Visualization Analysis of Academic Writing

Yu Li
School of Foreign Studies, East China University of Political Science and Law
Shanghai, China
liyuecupl@sina.com

Keywords: Academic writing, Visualization analysis; CiteSpace.

Abstract. The study on academic writing has been attracting increasing interest in the fields of education, linguistics, psychology, sociology, etc. Using the articles in SSCI and A&HCI databases in the Web of Science TM Core Collection between 2009 and 2018 as the data, this present paper employs the visualization software of CiteSpace 5.3.R4 as the research tool to analyze the mapping knowledge domains of academic writing. With the presentation and discussion of the general features, keywords, and cited references of publications, the development and the trends of academic writing are shown. The research results suggest that academic writing has become an independent discipline integrating with theories and methods in other disciplines; the more specific topics such as writers’ identity, gender have become the research hotspots; genre analysis, sociological analysis and cognitive analysis have become the main research perspectives in academic writing.

1. Introduction

Normally, academic writing refers to the writing activity with the purpose of academic communication, including book review, research proposal, literature review, curriculum paper, thesis proposal, research report, dissertation, journal paper, conference paper, and so on and so forth. The prominent distinction between academic writing and general writing is often considered to lie in whether the literature material is employed or not. Since the campaigns of “Writing across the curriculum” and “Writing in the discipline” in the 1980s, academic writing has arisen the great enthusiasm in relevant fields and numerous achievements have emerged. Academic writing is an effective way to break the limit of physical conditions to make scholars’ voice heard and understood. As a branch of writing research, research on academic writing focuses on academic discourses and writing activities. With its own research objects, research methods and research groups, academic writing has become an independent discipline.

Based on the idea of Big Data, this present paper explores the great value of literature. The core of big data lies in not only the large amount of data, but also the transfer of the ways of thinking and computing. Three essentials should be noted here: rather than samples, all the data are needed; rather than accuracy, the efficiency matters; rather than the causal relationship, the relevance is the focus.

2. Literature review

The study of academic writing has a history of over half a century since the publication of Barber’s paper on measurable characteristics of science prose. During this period, from the perspectives of cognition, society, culture and discourse, researchers have been exploring how to develop the academic writing ability, covering the topics of academic writing environment, writing process, textual features and the relationships among them. Generally, researches of academic writing focus on both the theoretical studies and empirical studies.

The theoretical assumptions on academic writing began from the mid-20th century. Based on cognitive writing model and the output model of written language, Paltridge et al. put forward the second language academic writing model. This model suggests that, just like the general writers,
the academic writers need to experience the cognitive process to produce the products. Another important assumption is Murray et al.’s academic writing structure model, which consists of the nine influencing factors: writing triggers or blockers, writing facilitators, writing prohibitors, engaging in efforts to write, environmental moderators, individual moderators, achieving effective writing outcomes, intrinsic rewards and extrinsic rewards [4]. Comparing the two models, we find that they have different focuses because of different perspectives and theoretical basis. The former puts more emphasis on the inner factors of academic writing, the interactions among the cognitive processes and their functions on texts, while the external factors are ignored. On the contrary, the external factors are stressed in Murray et al.’s model, whose efforts are given to the interactions of the extrinsic processes and the rewards from writing. These two theories are complementary and the combination of them will be beneficial to the understanding of the features of academic writing.

The empirical studies of academic writing are carried out in the three aspects of writing environment, writing process and text analysis. The studies on writing environment involve the interactions among reader consciousness, writing purposes and texts [5], critical thinking in academic writing [6], the relationship between plagiarism and writers’ education background [7], writing motivation [8], writers’ self-recognition [9], writing tasks and syllabus [10]. Studies on writing process mainly focus on writing strategies and the sub-process of writing, such as the integrative writing strategies [11], the single writing strategies [12] and the means of writing [13]. The text analysis of academic writing mainly focuses on the linguistic features, such as metadiscourse [14], directives [15], adverbial markers [16], lexical bundles [17], genre [18], etc.

