
RESEARCH ARTICLE

A Scientometric Review of Syntactic Complexity in L2 writing based on Web of Science (2010-2022)

Song Ruimin

College of Foreign Languages, Ocean University of China, Qingdao, China

Corresponding Author: Song Ruimin, **E-mail:** ruiminsongstu@163.com

ABSTRACT

As an important construct in the field of second language teaching and assessment, syntactic complexity is closely related to the language proficiency and language development process of L2 learners. Using the visualization software of CiteSpace, this study conducts an in-depth scientometric analysis of 140 articles on written syntactic complexity published over the past 10 years (2010-2022), thus uncovering the current development and challenges faced by relevant studies. Specifically, a frequency analysis was firstly administrated to describe the overall development in written syntactic complexity research. Furthermore, the current study conducted a Document Co-Citation Analysis (DCA), which enables researchers to conduct a network of co-cited references to identify the underlying research hotspots and future trends. The results indicate that the study concerning automatic essay scoring is the most prominent cluster active from 2010 to 2021. In addition, Norris & Ortega (2009) is the most cited paper, followed by Ortega (2003) and Biber et al. (2011). Meanwhile, the bursts of detected papers demonstrate that McNamara et al. (2012) and Grant & Ginther (2000) generated the strongest citation burst with a burst strength of 3.14 and 3.09, respectively. The findings of the study would have implications for subsequent research on written syntactic complexity in the field of language teaching and language learning.

KEYWORDS

CiteSpace visualization analysis, syntactic complexity, L2 writing

ARTICLE DOI: [10.32996/ijllt.2022.5.1.3](https://doi.org/10.32996/ijllt.2022.5.1.3)

1. Introduction

Over the last few decades, the importance of syntactic complexity in the second language (L2) writing has been recognized, as evidenced in those studies that investigated the relationship between syntactic complexity and the quality of L2 writing text (Barrot & Gabinete, 2021; Jagaiah et al., 2020; Lu & Ai, 2015; Mohamadi Zenouzagh, 2020; Qin & Uccelli, 2020).

Syntactic complexity, also known as syntactic maturity, was defined by Ortega (2003) as the diversity and sophistication of syntactic structures in language production. Furthermore, Bulté & Housen (2012) made clear the content and boundaries of syntactic complexity by introducing the concept of language complexity. According to them, syntax complexity, a crucial dimension in language complexity, mainly involves the richness and elaborateness of syntactic structures. Syntactic richness refers to the diversity and change of syntax forms in different levels of phrase, clause, and sentence, while syntactic elaborateness refers to the sophistication of various syntactic forms in these three levels. Furthermore, to a certain extent, written syntactic complexity reflects L2 learners' ability to store and produce syntactic knowledge in the brain. Therefore, it is regarded as a gauge of L2 learners' language proficiency and the quality of L2 writing.

Over the past few decades, there has been a surge of studies investigating the relationship between syntactic complexity and the quality of L2 writing. Most researchers have carried out synchronic studies to examine how syntactic complexity reflects the quality of L2 writing with the change of different independent such as first languages of the writer, L2 proficiency, writing task, etc. A number of researchers have carried out cross-sectional studies to investigate the extent to which different syntactic complexity

measures reliably index L2 writers' proficiency (e.g., Kim & Nam, 2019; Kyle & Crossley, 2017). There are also some longitudinal studies that track the development or change of syntactic complexity over time (e.g., Bulté & Housen, 2014a; Kyle et al., 2021; Mazgutova & Kormos, 2015; Menke & Strawbridge, 2019; Yoon & Polio, 2017).

