

An Application Research of Corpus-based Approaches to Middle School Writing from the Perspective of the Unit Topic

Yuhan Li^{1*}

¹ Foreign Languages College, Yangtze University, Jingzhou, China

Email Address

1728546355@qq.com (Yuhan Li)

*Correspondence: 1728546355@qq.com

Received: 1 September 2023; **Accepted:** 12 October 2023; **Published:** 17 November 2023

Abstract:

The teaching of unit writing under the guidance of themes is a process of cultivating students' ability to integrate unit topics, express themes using language, and showcase their thinking. The advantage of using a self-built corpus to assist in unit topic writing lies in integrating teaching resources in a more efficient and clear manner, strengthening students' input of topic chunks, and providing writing support. This article combines the current research situation in China to sort out the theoretical basis and advantages of corpora in writing teaching. It explores the integration and expansion of unit topic resources through self-built corpora and the application of Antconc, in order to strengthen students' input of topic chunks, help students summarize and summarize the corpus under the theme, and comprehensively improve their writing ability.

Keywords:

Corpus, Teaching Writing, Unit Teaching

1. Introduction

Unit topic writing is the mainstream trend in English writing teaching today. The English Curriculum Standards for Compulsory Education stipulate that the English curriculum content consists of elements such as themes, discourse, language knowledge, cultural knowledge, language skills, and learning strategies. [1] In the new curriculum standard, the theme and discourse are reclassified as the content of the curriculum, and the scope, functions, and importance of the themes that should be covered by the compulsory education English curriculum are specified. In recent years, teaching around thematic significance has received widespread attention. Wang Qiang proposed the overall teaching design of the unit under the concept of grand conception, which requires teachers to focus on specific themes, enabling teaching to set goals around a complete unit theme, guiding students to generate grand conception based on the unit theme by learning, refining, and establishing associations with different individual discourse concepts. [2] Scholars have pointed out that holistic teaching is more effective in promoting students' deep understanding and

development of subject core competencies. The overall teaching of unit themes under the theme provides a new direction for writing teaching. However, in actual teaching, teachers often find it difficult to operate and lack specific ideas, methods, and practical strategies in teaching. [3] For example, in writing teaching, teachers still separate different discourses in the unit, and separate the writing section in teaching design, which cannot draw on effective language materials in the unit's thematic context. [4] Some teachers have not established a close connection between topic-based vocabulary, reading, grammar, and writing in the module, and only ask students to memorize writing templates to solve their writing problems. This will lead to problems such as insufficient vocabulary accumulation, improper language expression, and disordered discourse structure in students' writing. [5]

Nowadays, corpus linguistics has been widely used in foreign language teaching. Corpus based research has received a large number of scholars' attention in the field of foreign language teaching. Previous studies have demonstrated the advantages of corpora in strengthening lexical input, grammar teaching, promoting writing teaching, and expanding teaching resources. In the application of corpora in writing teaching, self-built corpora demonstrate their unique flexibility and operability. It proposed that teachers should organize excellent sample essays from students, compositions at the teacher's level, sample essays from textbooks, or expert level sample essays from other sources in the process of writing correction according to the source of the test paper, genre, theme, etc., to form a sample corpus and apply it in writing teaching. [6] Zhang Yu and Xu Shichang conducted a teaching experiment combining writing rhetoric and genre analysis teaching based on genre analysis and corpus linguistics theory, using an online writing corpus and establishing a small writing corpus. [7] The article demonstrates the promoting role of sample corpus in improving students' discourse structure, writing content, and writing skills in writing. Based on corpus linguistics and the former's research on corpus in writing teaching, this article focuses on solving the problem of unit topic writing and explores a new model of applying self-built corpora more efficiently in writing teaching.

2. Advantages of Using a Self-Built Corpus to Assist in Teaching Topic Writing

2.1. Integrating Teaching Resources

Chen Yanzhu explored the overall writing teaching strategy for the unit. [8] The author pointed out that the teaching of unit topic writing should include four teaching strategies: comprehensive analysis of textbooks, organic integration of teaching content, orderly implementation of teaching activities, and effective evaluation and feedback. Guo Junbin also emphasized that writing teaching should be guided by thematic significance and conduct overall writing research on unit texts. [9] Therefore, integrating unit texts and linking writing instruction with unit texts will change the traditional form of writing instruction that exists in isolation from unit texts. Based on this ideological perspective, teachers build their own textbook corpus. When teaching writing, by searching for the thematic vocabulary of the teaching unit, relevant information in the textbook can be efficiently presented to students for review and integration of the thematic resources of the unit. This can not only help students integrate textbook knowledge, but also improve their reading ability in the process of selecting useful information. By using this method, students can draw on thematic

information from different types of text materials such as reading and listening, thereby achieving overall teaching under unit topics.