Overall, the theoretical studies of academic writing have presented multi-perspectives. Although these theories are based on empirical studies to some extent, more empirical support will be necessary for the complexity of academic writing per se. Various writing environments and writing groups should be taken into consideration in future discussion. There are several implications we can get from empirical studies of academic writing. Firstly, academic writers should improve their consciousness in society and culture and know about the expectations from discourse community and disciplines. Secondly, writers had better know more about intertextual strategy to avoid plagiarism which is often unintentionally caused because of the improper borrowing from other sources. Thirdly, writers, especially second language writers, should master not only the suitable linguistic expressions, but also the proper way to present writers’ identities. Last but not the least, the qualitative research method and the integrative method should be attached more importance in the future.

3. Research methodology

3.1 Research questions

The study attempts to answer the following three questions:

1) What are the general trends of the research field of “academic writing” in the recent decade?
2) What are the citation hotpots in the land of “academic writing” publications?
3) Which theories and methods have influenced the development of academic writing?

3.2 Research instrument

The research instrument employed in this study is CiteSpace 5.3.R4 designed by Chaomei Chen as a tool for progressive knowledge domain visualization [19]. CiteSpace is a Java application for analyzing the scientific literature to get the development of a field in a visualizing way and the interrelationship between knowledge [20].

The source of the input data is the citation indexes of SSCI and A&HCI in the Web of Science TM Core Collection. The SSCI is an authoritative citation database in social sciences with over 3000 leading academic journals across more than 50 disciplines. A & HCI stands for Arts and Humanities Citation Index and covers more than 1,700 arts and humanities journals.

With the topic of “academic writing” in “Basic Search”, the time span from the year of 2009 to 2018, and the document type of “article”, we finally get the search results of 4,632 articles. Then, we
download and save the full records and cited references of all the articles with the file format of plain text. After that, we input the data into the CiteSpace software to gain the development and the trends in the field of academic writing.

3.3 Research procedures

1) The annual publications in academic writing will be analyzed to show the general trends and people’s interest in the study of this field;
   2) The co-occurring keywords will be presented to see the research hotspots by counting their frequency and the connections among them;
   3) The co-citation references will be used to know about the representative literature and the development of this field by depicting the citing and cited references in this field.

4. Results and discussion

4.1 General trends and features

![Annual publications in the field of "academic writing" (2009-2018)](image)

The statistics (Figure 1) suggest that the publications are growing gradually during the recent ten years despite some fluctuations. The annual amount of articles is over 300 and it reaches the peak of 550 in the years of 2016 and 2018. It means that academic writing has received extensive attention and massive achievements have emerged.

The distribution of these articles in different journals tells that a large amount of the literature was published in the journals which are devoted to language teaching and learning. Besides, some journals in the fields of linguistics, psychology and communication are also involved, which suggests that the research of academic writing is interdisciplinary.

4.2 The visualization analysis of research hotspots in academic writing

The little space-occupying keywords are often the essence of an article for its representing of the themes. The frequently used keywords are often the hotspots in a certain field. In this part, the network and highly ranked co-occurring keywords will be presented by this software.

On the CiteSpace interface, the “time slicing” is set as “from 2009 to 2018”, the “year per slice” as “1” and the “node types” as “keyword”, with other default values unchanged. The knowledge mapping of the co-words is illustrated in Figure 3 and the top 20 ranked keywords from the operating results are shown in Table I.
Table I. The top ranked keywords of academic writing (2009-2018)

<table>
<thead>
<tr>
<th>No.</th>
<th>Keywords</th>
<th>Frequency</th>
<th>No.</th>
<th>Keywords</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>student</td>
<td>394</td>
<td>11</td>
<td>academic achievement</td>
<td>153</td>
</tr>
<tr>
<td>2</td>
<td>academic writing</td>
<td>324</td>
<td>12</td>
<td>science</td>
<td>138</td>
</tr>
<tr>
<td>3</td>
<td>education</td>
<td>284</td>
<td>13</td>
<td>literacy</td>
<td>135</td>
</tr>
<tr>
<td>4</td>
<td>language</td>
<td>218</td>
<td>14</td>
<td>achievement</td>
<td>128</td>
</tr>
<tr>
<td>5</td>
<td>higher education</td>
<td>205</td>
<td>15</td>
<td>discourse</td>
<td>121</td>
</tr>
<tr>
<td>6</td>
<td>performance</td>
<td>189</td>
<td>16</td>
<td>identity</td>
<td>114</td>
</tr>
<tr>
<td>7</td>
<td>English</td>
<td>179</td>
<td>17</td>
<td>skill</td>
<td>109</td>
</tr>
<tr>
<td>8</td>
<td>writing</td>
<td>173</td>
<td>18</td>
<td>university</td>
<td>109</td>
</tr>
<tr>
<td>9</td>
<td>knowledge</td>
<td>170</td>
<td>19</td>
<td>gender</td>
<td>108</td>
</tr>
<tr>
<td>10</td>
<td>children</td>
<td>158</td>
<td>20</td>
<td>academic literacy</td>
<td>104</td>
</tr>
</tbody>
</table>