All these studies have provided useful insights into how the construct of syntactic complexity should be conceived and utilized in L2 writing research and pedagogy (Lu & Ai, 2015). However, given the multi-dimension of written syntactic complexity and the disparity in research objects, research methods, or research designs, there's always a lack of consistency in the relationship between syntactic complexity and L2 writing quality. Therefore, several reviews tried to uncover the relationship between syntactic complexity and L2 writing quality (Frantz et al., 2015; Jackson & Suethanapornkul, 2013; Jagaiah et al., 2020). However, they have the limitation of narrow breadth and scope (Chen et al., 2009). For instance, Jagaiah et al. (2020), from 36 publications in the literature, examined how syntactic complexity measures (SCMs) varied by genre, grade level, students' writing ability, and writing quality. However, because of the small quantity and the lack of relative high-quality studies, it is hard to reach a consistent definition of SCMs, thus struggling to find the connection behind these 36 studies.

In view of this, making use of bibliometrics and the CiteSpace software, based on a sample of 131 articles on syntactic complexity between 2010 and 2022 in the WOS corpus, this research conducts data mining and analysis of researches on syntactic complexity in terms of their annual publication count, highly cited studies, core scholars, and research hotspots in this field. Furthermore, based on these results, the primary goal of the current study is to systematically examine how written syntactic complexity is varied by different impacting factors. The results will have implications for L2 writing and pedagogy.

2. Methodology

2.1. Data source

The sample sources of data selected in this study are listed in Table 1. The data were completely selected from Core Collection (2010–2022) in the Web of Science database. Web of Science (formerly ISI Web of Knowledge) is the premier research platform for information in the hard sciences, social sciences, arts, and humanities. And Web of Science Core Collection is the world's most trusted publisher-independent global citation database (Zhang, 2020).

Two rounds of data search were conducted in the Web of Science Core Collection on 16th December 2021. In the first round, the period was set from 2010 to all years in 2022. The topic search of "syntactic complexity" OR "syntactic maturity" is mainly used to cover recent research status in the field of linguistics. The search scope was restricted to the linguistics category to ensure accurate results, and document type was limited to "Article" and "Review". There were 847 publications from linguistics journals from 2010 to 2020. The publications were from several different branches of linguistics, such as psycholinguistics, cognitive linguistics, applied linguistics, and etc. Therefore, to unveil the relationship between syntactic complexity and L2 writing, those that have nothing to do with this topic were excluded from the rest of 847 publications. Through a systematic reading of both title and abstract of these articles, only 140 articles were further considered and analyzed. The data were subsequently downloaded from the Core Collection of Web of Science into text-formatted files.

Table 1. Data Source

Data Sources	Web of Science (WOS) Core Collection
Search Method	Topic Search
Time Span	2010-2022
Document Category	Article; Review
Search Quantity	140

2.2. Data analysis

In this article, we capitalize on CiteSpace software (Version 5.7) to analyze core scholars, co-cited references and research hotspots in the field of syntactic complexity. CiteSpace, an information visualization analysis program, is widely used to scientifically analyze the potential knowledge in certain fields (Chen, 2004; Chen, 2016; Chen et al., 2010). It has been widely used in that the knowledge graph it draws can capture the evolution of a certain knowledge field by virtue of documents on citation node and co-citation cluster labels as research frontiers (Aryadoust & Ang, 2021; Lim & Aryadoust, 2021).

On the whole, two rounds of data analysis were implemented. The first round of frequency analysis revealed the number of written syntactic complexity articles published annually, the names of the journals wherein the papers were published, and the names of the most productive authors and universities/institutes wherein the authors were residing when the papers were published. In the second round of data analysis, we conducted a document co-citation analysis to explain the co-citation relationship among keywords and references, thus tracing the trend of written syntactic complexity.

