Study of Corpus' Influences in EAP Research (2009-2018): A Bibliometric Analysis in CiteSpace

Chunyan He[1] & Xueyan Wei[1]

[1] School of Foreign Languages, China University of Geosciences (Beijing), Beijing, P.R. China

Correspondence: Xueyan Wei, School of Foreign Languages, China University of Geosciences (Beijing), Beijing, 100083, China.

Abstract

Using 328 papers published in SSCI (Social Science Citation Index) journals from the Web of Science between 2009 and 2018, this study analyzed corpus' influences in EAP (English for Academic Purposes) research with mapping knowledge domains CiteSpace. This study found that the relationship between international EAP studies and corpus tends to increasingly tight in the past ten years. Corpus in EAP study mainly acts as the methodology that EAP is studied by means of self-building or existing corpora using retrieval software together with statistical analysis method. International scholars pay more attention to writing teaching, academic literacy and discourse/discourse analysis in EAP study, and the co-cited references mostly tend to the genres study of EAP.

Keywords: CiteSpace, EAP, corpus, SSCI journals, mapping knowledge domains

1. Introduction

EAP (English for Academic Purposes), as a branch of ESP (English for Specific Purposes), is one of the directions in the college English teaching reform. EAP can be divided into English for EGAP (General Academic Purposes) and ESAP (English for Specific Academic Purposes). With the development of the "Double Top" university plan, the traditional teaching model of college English has stepped into a bottleneck period. A thorough reform must be carried out to break through this bottleneck and that is precisely the chance of EAP. Consequently, a variety of study achievements related with EAP spring up, such as monographs, journals, and textbooks. In recent years, corpus-based teaching and research methods have been introduced into EAP teaching and study and have become an innovation in EAP in terms of the methodology. Therefore, it is necessary to grasp the overall development of EAP in the past ten years, to clarify the hot issues of international EAP in the past ten years, and to explore the breadth and depth of corpus-based or corpus-driven EAP research.

Using the CiteSpace Software, a mapping knowledge domains tool, this paper analyzed corpus' influences in EAP study in the past ten years (2009-2018) with quantitative and qualitative research. Through the mapping knowledge domains, the dynamic trends of this field can be presented directly and accurately.

2. Methodology

2.1 Data

All the literature reviewed in this paper is from the SSCI journals of Web of Science (*http://webofscience.com*). As search query was limited to "EAP" and "Corpus" as the topic, the number of literature was too few to do a visualization analysis. Thus, the search query was expanded to "EAP" as the topic and the timespan was from 2009 through 2018 (ten years). The research categories were refined into linguistics, education educational research, educational research, language linguistics, and education scientific disciplines. Eventually 328 journal papers were collected after filtering on August 25, 2019.

2.2 Data Analysis Tool

This paper uses the mapping knowledge domains to analyze the dynamic tendency of corpus-based EAP studies in the past ten years. CiteSpace also called citation space, as a mapping tool developed by Dr. Chen, C. at College of Computing and Informatics of Drexel University, is designed to make it easy for users to answer questions about the structure and dynamics of a knowledge domain (Chen, 2014). It is a software for visualizing,

and analyzing bibliometric data; therefore, it is used in Domain Map Analysis, Bibliometric Analysis, and Scientometrics, and it can also be used to help researchers find research topics and write their literature review (Li, 2017). The software version used in this paper is CiteSpace5.4.R4.