The top 20 keywords are listed in Table I by citation counts. It is shown that the topics closely related with academic writing include the study on writing process and writers’ personal characteristics (“students”, “children”, “identity”, “gender”); the study on the assessment and evaluation of writing (“performance”, “achievement”, “skill”); the study on writing contexts (“knowledge”, “science”, “literacy”, “language”, “English”, “university”); and the study on writing instructions (“education”, “discourse”, “instruction”, “higher-education”). These keywords reflect the focuses and hotspots in the field of academic writing.

Other aspects concerning the keywords in Figure 3 and Table I can also be observed. “Student” and “academic writing” are the super-ordinate terms referring to the general categories of the topic. “Literacy” and “academic language” stress the academic competence and the social practice. “Literacy” is also one of the major academic perspectives in “higher-education”, which is also in the list of the top ranked keywords. In recent years, there is a trend for the study of literacy into the elementary education, hence “children” in the list. “Identity” is involved with the research topic in discourse analysis. “Discourse” implies that the study on academic writing puts the basic theory and perspective on its agenda. Compared with other keywords, “identity” and “gender” indicate that the research theme gradually go into a deeper and more detailed way.

4.3 The document co-citation network of academic writing

On the CiteSpace interface, the “Node Types” is set as “cited references” with other parameters unchanged. The operating of the software results in Figure 4. The top 10 ranked cited references by citation counts are presented in Table II.
Table II. The top ranked cited references by citation counts

<table>
<thead>
<tr>
<th>No.</th>
<th>Frequency</th>
<th>Cited references</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>70</td>
<td>Vygotsky, L. S. 1978. Mind in society: The development of higher mental process.</td>
</tr>
</tbody>
</table>

In the theoretical framework of discourse analysis, Hyland classifies the research methods of academic writing into three categories: textual analysis, contextual analysis and critical analysis [21]. Textual analysis includes genre analysis, corpus-based analysis and multimodal analysis. Contextual analysis covers sociological analysis, socio-historical analysis and anthropological analysis. Critical analysis consists in critical discourse analysis and academic literacy. Most of the research methods are shown in the cited reference map.

The purple circles are used to signal the importance of the references. The sizes of the nodes represent the frequency. The bigger the node, the more frequently cited the reference. Combining Figure 4 and Table II, we can clearly see that Swales’ two monographs *Genre analysis: English in academic and research settings* and *Research genres: Explorations and applications* are the most frequently cited references. In the former book, Swales demonstrates how to use a genre-centered approach to develop the teaching of academic writing. In his latter book, Swales reassesses genres’ producers and consumers and the contexts in which they occur. Genre and genre analysis have always been hot topics in academic writing. Most traditional studies often employ the research method of genre analysis to explore the schematic structures of different genres and generic structure potentials of different disciplines. Since Swales’ framework of move analysis, studies on the moves of different components of thesis (such as abstract, introduction, conclusion, acknowledgement, etc) never stop. The continuing studies on the “Move-Step” structure cover the following topics: the explorations on the discourse patterns of specific disciplines such as linguistics, economics, biology, medicine, news, laws, higher education, etc.; the move analysis of certain parts in some academic papers; the genre analysis of the distinctions in academic articles written by first language writers and second language writers.

In the mid-1990s, there appeared socio-cultural view on writing, where writing was viewed as a social, cultural and high-risk activity. This trend is reflected in Hyland’s *Disciplinary discourses: Social interaction in academic genres*, which looks at academic writing as an interactive and cognitive project. Besides that, more and more attention has been paid to genre-based instruction in academic writing in recent years. The static analysis based on genre plays an important role in learners’ communication for its focus on language features and the reasons behind them. In fact, genre has become a strong tool for non-native language learners to know about the textual features of academic papers [22].