3. Results

3.1 Descriptive statistics

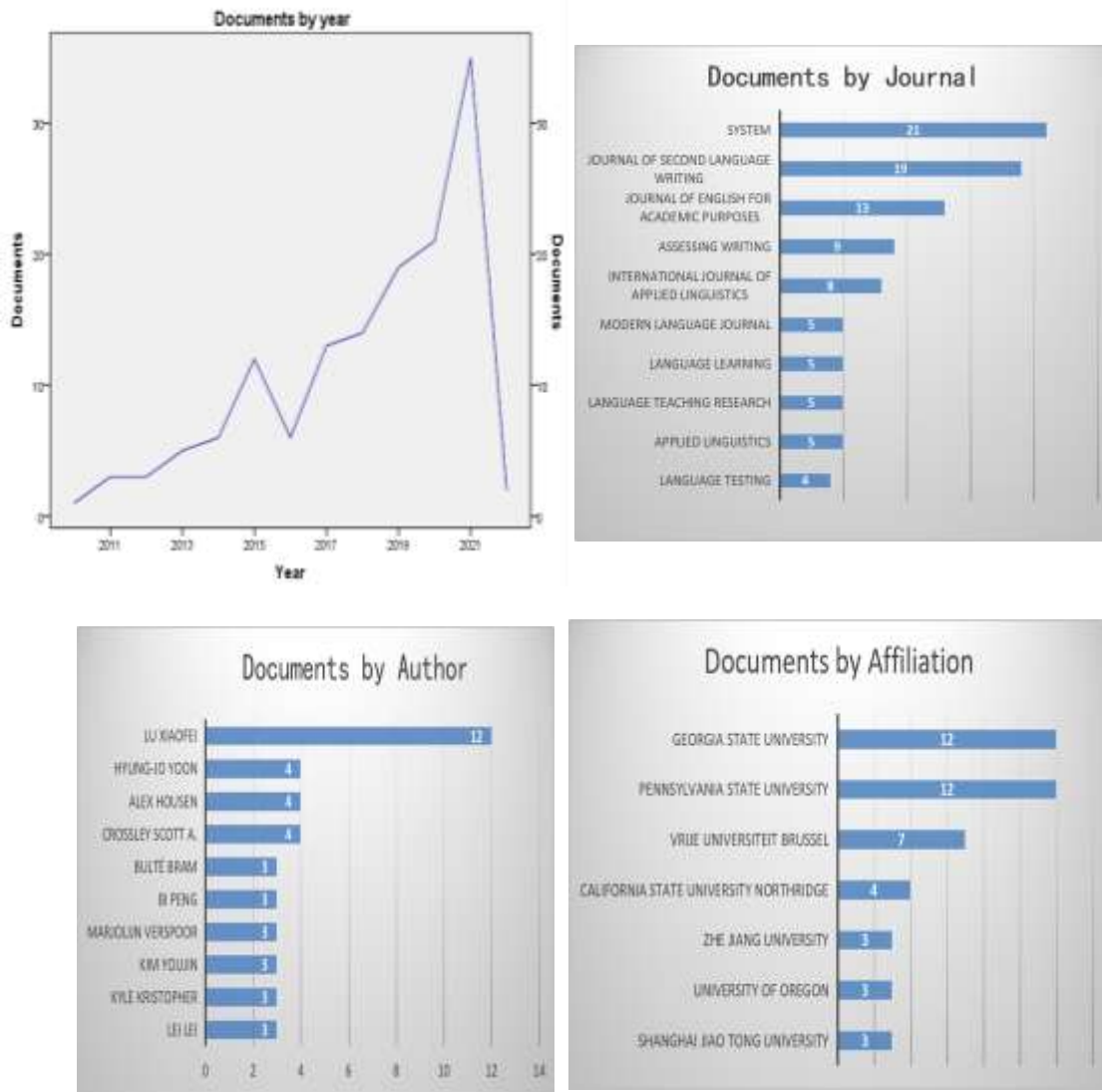


Fig. 1 Frequency Analysis of 140 Articles on Written Syntactic Complexity

Fig. 1 shows the frequency analysis results of 140 articles on written syntactic complexity. Before discussing the information behind the first figure, it should be noted that, given the time limitation, the first figure cannot cover all the articles published in 2022. Overall, the first figure indicates that the annual publication counts of documents covering a time span from 2010 to 2022 appear to show exponential growth, especially from 13 articles in 2017 to 35 articles in 2021.