3. Results and Discussion

3.1 Publishing Situations of Journal Papers

The total publications of a certain field, one of the effective indexes for the development of the discipline, show the degree of maturity and vitality in this field (Xiao, 2017). Figure 1 shows that the overall growth of EAP studies in the international SSCI journals has tended to steady in the past four years. In 2012, the number of publications is in a sharp increase of 4 percent.
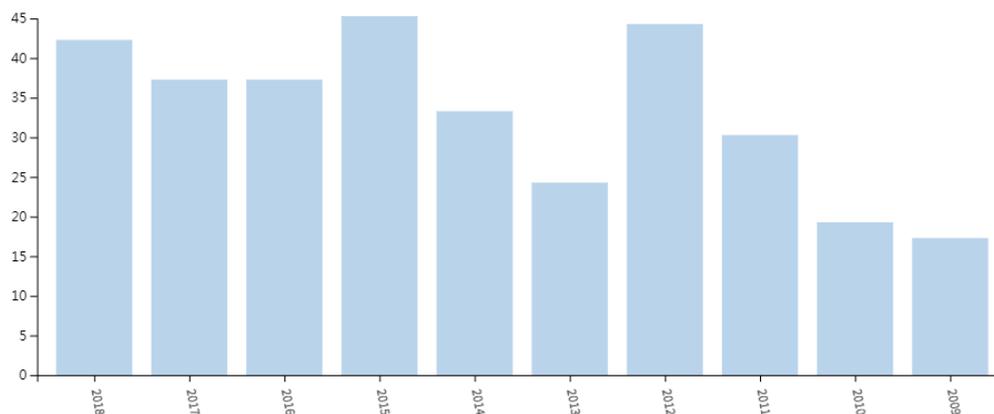

Figure 1. Annual Publications in EAP Studies (2009-2018)

The analysis of the published journals can not only understand the spatial distribution characteristics of EAP studies, but also master the core journals in this field, so as to provide a basis for literature collection and management, and provide important information references to others (Xiang, 2015). It is not difficult to find from Figure 2 that papers in EAP studies in a decade are mainly published in the Journal of English for Academic Purposes (138), English for Specific Purposes (34), System (19), Journal of Second Language Writing (17), etc. From the characteristics of these publications, the Journal of English for Academic Purposes (IF (Impact Factor) =1.732) and English for Specific Purposes (IF=1.704) cover all aspects of EAP and ESAP. System (IF=1.930) mainly studies educational technology and foreign language teaching (Applied Linguistics), while Journal of Second Language Writing (IF=4.2) focuses on the teaching and learning of L2 (the second language) writing, with high impact factors. Therefore, EAP studies mainly focuses on the field of applied linguistics, especially foreign language teaching, and also includes researches on educational technology in teaching, including the corpus-based EAP studies. For example, Larsen-Walker and Melissa wrote in the paper *Can Data Driven Learning address L2 writers' habitual errors with English linking adverbials?* in 2017 that studied the influence of DDL (data-driven learning) in *student*s to use conjunctive adverbs correctly in their essays with help of the Michigan Corpus of Upper-level Student Papers (MICUSP). In 2017, Crosthwaite and Jiang built a diachronic mini-corpus with 200,000 token consisting of student course reports and papers for one semester to study whether EAP would influence the academic stand in writing of L2 learners.
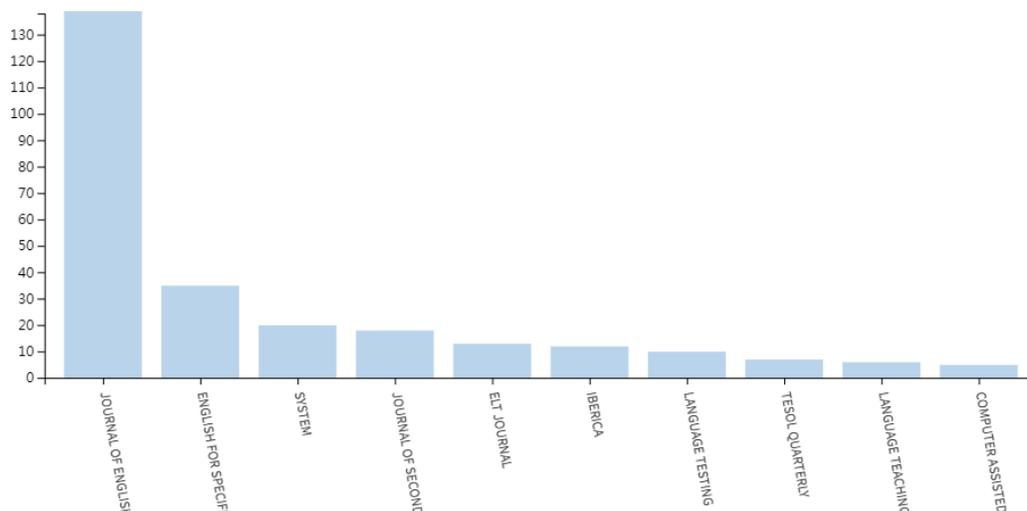
Figure 2. Major journals in EAP studies (2009-2018)