Following Hyland’s work is Lave and Wenger’s *Situated learning: Legitimate peripheral participation*. In the early 1990s, the anthropologist Lave and Artificial Intelligence expert Wenger worked together to build up the theory of “Communities of Practice”, which is distinguished from traditional learning theories in the aspect that learning should be a naturally developing process in a real communicative environment where learners can actively participate. The application of this
theory can be shown in three different ways. Firstly, on the macro-level, based on their discourse modes, language learners can be classified into different socio-cultural groups to verify their community relationships and identity, which is helpful to assess learners’ writing competence in a relatively objective way. Secondly, in the process of writing training, a series of feature descriptions in this theory can show how successful language learners gradually become “core members” from “peripheral members” by stimulating their speech act modes. Lastly, the theory helps to differentiate writing groups, writing levels, writing goals in different genres.

In Longman grammar of written and spoken English, Biber and his colleagues describe the syntactic phenomena of English in detail. The work is based on corpora of different registers: transcribed conversation, fiction, news and academic prose. The grammar phenomena are illustrated by the examples from the corpora. It is an ideal tool for the serious language users in case there are any questions of subtle aspects of English grammar and lexis. There is no doubting that this work has been an indispensable reference for the analysis of both spoken English and written English.

Number 6 in this list is Lea & Street’s article Student writing in higher education: An academic literacies approach discussing the issues of student writing in higher education by outlining a framework of “academic literacies”. This research takes into consideration the important topics such as writing requirements, writing across courses, plagiarism, feedback, etc. The themes of students, student-tutor interactions and the institution are also emphasized. A fresh perspective is provided to present debates about what is “good writing” and what is “poor writing”.

The next highly ranked reference is Ivanić’s Writing and identity with the subtitle The discoursal construction of identity in academic writing. Just as the title suggests, the book discusses several important issues between writing and identity by adopting case studies: discourse and identity, literacy and identity, identity in academic writing, academic community membership, the sense of self and the role of reader in the discoursal construction of writer identity.

In the 1970s, the revolution in the field of writing teaching and researching made people rethink the nature of writing, which was then regarded as a cognitive activity and a process of finding out meaning. Since the 1980s, cognitive psychology developed quickly, which provided a completely fresh teaching concept. Among various models in cognitive psychology, Vygotsky’s theory of “zone of proximal development” in Mind in society: The development of higher mental process, which was based on children’s I. Q. competence, was paid special attention. This theory holds the view that cooperation idea should be fully applied in classroom teaching, which helps to lessen the anxiety in writing process, enhance language learners’ interest in writing, fulfill the competence development steps from “painful writing” to “happy writing”, and finally to “academic writing”.

In The psychology of written communication, Bereiter and Scardamalia raised the knowledge-telling model and knowledge-transforming model to make the difference between the high-level writing and low-level writing. The knowledge-telling model reflects human’s natural ability, which is naturally gained from daily life. The knowledge-transforming model reflects human’s problem-solving ability, which is achieved after long-term and special training. They hold the view that the purpose of writing is to communicate successfully for writing experts, but to keep down a certain topic from memory for the green hands.

The last one in the list belongs to Halliday. As one the various theories in discourse analysis, Halliday’s An introduction to functional grammar shows that his Systemic Functional Linguistics becomes an outstanding one. Halliday’s theory is the one of language in use, which tells how texts construct meaning, and how language resources are systematically and functionally organized. The main purpose of the systemic functional grammar is to look at the grammatical system as a resource for making meaning.

These top ranked references have shown the trends of the study on academic writing. With different focuses and perspectives, the nature of writing is understood in a deeper and fuller way. Considering these top ranked references and other literature in the recent ten years, the development of academic writing can be summarized in these three aspects.