The outputs of documents by the journal are illustrated in the second figure of Fig.1. These articles on written syntactic complexity were carried in 45 journals, which meant that the study on written syntactic complexity decentralized in various research fields. According to the second figure, *System* ranked the top in the number of published articles (21), followed by *Journal of Second Language Writing* (19) and *Journal of English for Academic Purposes* (13).

In terms of the most productive authors, the third figure illustrates that Lu Xiaofei was at the top of the publication list (12 articles), followed by Yoon Hyung-Jo, Housen Alex, Scott A. Crossley (4 articles). As for the affiliations wherein the authors were residing, the fourth figure shows that Georgia State University and Pennsylvania State University topped the list with 12 articles, while Vrije Universiteit Brussel came in second with 7 articles.

3.2 Document co-citation analysis

As a widely-used visualization tool in scientometric analysis, Document Co-Citation Analysis (DCA) enables researchers to conduct a network of co-cited references to identify the underlying research hotspots (Chen, Ibekwe-SanJuan, & Hou, 2010; Lim & Aryadoust, 2021).

3.2.1 The knowledge mapping for keywords

The paper's keywords may highly reflect the scheme and thought of the paper, and it is feasible to determine the research hotspots in the field of discipline according to the frequency of the keywords (Ma et al.,2020). Fig. 2 shows the knowledge mapping for keywords in the written syntactic complexity research field. Overall, the modularity Q index ranges from 0 to 1, and the mean silhouette ranges from -1 to 1. Higher modularity Q index indicates higher reliability, and a higher mean silhouette indicates better homogeneity and vice versa (Lim & Aryadoust, 2021). In the current study, the modularity Q index and the weighted mean silhouette metric for the DCA network were 0.4842 and 0.7743, respectively, indicating a highly acceptable level of reliability and homogeneity for the network. As shown in Fig. 2, excluding those small clusters that had little or no connection and those clusters that are not clearly labelled, a total of 8 major clusters with keyword labels were automatically mined by CiteSpace Version 5.7 software, which is academic writing (cluster #1), complexity (cluster#2), syntactic language development (cluster#3), teaching materials (cluster#4), linguistic complexity (cluster#5), automatic essay scoring (cluster#7), syntactic elaboration (cluster#8), and education (cluster#9).