3.2 Publishing Situations of Authors

Through analyzing the authors of the selected papers, the number of published papers in the decade shown in Figure 3 is obtained. In terms of the total amount, the authors of the International SSCI Journals on EAP studies in the past decade mainly include Atai (7), professor of the University of Karachi, Iran. Engaging in the field of applied linguistics, he is the editor of Iranian Journal of Applied Linguistics now. He meanwhile is a high-yield author of international EAP studies field who almost published one paper each year in the International SSCI Journal during these ten years. The other author is Hyland (6), the professor of Applied Linguistics at the University of East Anglia, the internationally well-known applied linguist, and honorary professor at the University of Hong Kong and the University of Warwick. He is a leading scholar in academic discourse and L2 writing. There are also 6 authors who have published 4 papers respectively, including Crosthwaite, who mainly studies corpus-based EAP studies, Deroey, Katrien L.B., who is passionate about discourse analysis, especially discourse analysis of academic lectures, Hamp-Lyons, Liz, who focuses on the assessment of EAP, Wette, who studies the empirical research of EAP, James, who tends to the learn transfer in second language acquisition, as well as Tribble, who concentrates on the theoretical research in EAP and the study of teaching methods.
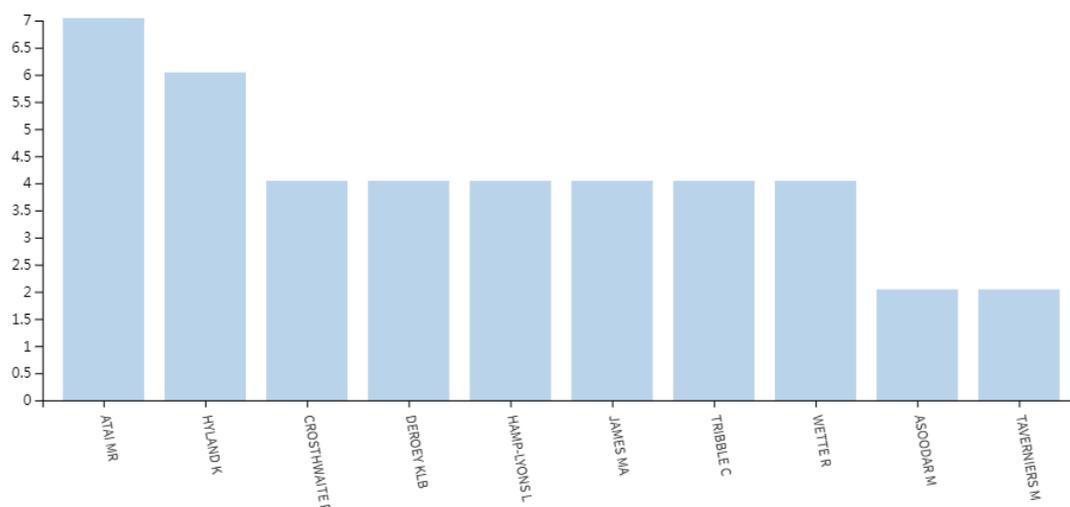

Figure 3. Publishing situations of authors (2009-2018)

3.3 Study of Cited Situations

The sizes of the nodes (dots) in the mapping of co-cited authors, references, and journals represent their cited frequency (Li, 2017: 28). The analysis of co-cited authors can identify highly-cited authors in a certain field, and therefore also can generate researcher clusters or groupings to identify research topics and the distribution of their subject areas (Li & Chen, 2016: 170). The analysis of co-cited references can not only identify high-impact references in a certain field, but also explore the tendency and hotspots (Xiang, 2015). Furthermore, the analysis of co-cited journals can identify the main sources of knowledge in a certain field (Li & Chen, 2016: 170), promoting learners and researchers to find their study focuses while reading or submitting papers.