1) The study of academic writing in the perspective of systemic functional linguistics. These studies have revealed the language features and interpersonal functions of academic discourses,
which have shown the common features of academic discourse community and the disciplinary distinctions, genre distinctions and the discourse differences between second language learners’ writing and expert writers’ writing. These research fruits can be further divided into three aspects. Firstly, the essential analysis framework has been created, employed and revised. For example, in the study on the interpersonal meaning of academic discourse, there exist various theories and research approaches, among which the Appraisal Theory has always been preferred by the discourse analysts for its comprehensive and systemic analytic framework. Hyland (2005) puts forward the concepts of stance and engagement [23] and shows the distinction of the interpersonal meaning in academic discourses in social science and natural science. Secondly, the explored objects are increasingly integrate. For example, there are the vocabulary and grammar (such as the academic glossary, shell nouns, adverbs, clusters, cluster collocation, conditional clause, voice, tense, etc.), interpersonal features (such as stance markers, engagement markers, self reference, metadiscourse, literature citation, and various evaluative resources) and genre structure (such as macro structure of academic discourses, book reviews etc.). Thirdly, the research perspectives have developed from macro to details. The focus has been transferred from the discourse features of all the disciplines to the branch genre in a specific discipline. For example, in the study of genre structure, early studies are concerned with the discourses’ macro structure, then they begin to go into the structure potentials of the inner discourse, and recently they only explore a single move.

2) Studies on writing practice in the view of academic competency. Since the mid-1980s, studies in writing has been transferred from the result-oriented to the process-oriented, and then the post-process-oriented, which has pushed forward people’s concern with social-cognitive process of academic English writing and teaching. Firstly, studies on the social-cognitive process of academic writing cover the case analysis of the writing process from the diachronic perspective. For example, Cheng (2008) follows a student’s writing literacy process reveals how a second language learner develops the consciousness of academic criticism, genre structure and writing rhetoric competency [24]. Secondly, there are studies focus on the issue of international publication. Some of them have revealed the complexity and strategies to publish in international journals when the writers are from the non-English speaking countries where English is used as an additional language. Other studies have shown the paper quality and the strategies of these writers. Thirdly, the teaching researches on academic writing mainly include the teaching strategies of writing, the development and employment of teaching measures, and the design and evaluation of curriculum. These studies have shown that the teaching idea of EAP (English for Academic Purposes) has been applied into teaching practice and developed and revised in the continuing experiment. Most studies are concerned with the writing teaching strategies in EAP, which is involved in every element in writing process and writing achievement. Among the studies on the development and evaluation of teaching method, corpora are a research hotspot, which includes how to build a corpus and how to use the corpora to improve the teaching of academic writing.

3) Studies on theories and methods of academic writing. They cover the following six aspects. Firstly, a series of important concepts have been properly defined, such as English for academic purposes, academic English competency, discourse community, writers, authors, etc. Secondly, some descriptive analysis framework has been developed and revised, such as hedges, the self representation of the writer’s voice, the evaluation categories in academic articles, etc., which facilitate the empirical research on academic writing. Thirdly, some research methods have been evaluated. For example, Flowerdew (2005) reviewed the criticisms against corpus-based approaches and put forward an integration of corpus-based and genre-based approaches to text analysis [25]. Fourthly, scholars in this field interpret the teaching principles and ESP learning models in academic writing from various perspectives. Fifthly, scholars reflect on the interpretation of some concepts and discuss the controversial issues in researching and teaching. Finally, the teasing of the development of academic writing has been carried out from the perspectives of systematic functional linguistics, cross-cultural rhetoric, social-cognitive theory, social contractivism, social politics, etc.
5. Summary

The knowledge mapping analysis based on big data thinking provides a sound and strong tool and a detailed and complete data basis to know about the development and hotspots in studies on academic writing. The study has found that 1) publications on academic writing in the recent decade has shown an upward trend with a few fluctuations between years, which suggests people’s increasing interest in the study of this field; 2) the hot areas of study cover writer’s identity, gender and the writing instruction both in higher-education and elementary education; 3) the frequently cited references suggest three primary research methods of academic writing: genre-based analysis; sociological analysis and cognitive analysis.

Future studies on academic writing can be furthered from the following aspects: relevant studies from the perspective of systemic functional linguistics should take into consideration the influence of social context on writing practice rather than the over-emphasized focus on text; the integration of the qualitative study of discourse analysis and the qualitative study of corpus-based approach can benefit from the systemic interpretation of the former and the descriptive advantage of the latter; when cross-cultural rhetoric theory is employed to interpret the cultural differences in language use in academic writing, the multiple and dynamic culture should be taken into consideration rather than the single and static culture; it is necessary to place academic writing into the broader research field of second language writing, which has transferred from the disciplinary nature to the interdisciplinary nature and even the transdisciplinary nature. In a word, the basic way to develop the academic writing is the multiple perspectives and the integration of various disciplines.

References