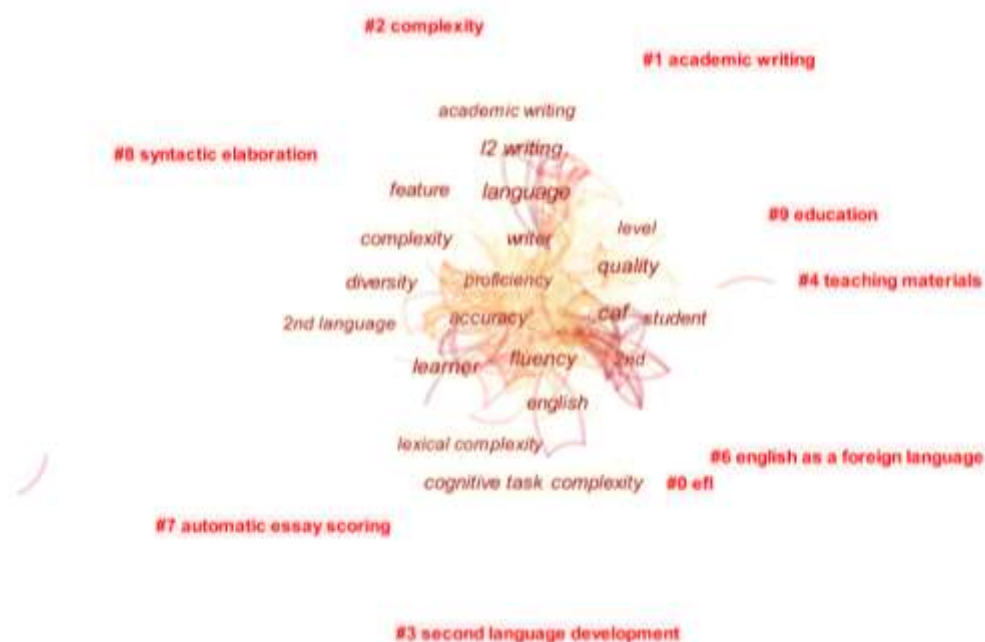
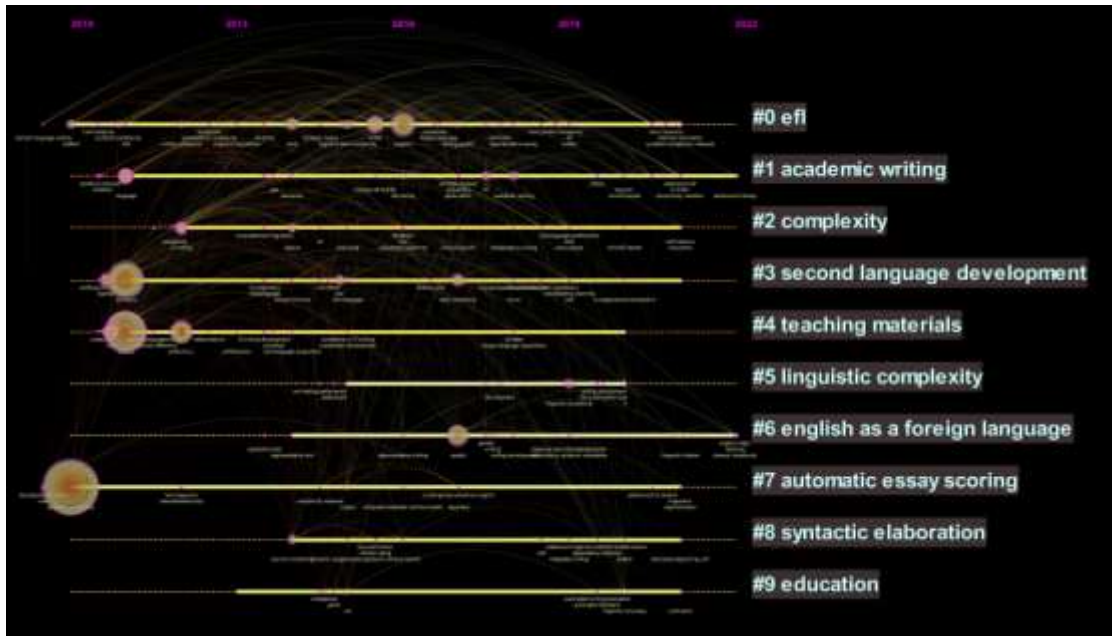


Fig. 2 knowledge mapping for keywords (modularity Q index = 0.4842; weighted mean silhouette metric = 0.7743)

On this basis, the current study drew a timeline of keywords through the co-citation analysis. Figure 3 presents the major clusters in the data on a horizontal line to provide an overview of the development of the research on written syntactic complexity over time (modularity Q index = 0.4842; weighted mean silhouette metric = 0.7743). Overall, the strength of the connection between nodes is presented as the size of the outer purple rings of each node (Aryadoust & Ang, 2021). As illustrated in Fig. 3, the longest cluster depicted in the timeline view from 2010 to 2022 was cluster 7, which was labelled as "automatic essay scoring" and was active from 2010 to 2021. The following large nodes in clusters were cluster 4 (teaching materials) and cluster 3 (second language development), implying that the publications within these two clusters were highly cited. Furthermore, the timeline view also provided chronological information on the duration of activities of each cluster in that the length of each line represents the lifetime of the cluster (Chen, 2016). For instance, cluster 2 (i.e., academic writing) and cluster 7 (automatic essay scoring) had the longest active duration of approximately 11 years. In contrast, cluster 5 (linguistic complexity) had the shortest lifetime of approximately 5 years from 2015 to 2020.

