Journal of Education and Educational Research
ISSN: 2957-9465 | Vol. 4, No. 3, 2023

A Review of Research on Corpus-Based Academic English Writing

Yujia Qi

School of Southwest Petroleum University, Sichuan, China

Abstract: Academic English is the main medium of international academic communication for contemporary scholars, and it is also the main genre of English writing in graduate education. The linguistic community has also emphasized the study of combining corpus research with academic English. In order to have a comprehensive understanding of the research trends of corpus-based academic English writing in the past ten years, this study compiles the relevant research papers on corpus-based academic English writing from 2013 to 2022 and summarizes the hot issues and shortcomings based on the overview of the literature and the content of the research. The study finds that the research on corpus-based academic English writing in China mainly focuses on teaching, genre, register, and interpersonal meaning and found the two trends of the research on self-constructed corpus and interpersonal meaning. Two areas for improvement have also been identified: the unevenness and limitations of relevant research in China.

Keywords: Academic English Writing; Corpus; Review.

1. Introduction

In recent years, the rapid development of corpus linguistics at home and abroad has led to the flourishing of corpus-based research on academic English writing. The development of academic English writing also facilitates bilingual scholars to express their opinions correctly and effectively in international journals, dissertations, project proposals, and textbooks. Corpora provide rich and authentic linguistic materials for academic English writing, thus giving rise to new research dimensions and methods [1]. However, compared with foreign studies, domestic studies have not combined corpus and academic English writing closely enough, and the relevant studies need to be more comprehensive. It's necessary to investigate this area's current state of research in the past ten years. This review aims to provide an overview of recent research on corpus-based academic English writing and practical guidance for teaching academic writing. Given this, to understand the current research situation of corpus-based academic English writing, the author has analyzed the research results in China National Knowledge Network (CNKI) with the keywords "corpus," "corpus-based," "academic English writing," and "academic writing." The publication dates were set to 2013-2022. As important journals in a field can often reflect the research progress and hot trends in the field, the author set the source of literature as "CSSCI" journals and then manually eliminated the literature that is not related to the research, such as book reviews, non-English academic writing, etc., and finally obtained 26 relevant literatures. This review analyzes the retrieval literature from the literature overview and research content and summarizes the hotspots and shortcomings of the current corpus-based research.

2. Research Methodology

Given this, to understand the current research situation of corpus-based academic English writing, the author has analyzed the research results in China National Knowledge Network (CNKI) with the keywords "corpus," "corpus-based," "academic English writing," and "academic writing." The publication dates were set to 2013-2022. As important journals in a field can often reflect the research progress and hot trends in the field, so the author set the source of literature as "CSSCI" journals and then manually eliminated the literature that is not related to the research, such as book reviews, non-English academic writing, etc., and finally obtained 26 relevant literatures. This review analyzes the retrieval literature from the literature overview and research content and summarizes the hot spots and shortcomings of the current corpus-based research.

3. Research Results and Discussion

3.1. The General Publication of Relevant Research in China

The increase or decrease in the number of published literatures can reflect its development to a certain extent, as can be seen in Figure 1, the number of published literatures reached the maximum between 2015 and 2017, and since 2017, although there is core literature every year, the number is not as high as in previous years. This data is less than the number of corpus-based academic English writing papers worldwide summarized by Luo Qin, and this result of comparisons can reflect that domestic scholars do not pay enough attention to this field, and there is still a lot of room for China to continue to conduct in-depth investigations in this field [2]. From Figure 2, we can see the literature published in the domestic core journals, and the most significant number of papers is published in the journal Foreign Languages and Their Teaching, with six articles, followed by Foreign Language Word and Foreign Language Education, with 4 and 3 papers, respectively. Through organizing and analyzing the selected documents, we found that 21 documents used self-constructed corpora, and most of their corpus sources were selected from Chinese learners' English Master's theses, international journal articles, and English native language dissertations, while three documents used existing corpora, namely, MICUSP (Ädel&Römer 2012) developed by the School of English at the University of

Michigan, the Beijing Collection of Academic Research Essays (Beijing CARE) constructed by the Beijing University of Aeronautics and Astronautics, and Chinese Learner English Corpus (CLEC). In terms of analysis methods, comparative analysis is more prominent, but multidimensional analysis is also increasingly adopted by scholars.