2.2. Strengthen Theme Block Input

The selection and distribution of words are closely related to the theme expression, and the choice of theme determines the use of words and students' understanding and expression of the theme. [10] In the use of lexical chunks, the use and distribution of vocabulary in students' compositions are closely related to the theme of the composition. A certain topic has a unique vocabulary group, and these words have an extraordinary recurrence rate in the corpus, indicating that students have a high consistency in their expression of a certain topic in terms of vocabulary selection. So, the teaching of writing under the theme and the input of lexical chunks under the theme are the core of teaching. In the research on corpus-based chunk teaching, many scholars have confirmed that corpus assisted chunk teaching can improve students' writing grades and abilities. For example, Long Yun and He Anping demonstrated that teachers can generate topic word lists and basic word lists through corpora, provide learners with information on word frequency and distribution, analyze the relationship between keywords and sentence levels, and promote students' writing ability. [11] Chen Bimei improved students' writing ability through corpus retrieval, such as auxiliary chunk usage induction, chunk comparison, chunk sentence making and translation, after a 12-week classroom teaching experiment. [12] However, Xu Fang found in comparing the writing of Chinese learners' papers with the corpus of international scholars' papers that the composition of language proficiency is not isolated, but inevitably rooted in various contexts. [13] Only by acquiring the inherent expressions in a specific context is the key to acquiring second language writing ability. Even if the learners have a large number of lexical chunks and use them inappropriately or out of context, their writing ability is also problematic. Therefore, based on the theory of lexical chunks, the language chunks that students need to acquire are determined through context and related topic classification, so that students can master high-frequency and conventional expressions in the context in order to more efficiently improve their writing skills. By combining corpora with themes, it is possible to promote students' acquisition of thematic chunks in a more efficient and intuitive way. That is, teachers can collect writing samples from the same theme and build their own corpus of sample articles classified by theme. By searching for keywords or specific words, students can summarize and summarize the theme blocks in the sample corpus based on the presented search results, allowing them to imitate and apply them in writing, thereby enhancing their input of theme blocks.

2.3. Provide writing support

In writing teaching, providing students with pre reading materials to provide output and content support for writing can help improve learners' writing skills and promote creative imitation of the materials. When learners are provided with pre reading materials, they repeatedly use the words and phrases from the previous text to imitate learning, which is also a form of scaffolding learning. [14] The pre reading materials provided to students through a small sample corpus have their own huge advantages. Firstly, a small-scale corpus is designed by teachers themselves, with clear objectives and targeted collection of corpora, which is suitable for the difficulty and interests of the teaching objects. Secondly, the corpus can be continuously expanded, with

openness and timeliness. Teachers classify the collected corpus according to certain principles, such as register, style, source, difficulty, etc., and continuously expand them. Thirdly, teachers annotate the thematic information of the discourse at the beginning, allowing them to analyze discourse of the same type or theme, thereby making more accurate language generalizations, which helps expand students' knowledge under the same theme. [15] At the same time, a corpus can provide students with keyword information, vocabulary frequency information, etc. of the discourse, helping them conduct in-depth analysis of the model text. Long Yun and He Anping explored the role of corpus keywords in refining the main content of a discourse, clarifying the context of an article, and outputting language expression. [12] Confirmed the objectivity, intuitiveness, and generative nature of corpora in text analysis. In addition, by analyzing words, the semantic relationships related to them can be derived, forming a semantic network. Students can obtain information about the main theme of the article, such as descriptions of location, time, and people. These high-frequency words can be searched in "Concordance" to view their distribution and conduct in-depth discourse analysis. In the process of analyzing search results, students learn how to interpret text and have a clearer understanding of the target text, gradually building a knowledge framework.

3. Application of Self-Built Corpus to Assist Unit Topic Writing

This article is based on corpus linguistics, combined with contemporary compulsory education English curriculum standards and actual secondary school teaching situations, to establish a corpus of teaching materials based on the People's Education Press for the second volume of the eighth grade of junior high school English, as well as one of the 12 sample corpora that match the theme of the junior high school unit. The main theme of the seventh unit, "Protecting Animals", is used as a case study, and the self-built corpus is applied in writing teaching classrooms.

3.1 Method of building a self-built corpus

The factors that need to be considered in establishing a corpus include the purpose of building the corpus, the variety of corpus, sampling standards, scale, representativeness, equipment, storage methods, and format. [16] The purpose of building a corpus is the most important, directly affecting the type and scale of the corpus. The establishment of a textbook corpus and a sample corpus is aimed at applying them to unit topic writing and promoting students' acquisition of topic chunks more efficiently. Firstly, teachers need to collect and organize language materials. In establishing a corpus based on topic-based classification, when selecting corpus, it is not only necessary to pay attention to the source, topic, and difficulty of the corpus, but also to consider the interests, levels, and needs of the teaching objects. The corpus should balance knowledge, timeliness, and education. [15] The sources of corpus can come from the internet, teaching reference books, excellent model essays polished by teachers, etc. Collect the corpus and convert it into txt format, classify and save it, and name it uniformly. Secondly, annotate the corpus. In order to enable users to better operate and understand the corpus, the information in the corpus should be supplemented and explained. When establishing a textbook corpus, teachers need to first scan the textbook, perform text recognition and manual proofreading, and perform text cleaning, formatting, and meta information input as required. Finally, when processing non textual parts such as images and graphics, the image content should be annotated.