3.3.1 Situations of Co-cited Authors

According to the co-cited authors network (Figure 4), scholars with highly-cited frequency from 2009 through 2018 mainly include: Hyland (175), the leading scholar in contemporary academic discourse and L2 writing; Swales (141) and Dudley-Evans (48), the proposer of genre analysis theories on academic discourse; Biber (63), the corpus linguistics and grammarian; Flowerdew (62), the expert in discourse analysis; Coxhead (43), the founder of The Academic Word List (2000); Leki (40), the expert of L2 writing; Halliday (39), the functional linguist; and Nesi (35), the main builder of BASE (British Academic Spoken English) and BAWE (British Academic Written English) corpus and the expert of discourse analysis in EAP. These cited authors are internationally renowned scholars in the fields of EAP and corpus linguistics. Figure 4 shows that there is only one co-cited cluster. This cluster mainly consists of Hyland, Swales, Biber, Dudley-Evans, Charles, Leki, Coxhead, and Simpson, who engaged in EAP teaching, discourse analysis, and corpus linguistics studies, which suggests that EAP studies are more closely related to applied linguistics with more cross-citations. In addition, there are many red citation burst in Figure 4, among which the scholars with high cited burst index are Hamp-Lyons, Thompson, Schmitt, Read, Wingate, Atai, Storch, Ivabic, and Braine. Their papers cited mainly focus on the field of applied linguistics and higher education, which indicates that the direction in EAP studies has turned to the field of higher education in recent years.
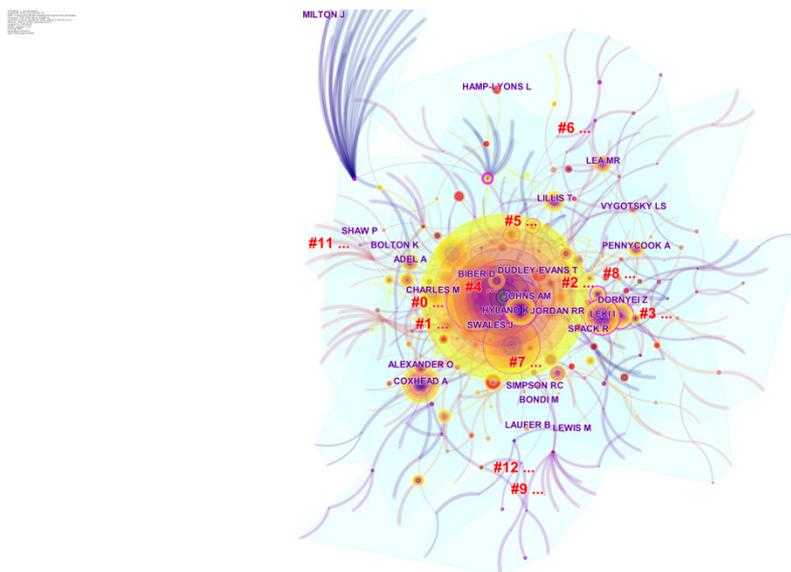

Figure 4. Network of Co-cited authors and their cluster in International SSCI (2009-2018)---Co-Cited Authors

According to the silhouette value, Figure 5 shows that the core research areas of EAP include: EAP (#10), intercultural communication (#12), institutional language policies (#11), cognitive processing (#9), thesis and dissertation writing (#4), learner conceptions (#7), advanced student writing (#0), L2 student (#5), models and modeling (#2), newspaper discourse (#6), teacher actions (#3), paraphrase (#8), as well as coherence (#1).
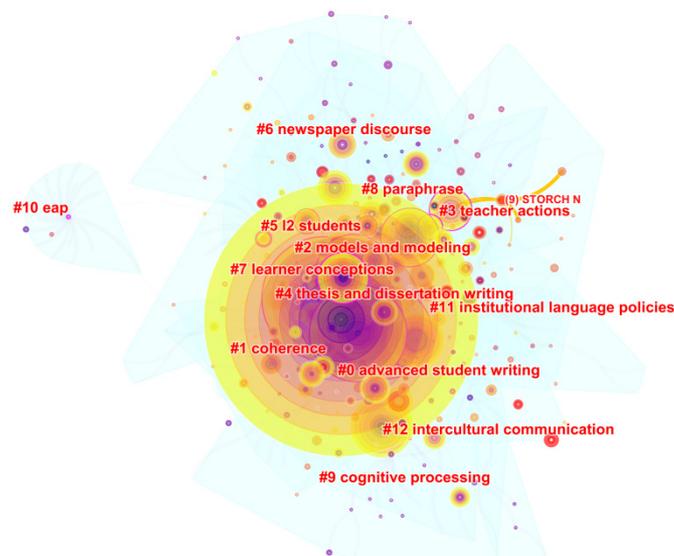
Figure 5. Network of Co-cited Authors and their cluster in International SSCI (2009-2018)---Research Field