Fig.3 Timeline visualization from 2010 to 2022 (Modularity Q index= 0.4842, mean silhouette= 0.7743).



3.2.2 The knowledge mapping for co-cited reference

Fig. 4 represents the knowledge mapping for the co-cited references network (2010-2022) in the field of written syntactic complexity. Normally the importance of references was analyzed with the citation frequency, centrality, and explosiveness of reference. The five top-cited articles in written syntactic complexity are shown in Table 2.

As shown in Figure 4 and Table 2, the most cited paper, which examined challenges faced by the current practices in the measurement of syntactic complexity, proposed that measurement practices in relation to CAF must become considerably more organic and sustainable, thus shedding new light on the subsequent research on the construct of syntactic complexity (Norris & Ortega, 2009). The

The second most cited paper is Ortega (2003), systematically investigating the relationship between L2 proficiency and measure of syntactic complexity in college-level L2 writing and finding that the relationship between L2 proficiency and L2 writing syntactic complexity was not static. Instead, it varied systematically across factors like second or foreign language learning context and the methods of defining and rating the level of L2 proficiency. Meanwhile, Ortega examined four measures of syntactic complexity which can better detect between-proficiency differences and proposed an observation period needed for substantial changes in the syntactic complexity of L2 writing. As the third most cited paper, Biber et al. (2011) challenged the validity of using T-units and clausal subordination to assess the development of grammatical complexity in L2 writing, especially in academic writing, and proposed a radically new approach that is most strongly characteristic of advanced academic writing. Lu (2010) put forward a computational system for automatic analysis of syntactic complexity in second language writing, which provided technical support for the automatic analysis of L2 writing text. Bulté and Housen (2014) ascertained the measures of syntactic complexity, which can sensitively reflect the short-term changes in English L2 writing proficiency. To sum up, all these highly cited papers give impetus to the development of studies on written syntactic complexity.