3.2 Application of Self built Corpus

The unit theme of Unit 7 in the grade eight volume of the People's Education Press is "Geographic Knowledge". Under this theme, students' writing topics include small themes such as protecting the environment and caring for animals. The writing training topic for this unit is to protect giant pandas. In the unit writing teaching of this unit, the teaching process mainly consists of five parts: link 1: teacher introduction, link 2: establishing a framework, link 3: summarizing and summarizing, link 4: independent writing, and link 5: writing feedback. In step one, the teacher can briefly summarize the writing tasks of this lesson and use promotional videos on panda protection as the course introduction. In this section, students will have a comprehensive understanding of the topic and be interested in writing, gaining a preliminary understanding of information related to panda conservation. In step two, the teacher began using a self built corpus to gradually build a writing framework for students. Encourage students to integrate relevant language materials related to this topic in the textbook and prepare adequately for their writing. In link three, the teacher helps students sort out the writing framework, summarize the previous corpus, and form the outline of the composition. The fourth stage is students' independent writing, where they apply the input chunks to their own compositions and choose the appropriate corpus according to the situation to create their own compositions. During this process, the teacher provides writing guidance to students, reminding them to pay attention to various aspects such as the structure and format of the article, word spelling, and grammar usage. In the final stage, peer evaluation and teacher classroom analysis will be conducted to further consolidate the commonly used lexical chunks and sentence structure features under this topic. Among them, links two and three particularly highlight the application of self built corpora. The following are teaching examples.

The teacher will open the established corpus of the eighth grade textbook of the People's Education Press in Antconc, and search for "baby panda" and "adult panda" respectively. All relevant information related to the keyword in this unit will appear. Based on the results of searching the textbook corpus (i.e., Figure 1), students are asked to identify relevant information about baby pandas and adult pandas, circle key information, and gradually complete their own outline.

Left Context	Hit	Right Context
live up to to years. 1d Take turns telling your classmates about pandas. A: A	baby panda	cannot see. B: An adult panda weighs many times more than a baby
panda is about to kilos. 2. At birth, a baby panda is about cm long. 3. A	baby panda	is not black. It is and it has no . 4. A panda can live
Is a baby panda also black and white? Teacher: No, you would never guess. A	baby panda	is pink and it cannot see. It also doesn't have any teeth.
right. And the adult pandas weigh around 100 kilos. Boy 2: What about the size of a	baby panda?	How small is it? Teacher: At birth, it's only around 15 cm long.
Take turns telling your classmates about pandas. A: A baby panda cannot see. B: An	adult panda	weighs many times more than a baby panda. 2a Talk about pandas in
baby pandas. 2. In , pandas have become so that they are now a symbol of China. 3.	Adult pandas	do not have babies very , and some of the babies only live for
Do you have any questions about baby pandas? Boy 1: Are they much smaller than an	adult panda?	Teacher: Yes. When they're born, they only weigh about 0.1 to 0.2 kilos. Boy
How small is it? Teacher: At birth, it's only around 15 cm long. but an	adult panda	is around 150 cm long. Girl: Is a baby panda also black and white?
born, they only weigh about 0.1 to 0.2 kilos. Boy 1: Wow, that's small! I know an	adult panda	weighs many times more than a baby. Teacher: That's right. And the
every two years. The babies often die from illnesses and do not live very long.	Adult pandas	spend more than 12 hours a day eating about 10 kilos of bamboo. Many years
adult panda weighs many times more than a baby. Teacher: That's right. And the	adult pandas	weigh around 100 kilos. Boy 2: What about the size of a baby panda? How

Figure 1. Corpus of Textbooks.

The teacher selects a sample corpus with the theme of “protecting animals”, opens it in Antconc, searches for the keyword “people”, and asks students to imitate their writing sentence structures, lexical chunks, grammar, etc. (as shown in Figure 2). Teachers can design sample sentences related to students’ reading and translation, guiding them to use them in writing, such as “people are taking an important measure to...”, “people should achieve that...”, “people are working to help...”.