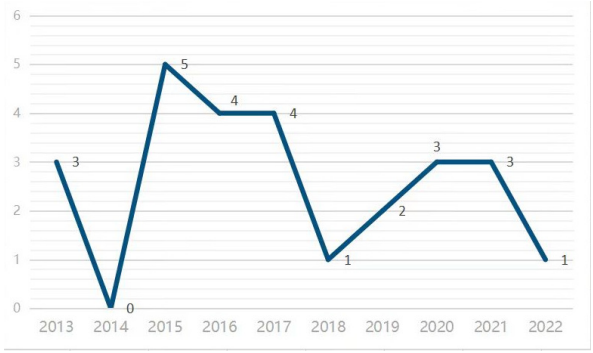


Figure 1. 2013-2022 CSSCI Journal literature trends

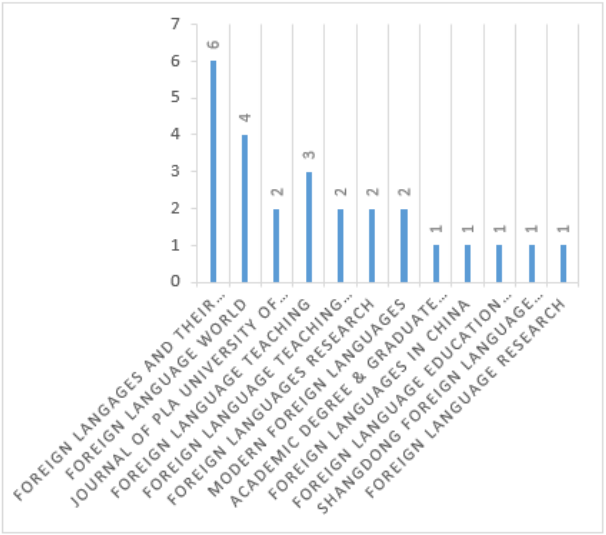


Figure 2. Number of CSSCI journals from 2013 to 2022

3.2. Studies on Corpus-based Academic English Writing

To better categorize the selected literature, the review refers to the classification methods of Coffin & Donohue (2012) and Gardner (2012). It analyzes the research contents of the 26 selected CSSCI journal articles as the basis of analysis. By defining the concepts, the review tends to classify the corpus-based academic English writing research into teaching research, discourse structure research, register feature research, and interpersonal meaning research.

3.2.1. Sub-section Headings Studies Concerning the Teaching Research

The richness of the corpus and the simplicity of searching have made corpora important to learners, researchers, and teachers. Before the prevalence of corpora, foreign language teaching was often based on the intuition and experience of educational experts, and the compilation of teaching materials was not based on authentic language materials, which was detached from the actual use of language [3]. However, corpora provide rich and authentic language materials for foreign language writing teaching, which is more conducive for teachers and learners to analyze the language in multiple dimensions and explore its essence through the surface structure of the language [4]. According to the British Academic Language Proficiency (AL) theory, research on teaching and learning can be divided into research on EAP writing teaching strategies, development, and use of instructional materials, as well as curriculum program design and evaluation. These researches are often based on experiments, and by deepening the concept of EAP to the practical level, the concept is updated and developed through continuous adjustments and experiments. Among them, there are rich research results on EAP writing teaching strategies in China, especially in examining the elements of academic writing in the EAP program or discussing writing teaching strategies and their implementation effects, covering all aspects of the writing process and outcomes. For example, Jiang Feng [5] conducted an 18-week English for Academic Purposes (EAP) experiment to empirically analyze the applicability and effectiveness of this teaching pathway in developing academic discourse competence. The second aspect examines the development, utilization, and evaluation of teaching and learning materials, including how to build a corpus and how to use corpus software to enhance teachers' teaching of academic English writing. Lou Baocui [6] explored the difficulties in graduate students' academic English writing through comparative analysis. The study found that academic English writing training can solve the problems of vocabulary collocation and pragmatic accuracy in graduate students' academic English writing and that through the indexing training of professional corpora, graduate students' linguistic awareness and rhetorical skills in writing academic English can be improved in this field. The third aspect is related to the design and evaluation of writing programs, such as the study on the difficulty of writing bilingual abstracts for Master's degree students in science and engineering. According to the British theory of academic language proficiency (AL), the three aspects of teaching and learning are relatively independent. According to the British Academic Language Proficiency (AL) theory, research in each of the three areas of study is relatively independent, focusing on either the text of writing, the process of writing practice, or both perspectives. Therefore, there is much room for integrating the three aspects of research.