3.3.2 Situations of Co-cited References

The network of co-citation references in Figure 6 indicates the authors and cited frequency of highly-cited references from 2009 through 2018 are Nesi. (2012) (21), Hyland. (2006) (15) (2008) (14) (2007) (10), Swales. (2004) (13), as well as Simpson-Vlach. (2010) (10). According to the years of these highly-cited references, most cited references are the current EAP research achievements rather than the classical literature. For example, the monograph written by Swales. in 1990 *Genre Analysis: English in Academic and Research Settings* is not the highly-cited work. His high-cited works are written in this century, especially *Research genres: Explorations and applications* published by Cambridge University Press in 2004, which ranks the fourth in the citation list. The top highly-cited reference of Nesi. is Genres across the Disciplines: Student Writing in Higher Education, co-authored with Gardner, the theme of which is corpus-based interdisciplinary genre study. Based on the analysis of the co-cited references, the current focus of EAP studies is genres as well as teaching and learning, and the methodology focuses on corpus-based method.

Furthermore, referring to clustering situations of these cited references with topics in Figure 6, three major matrices are displayed: the first matrix is located at the upper left of the Figure where there are these scholars as Evans, Lillis, Hyland, Swales, and Harwood. who mainly engage in discourse analysis in EAP. The second matrix is located at the upper right, headed by Bibe. who mainly engages in corpus linguistics and grammar. The third matrix, the largest one, represented by Swales, Nesi, Hyland, Gardener, Tardy, Chen, and Liu, concentrates on writing teaching and educational technology. It is noticeable that Hyland, spans both first and third matrices. This matrix shows that corpus linguistics and the corpus-based approach play an important role in EAP studies. In addition, some red dots called citation bursts in Figure 6 can be found. A citation burst, Chen (2014: 28) explained, provides evidence that a particular publication is associated with a surge of citations, i.e. an extraordinary degree of attention from its scientific community. Furthermore, if a cluster contains numerous nodes with strong citation bursts, then the cluster as a whole captures an active area of research, or an emerging trend (Chen, 2014: 29). Here, more than one author appears the strong citation bursts displayed in the third matrix of Figure 6. They are Hyland, Swales, Nesi, Gardener, Wingate, Simpson-Vlach, and Hirano, among whose papers in applied linguistics and higher education. It indicates that the active area of EAP studies has been English teaching research of higher education in recent years.
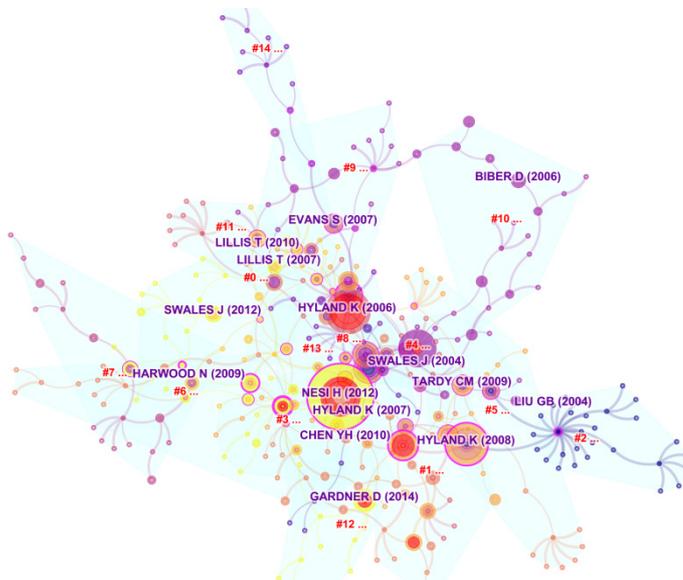
Figure 6. Network of Co-cited References and their Cluster in International SSCI (2009-2018)