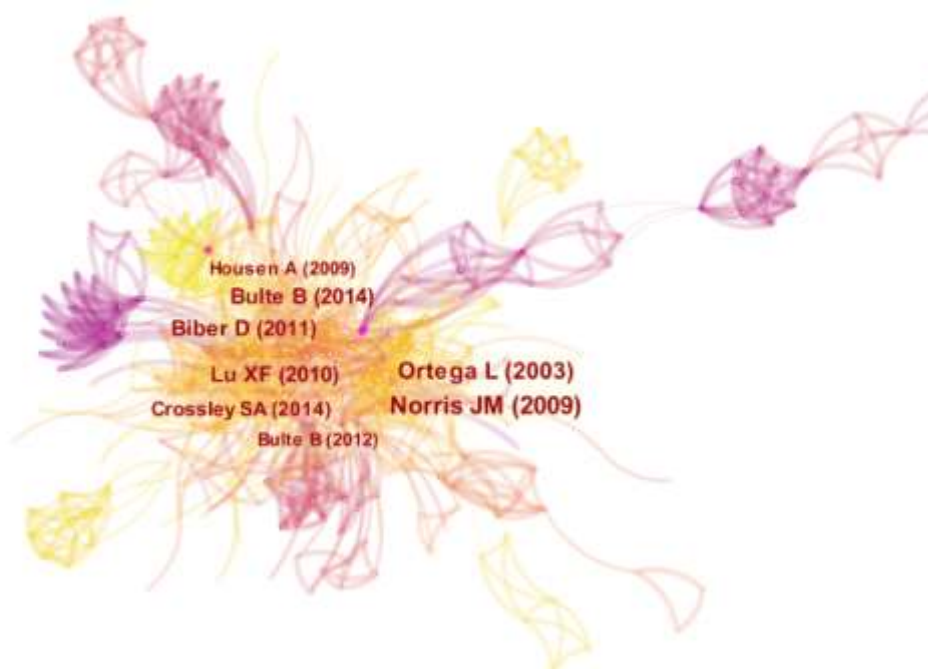


Fig. 4 Knowledge mapping for co-cited reference

Table 2. The five top-cited articles in written syntactic complexity

Cited frequency	Author (year)	Title	Source
92	Norris (2009)	Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity	<i>Applied Linguistic</i>
75	Ortega (2003)	Syntactic Complexity Measures and their Relationship to L2 Prociency: A Research Synthesis of College-level L2 Writing	<i>Applied Linguistics</i>
57	Biber (2011)	Should We Use Characteristics of Conversation to Measure Grammatical Complexity in L2 Writing Development?	<i>TESOL QUARTERLY</i>
57	Lu (2010)	Automatic analysis of syntactic complexity in second language writing	<i>Intemational Journal of Corpus Linguistics</i>
48	Bulte (2014)	Conceptualizing and measuring short-term changes in L2 writing complexity	<i>Journal of Second Language Writing</i>

Aside from the analysis of top-cited articles, the current study performed a burst detection analysis to identify the most influential publications in the field of written syntactic complexity. The top 10 publications with the strongest document co-citation burstness are shown in Table 3. A burst detection analysis enables researchers to explore the research trends of a research field (Chen, 2006; Wang et al., 2019) and reveal the future trends to some extent (Guo, 2017).

For example, the burst group with the end year of 2022 suggests that their citation burst will probably continue in the future, thus reflecting the future trend in the field. The most common future trend is the automatic analysis of syntactic complexity. Lu (2017) examined the rationality of using the automated measurement tool (L2 syntactic complexity Analyzer) to analyze the complexity of syntactic structures in corpus-based writing text. Another notable finding in Table 3 is that McNamara et al. (2012) and Grant & Ginther (2000) generated the strongest citation burst with a burst strength of 3.14 and 3.09, respectively.

Table 3. Top 10 publications with the strongest document co-citation burstness

Author	Year	Burst Strength	Begin	End	Span
Verspoor (2008)	2008	2.74	2014	2017	4
Jeff & Polio (2014)	2014	2.68	2014	2015	2
Skehan (2009)	2009	2.34	2014	2016	3
Verspoor (2012)	2012	2.84	2015	2018	4
Byrnes (2009)	2009	2.7	2015	2016	2
Grant (2000)	2000	3.09	2018	2019	2
Vyatkina et al. (2015)	2015	2.69	2018	2019	2
Robinson (2005)	2005	2.44	2015	2017	3
McNamara et al. (2014)	2014	3.14	2020	2022	3
Lu (2017)	2017	2.71	2020	2022	3

4. Discussion

As illustrated in Figure 2, a total of 8 major clusters with keyword labels were automatically mined by CiteSpace Version 5.7 software, which are academic writing (cluster #1), complexity (cluster#2), syntactic language development (cluster#3), teaching materials (cluster#4), linguistic complexity (cluster#5), automatic essay scoring (cluster#7), syntactic elaboration (cluster#8), and education (cluster#9). In this chapter, 8 clusters are divided into four groups according to the keyword of paper in these clusters.

4.1 Academic writing

The automatically extracted label of cluster #1 was ‘academic writing’. The cluster focused on (i) synchronous and cross-sectional corpus-based study on the measures of syntactic complexity across different factors like genre, register, discipline, L2 proficiency, etc., and (ii) longitudinal study on how syntactic characteristics of L2 learners’ academic writing change over time.

In terms of the first focus, many researchers capitalized on different corpora to identify the use of grammatical complexity features in academic writing (e.g., Ansarifard et al., 2018; Casal et al., 2021; Larsson & Kaatari, 2020; Liu & Li, 2016; Nasser, 2021; Staples et al., 2016; Zeng, 2021). For instance, Staples et al. (2016) used a cross-sectional corpus to analyze phrasal and clausal complexity across the level of study, discipline, and genre. It is found that the extent of the differences across academic levels is mediated by discipline and genre.