3.2.2. Studies Concerning the Discourse Structure Research

Discourse structure research mainly explores the macro-structural features of various sub-discourse categories of academic discourse, as well as the structural characteristics of different parts within the discourse [7]. For example, Cao Yan and Xiao Zhonghua [8] built their comparative corpus of English abstracts of scientific and technical papers to examine the differences in the discourse communicative functions of Chinese and foreign English abstracts from five dimensions. The study shows that native speakers are more adept at using various adverbs to enhance their tone as a way of emphasizing their stance and attitude, and they like to use the first person and interactive writing. However, Chinese authors' extensive use of pre-modifiers and passive verbs makes it difficult for readers to read and understand and deviates from today's international academic writing style of simplicity. The number of published studies on language structure in China is small, and scholars mostly analyze the academic essays of L2 learners and native English academic discourse on a macro level to compare the general characteristics of Chinese and foreign academic essays as well as the disciplinary

differences between the two languages. There is still a lot of room for this kind of research, such as the discussion of specific discourses in domestic and foreign academic groups and the developmental characteristics of Chinese scholars.

3.2.3. Studies Concerning the Register Feature Research

The study of register features focuses on what lexical and grammatical forms are used in academic English writing and whether these linguistic forms are unique to the academic discourse community or common to the discipline [7]. Lou Baocui [9] examined the use of enclitic nouns in graduate students' academic English discourse and found that graduate students' awareness of the use of enclitic nouns is relatively low, the use of class symbols lacks diversity, and the frequency of standardization is significantly lower than that of native speakers. Some researchers [10] analyzed the use of reporting verbs in 112 interdisciplinary (natural sciences, social sciences/humanities) and cross-level (undergraduate, master's) research papers written by native speakers based on the framework of intervention theory. The results show that there is close to significant interaction between majors and levels. From the perspective of academic language features, there are also related studies on the phenomenon of colloquialization of academic language. Wang Li and Wang Nan [11] conducted a comparative study on the tendency to colloquialize the three grammatical items in Chinese English learners' master's and doctoral dissertations and international journal papers. The study found that compared with international journals, master's and doctoral dissertations have significant speaking tendencies, namely doctoral dissertations tend to use quasi-modal verbs, while Master's thesis tends to use negative diminutive forms, BE verb diminutive forms, and willingness/prediction quasi-modal verbs. While using obligatory quasi-modal verbs, doctoral dissertations are much more significant than Master's dissertations. The study of register features does not exist on its own. Still, it is often related to the study of interpersonal meaning, which has led to the richness and diversity of publications in this field in China.

3.2.4. Studies Concerning Interpersonal Meaning Research

From the perspective of functional linguistics, interpersonal meaning researches emphasize the sociality of language and interpersonal dynamic significance so as to seek to reveal the characteristics of interpersonal meaning, including authorship and stance. The common questions are how to deal with the relationship between authors and readers and whether and how different native language environments, second language levels, and writing experiences affect the interpersonal meaning of academic texts. In the past ten years, research on interpersonal importance in China has focused chiefly on identity construction and stance expression. Domestic scholars mainly selected domestic English master's degree theses and international journals or dissertations to examine the characteristics of authors' self-mentions by adopting a comparative approach. Considering the differences in disciplines, Wang Jingjing and Lv Zhongling [11] compared English doctoral dissertations of science and engineering students with international journals, and the results demonstrated that Chinese learners are less aware of constructing academic identities and intentionally hide their authorship. Chinese scholars have a weak sense of interpersonal communication and insufficient self-expression in the academic discourse circle, which makes the authors' participation low and easy to create an objective and detached image of academic identity. Regarding stance expression, there is also a wealth of relevant results in China. Xu Fang [7] subdivided authorial stance markers into epistemic, attitudinal, and authorial presence. In addition to examining authorial stance from the categorization, some scholars have also examined scholars' expression of academic stance in China through the use of signaling nouns, reporting verbs, and evaluative adjectives. In terms of methodology, scholars mainly adopt a comparative approach, comparing the characteristics of Chinese learners' stance expressions with those in foreign journals. It is found that Chinese learners generally tend to use fewer linguistic features to express the author's stance, which is characterized by lower objectivity, lack of critical thinking, and unclear stance, thus affecting the persuasiveness of academic texts. According to the use of the three stages of a Bachelor's degree, Master's degree, and Doctor's degree, the awareness of the author's stance is more evident in the Master's degree. In addition to the above two points, interpersonal meaning research emphasizes various evaluative linguistic resources, such as evaluation strategies, explicit or implicit evaluation, etc. Based on a large comparative academic English corpus, Bingxin Wang and Naixing Wei [12] conducted a comparative analysis to evaluate patterns and meaning in the academic writings of Chinese and Western scholars. The study shows that Chinese scholars tend to express the essence of difficulty and certainty of propositions or events. In contrast, Western scholars tend to express the essence of importance and probability. In addition, domestic research is becoming more and more detailed and complete. The grammatical forms of vocabulary examined include cased nouns, intervening markers, reporting verbs, lexical block collocation, conditional clauses, tense, and general linguistic features of the discourse. Based on the corpus, Xu Fang [13] compares intervening markers' use and development characteristics in L2 writing. The study found that student essays used significantly fewer intervening markers than journal essays but significantly more directive markers than journal essays. From a developmental point of view, the use of the three types of intervening markers showed three different trends.