3.3.3 Situations of Co-cited Journals

Understanding situations of co-cited journals can help scholars make better choice while reading literature and submitting their papers. Figure 7 is a network of co-cited journals and their cluster in international SSCI from 2009 through 2018. It shows that the highly-cited journals mainly are J ENGL ACAD PURP, TESOL QUART, ENGL SPECIF PURP, APPL LINGUIST, J SECOND LANG WRIT, SYSTEM, WRIT COMMUN, ELT J, LANG LEARN and MOD LANG J. From the perspective of clustering, there is no obvious camp. The research fields of these highly-cited journals are mainly applied linguistics, L2 writing under educational technology (corpus-based) as well as L2 acquisition. Finally, there are some cited burst journals highlighted with the red circle, including THESIS, HIGHER ED RES DEV, COMPUT ASSIST LANG L, ACAD DISCOURSE, ACAD LISTENING RES P, INTRO FUNCTIONAL GRA, J APPL LINGUISTICS and AILA REV (21 journals in total). From these active cited journals, the research direction of EAP studies in recent years has turned to educational technology, discourse analysis, academic listening, and grammar studies and so on. These study directions will become a new trend of EAP studies, including corpus-based research in educational technology.
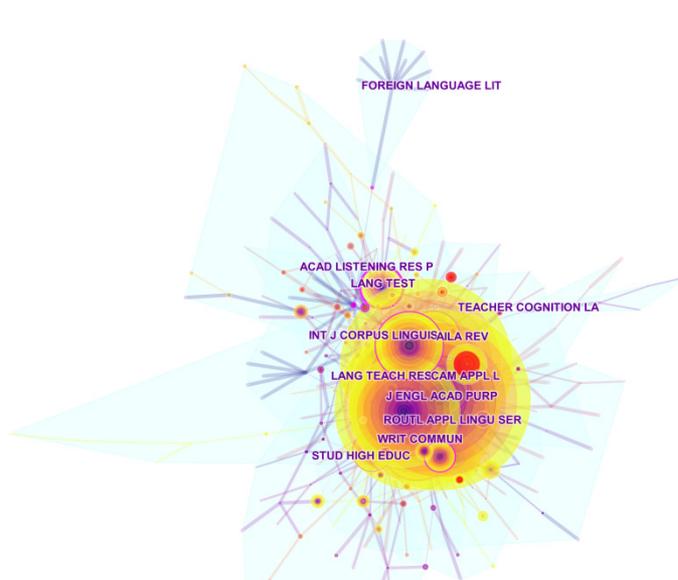

Figure 7. Network of Co-cited journals and their cluster in International SSCI (2009-2018)

3.4 Analysis of Hotspots and Trends in EAP Studies

The analysis of frequencies of topics and keywords can contribute to explore the hotspots and trends in EAP studies. Figure 8 shows the hotspots of EAP studies in international SSCI journals from 2009 through 2018 from the macro perspective. Besides the word "EAP", other four words locate in the middle of the network, including "English", "student", "language" and "academic writing". Their cross-shaped areas are several times those of other keywords, which reflects that they are the hottest research ranges in EAP. Moreover, other EAP studies also include discourse and academic literacy.
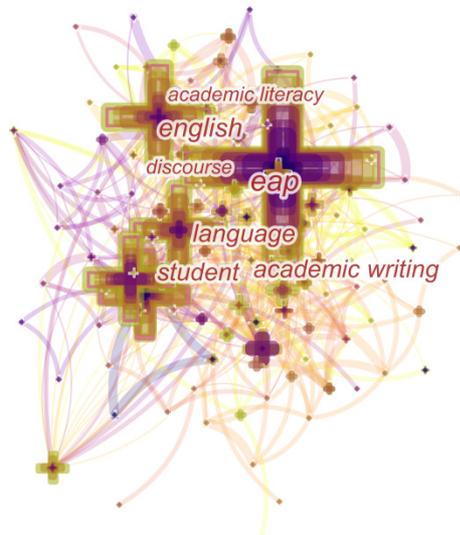

Figure 8. Hotspots and trends in EAP studies (2009-2018) (Macro)

