and Research*, 2009, 41(1): 54-60+81.
- [11] Ma, Guanghui. Definition, Classification and Recognition of Word Blocks[J]. *Journal of PLA College of Foreign Languages*, 2011, 34(1): 1-4+127.
- [12] Ministry of Education of the People's Republic of China. English Program Standards for General High Schools (2017) 2020 Revised[M]. Beijing: People's Education Press., 2020.05.
- [13] WRAY A. Formulaic language and the lexicon[M/OL]. Cambridge: Cambridge University Press, 2002.

- [14] Zhang Huiping. A Study on the Characteristics of Word Block Usage in Chinese Beginning English Writers[J]. Journal of Foreign Languages College of the People's Liberation Army, 2020, 43(4): 19-25+86+160.