Left Context	Hit	Right Context
So the number of them is falling quickly At present,	people	are taking an important measure to save pandas. People
their numbers , the rats and mice multiplied quickly. Luckily , some	people	are working to help save the animals. Some groups
these clothes sell at a high price. So some greedy	people	begin to kill animals in a large quantity. This
the United States alone are considered in danger. Why should	people	care? Because we need animals, and because once they
help save the animals. Some groups raise money to let	people	know about the problem. And they try to get
friends and part of our environment. Third, we should raise	people's	awareness to protect animals and our environment. In
and meat. In order to protect our resources of ecology,	people	should realize that these of any species is at
teaches children in cities about pandas and other endangered animals They send	people	to schools to tell children about their importance of saving
present, people are taking an important measure to save pandas.	People	want to find a home for them, and recently,

Figure 2. Corpus of writing.

4. Conclusions

Under the concept of overall unit teaching, this article proposes teaching exploration for unit topic writing. Teachers use their own textbook corpus to integrate resources in the textbook and closely connect writing teaching with discourse in the unit. In addition, using a sample corpus to provide students with imitation writing paradigms and strengthen the input of topic chunks.

Corpus assisted unit topic writing has the following advantages: firstly, students internalize the learned chunk knowledge, network it through topics, and transfer knowledge in future writing related to this topic or other related topics. Therefore, unit topic writing can help students change their previously isolated, shallow, and fragmented learning methods. Under the integrated unit theme, students use the word blocks under the theme as a framework to expand their critical thinking on the topic, and produce written results that accurately understand and express themselves appropriately. Secondly, corpus is based on digital driven language learning, which is more efficient, clear, and prominent in presenting the corpus to students, with clear focus, and efficient use of modern technological resources in foreign language teaching. Finally, students use their own cognitive strategies to summarize the language structure, meaning, and usage patterns of specific language forms through observation, comparison, classification, and summarization methods. They conduct in-depth analysis of sample texts, promote synergistic effects, and narrow the gap between students and sample texts, thereby improving students' writing abilities. Overall, by using a self-built corpus to assist in unit topic writing, the resources for unit learning are fully expanded and integrated, allowing students to focus on the sentence structures, lexical chunks, and grammar involved in the topic, and produce better compositions.

Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of this article.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

- [1] Wang, M.X. Utilizing a composition corpus to assist English writing teaching. *Middle School Foreign Language Teaching and Learning*, 2014, 12.
- [2] Wang, Q. Analysis of English subject teaching design based on the concept of grandeur. *Curriculum, Textbooks, and Teaching Methods*, 2022, 11.
- [3] Wang, C. The Basic Path of Implementing Unit Integrated Teaching Based on the Concept of Discipline. *Primary and Secondary School Management*, 2021, 07.
- [4] Ma, L. Exploration of Integrated and Orderly Design Unit Teaching. *Curriculum, Textbooks, and Teaching Methods*, 2012, 02
- [5] Tong, X.H. A Practical Study on Improving the Writing Level of High School Students through Theme Meaning Based Retelling. *English Learning*, 2019, 10.
- [6] Chen, Y.J. Exploration of Middle School English Writing Teaching from the Perspective of Unit Integration. *English Teacher*, 2020, 20.
- [7] Zhang, Y.; Xu, S.C. An Empirical Study on Corpus Based Genre Analysis in Teaching English Professional Writing. *Foreign Language Journal*, 2011, 06.
- [8] Chen, Y.Z. Exploration of English Writing Teaching Strategies Based on Unit Integration. *Teaching and Management*, 2022, 07.
- [9] Guo, J.B. Exploration of Teaching High School English Unit Text Overall Writing Based on Theme Meaning. *English Illustrated Magazine (High School Edition)*, 2021, 06.
- [10] Liang, M.C.; Li, W.Z. Corpus Application Tutorial. Beijing: Foreign Language Teaching and Research Press, 2010.5
- [11] Long, Y.; He, A.P. Teaching Exploration of Corpus Keyword Assisted High School English Summary Writing. *Research on Classroom Teaching in Primary and Secondary Schools*, 2020, 12.
- [12] Chen, B.M. The Application of Corpus Assisted Chunk Teaching in English Writing Teaching, Basic English. *Journal of the School of Foreign Languages, Shandong Normal University (Basic English Education)*, 2014, 16.
- [13] Xu, F. Teaching and Research of English Writing. Beijing: Foreign Language Teaching and Research Press, 2012, 6.
- [14] Wang, Q. The Synergistic Effect of Chinese Second Language Intensive Reading and Writing Continuation. *Research on Second Language Learning*, 2020, 01.
- [15] Xie, J.C. On the Construction of Personal Teaching Corpus. *Foreign Language Electronic Teaching*, 2003, 03.
- [16] Graeme, K. Corpus Linguistics. *Foreign Language Teaching and Research Press*, 2000.



© 2023 by the author(s); licensee International Technology and Science Publications (ITS), this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (<http://creativecommons.org/licenses/by/4.0/>)