While zooming in on the Figure 8, a more detailed study tendency chart in EAP is displayed in Figure 9. Figure 9 unfolds almost 20 hotspots of EAP studies in international SSCI journals from 2009 through 2018 from the meso perspective, such as corpus, learner, perception, applied linguistics, corpus linguistics, proficiency, literacy, etc. Therefore, whether corpus linguistics or corpus as a tool has occupied an important place in EAP studies for the recent ten years.
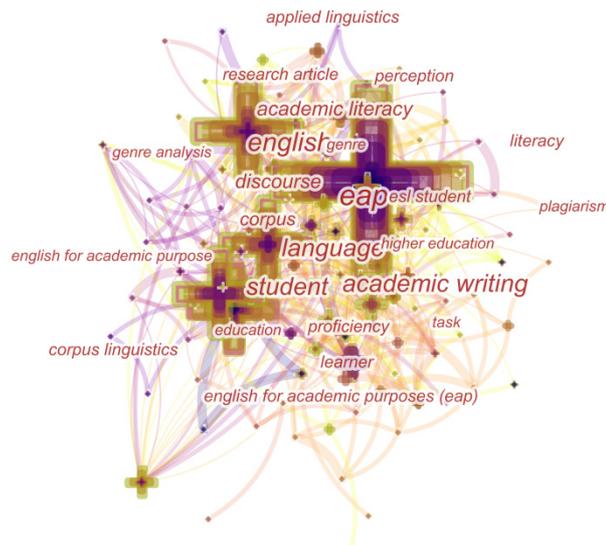

Figure 9. Hotspots and trends in EAP studies (2009-2018) (Meso)

The burstness of term reflects the active degree or an emerging trend of a topic at a specific duration (Li & Chen, 2016: 10). By clicking on Citation/Frequency Burst in CiteSpace, some surging terms present with red crosses. The surging topics include academic literacy in 2011 and academic writing in 2012. It shows that academic literacy and academic writing have been the new emerging trends in EAP studies in recent years, especially academic writing, which has become an indispensable hot topic. Corpus linguistics as a theory has entered into EAP studies since 2009, and has become a hotspot in this field by 2012, next to academic writing, discourse analysis and academic literacy. Therefore, corpus linguistics indicates new trends of EAP studies in theory and methodology.

4. Conclusion

Using CiteSpace software, corpus whether as a tool or a theory, its influences in EAP research from 2009 through 2018 are explored from the perspectives of publication, citation and research hotspots and trends. This study indicates that the relationship between international EAP studies and corpus tends to increasingly tight in the past ten years. That is to say, corpus' influences in EAP research have surged with the development of EAP study field. Corpus in EAP study mainly acts as the methodology that EAP is studied by means of self-building or existing corpora using retrieval software together with statistical analysis method. International scholars pay more attention to writing teaching, academic literacy and discourse/discourse analysis in EAP study, and the co-cited references mostly tend to the genres study of EAP.

Through the dynamic research on international literature, some suggestions about EAP study and teaching are proposed as follows.

First, EAP studies are not limited to college English. It can be studied more detailed according to the different disciplines and language learning skills so as to develop the study in depth.

Second, the corpus in EAP studies mostly lingers over the usage as a tool. In future, more research can be done from the perspective of corpus-driven and with the bottom-up induction to expand the field of corpus linguistics.

Third, experts and scholars in China major in either corpus or EAP studies, but rare scholars like Hyland, K and Nesi, H. who specialize in both fields and make large achievements. This is one of the developing goals of language researchers. Facing the diversified society, language researchers' fields specialized should also be diversified to make the breakthrough in the existing achievements.

Acknowledgments

References

Chen, C. (2014, December 13). *The CiteSpace Manual (Version 0.96)*. Retrieved Auguest 23, 2019, from http://cluster.ischool.drexel.edu/~cchen/citespace/CiteSpaceManual.pdf

Li, J. (2017, July 19). *CiteSpace Handbook in Chinese*. Retrieved Auguest 24, 2019, from http://blog.sciencenet.cn/blog-554179-1066981.html

Li, J., & Chen, C. (2016). *CiteSpace: Text Mining and Visualization in Scientific Literature*. Beijing: Capital University of Economics and Business Press.

Xiang, M. (2015). Pragmatic Studies: A Mapping-Knowledge-Domain Analysis. *Journal of Foreign Languages, 38*(06), 36-47.

Xiao, Y. (2017). A Study of International Advances in Research on Pragmatics (2006-2015). *Foreign Language Teaching and Research, 49*(05), 699-709+800.

Copyrights