4. Research Hot Spots and Shortcomings

According to the current research of academic writing based on a corpus, the self-constructed learner corpus has become a significant innovation and hot spot in research. Most of the studies use the corpus as a medium for comparative analysis of lexical features of international literature, revealing the characteristics and common problems of Chinese students' writing, providing guidance for writing teaching, and improving students' writing levels. In terms of research content, interpersonal meaning research is the mainstream of corpus-based academic English writing research, which explores the two main themes of authorial identity construction and position expression through different language forms. Corpus teaching research is gradually developing towards multi-perspectives and multi-dimensions, and teaching methods in which students are directly involved in corpus analysis and teaching research are progressively beginning to appear. In the past ten years, although the application and research of corpus in academic English writing have achieved fruitful results, there still needs to be improvement. Firstly, the development of the research

field needs to be balanced, such as teaching research related to writing program design and assessment of such research is less. Secondly, the research field is too limited, and there is insufficient attention to inter-disciplinarity in China. Although doctoral dissertations from science and engineering disciplines have appeared in recent years, the number of such dissertations is relatively small, and the research topics are also limited.

5. Conclusion

Based on the CSSCI literature of 2023-2022, this paper analyzes and summarizes the use of corpora in academic English writing and explains and analyzes it from the aspects of literature review and research content. The research on corpus in academic writing mainly focuses on four aspects, namely, teaching, language class, domain, and interpersonal meaning, and self-built corpus is the hot spot and highlight of the research in this field. At the same time, it is also found that the development in this area needs to be balanced in China. Future research should take into account the characteristics of corpus development in academic English writing and develop in a more balanced and diversified direction.

References

[1] Biber, D., Reppen, R . The Cambridge Handbook of English Corpus Linguistics. (Cambridge: Cambridge University Press, 2015).

[2] Q. Luo. Review of academic English writing based on corpus, Journal of Lanzhou Vocational Technical College, Vol. 37(2021) No.5, p.94-97+116.

[3] Y. Zhu. On application of corpus in English writing teaching, Foreign Language World, (2009), p.3.

[4] Friginal, E. Corpus Linguistics for English Teachers: Tools, Online Resources and Classroom Activities (New York: Routledge, 2018).

[5] F. Jiang. The development of China students' academic discursive competence and effect of teaching, (2021) No.321, p.06.

[6] B.C. Lou. Application of corpus approach in EAP teaching for postgraduates, Foreign Languages and Their Teaching, Academic Degree & Graduate Education, (2020) No.332, p.07.

[7] F. Xu. A critical review of academic English writing studies, Foreign Languages and Their Teaching, (2015) No.47, p.94-105+160-161.

[8] Y. Cao, Z.H. Xiao. A Multidimensional analysis model of research articles' English abstract, Foreign Language Teaching, (2015) No.36, p.06.

[9] B.C. Lou. A corpus-based study of shell nouns in the graduates' EAP writing, Foreign Language Teaching, (2013) No.34, p.03.

[10] J.W. Gao, M. Zhang, W.T. Chen, Revisiting reporting verb use from a theory-driven approach: based on advanced English-L1 learner corpus data across disciplines and levels, Foreign Language Teaching, (2022) No.39, p.05.

[11] J.J. Wang, Z.S. Lv, Study on the self-mentions in academic English writing of science and engineering, Foreign Language World, (2017) No. 179, p.02.

[12] B.X. Wang, N.X. Wei, Characteristic evaluative patterns and meanings in research articles by Chinese and Western scholars: a contrastive study, Foreign Languages and Their Teaching, (2018) No.303, p.06.

[13] F. Xu. The use and development features of engagement markers in L2 academic writing: a corpus approach, Foreign Languages and Their Teaching, (2013) No.269, p.02.