On the other hand, longitudinal studies of how syntactic characteristics of L2 learners’ academic writing change over time were implemented in the last decade. Mazgutova and Kormos (2015) compared two argumentative essays at the beginning and at the end of the EAP course to identify the differences in written syntactic complexity. However, the limitation of this study was that argumentative essays could not be boiled down to the category of academic writing, strictly speaking.

It is noted that recent years have witnessed a new trend in academic writing research that combines the analysis of syntactic characteristics and functional analysis of move structure (e.g., Lu, Casal, et al., 2021; Lu et al., 2020; Lu, Yoon, et al., 2021; Saricaoglu et al., 2021).

4.2 Syntactic language development

Publications in cluster#3 concentrate on (i) cross-sectional study on how the measures of syntactic complexity across the change of time, and (ii) longitudinal study on how syntactic characteristics of L2 writing change over time.

On the whole, cross-sectional studies are concerned with the extent to which different syntactic complexity measures reliably index L2 writers’ proficiency (e.g., Kim & Nam, 2019; Kyle & Crossley, 2017). In contrast, those longitudinal studies embark on tracking the development of or changes in syntactic complexity measures over time (e.g., Bulté & Housen, 2014a; Kyle et al., 2021; Mazgutova & Kormos, 2015; Menke & Strawbridge, 2019; Yoon & Polio, 2017). However, the choice of which components of syntactic complexity (global or fine-grained complexity) are related to proficiency and of which measures should be employed is critical to the results of the studies (Kuiken et al., 2019).

4.3 Syntactic complexity in education

The label was extracted from cluster #4 (teaching materials) and cluster#9 (education), which mainly deals with (i) the analysis of readability and syntactic characteristics of various teaching materials, like textbooks and exam materials, and (ii) with how syntactic

complexity measures vary across different factors affecting language learning, like feedback, task design, task complexity, and etc. (e.r., Golparvar & Rashidi, 2021; Lee, 2020; Robinson, 2005; Zhan et al., 2021)

Through the analysis of 3,368 adapted English texts solicited from a corpus of teaching materials, Jin et al. (2020) established syntactic complexity benchmarks to guide text adaptation practices in such contexts, having implications for grading the difficulty level of reading or teaching materials.

While many studies investigated the relationship between written performance and different factors in teaching classrooms, thus facilitating the improvement of teaching quality.

4.4 Syntactic complexity

According to the keywords in these clusters, cluster#2 (complexity), cluster#5 (linguistic complexity), cluster#7 (automatic essay scoring), and cluster#8 (syntactic elaboration) were labelled as syntactic complexity itself, focusing on (i) the construct of syntactic complexity in L2 writing, (ii) on the validity of syntactic complexity measures, and (iii) on the analyzer of syntactic complexity.

As previously noted, syntactic complexity is a multidimensional concept. That is because, on the one hand, syntactic complexity, coupled with lexical complexity, is one of the constructs of linguistic complexity (Bulté & Housen, 2012), while on the other hand, the concept itself involves syntactic diversity and syntactic elaboration, which both have different measures needed to analyzed (Ortega, 2003).

Therefore, choosing which components of syntactic complexity are related to proficiency and which measures should be employed is crucial to the final results. Many authors then examined whether those previous measures can fully reflect the difference in syntactic complexity(e.g., Jagaiah et al., 2020).

In terms of the analysis of these syntactic complexity measures, there are some widely known automated text analysis tools, like the Biber tagger, Coh-Metrix, the Tool for the Automatic Analysis of Syntactic Complexity (TAASSC), and L2 Syntactic Complexity Analyzer (L2SCA), making automatic essay scoring more efficient and trustworthy.

5. Conclusion

In this article, we conducted a scientometric analysis of papers on written syntactic complexity in the past 10 years (2010-2022) to explore current development and challenges faced by related studies. Through visualization analysis, the current research trends of syntactic complexity are summarized as follows:

1. The extensiveness of Research Contents:

At the macro level, studies on written syntactic complexity exhibit an interdisciplinary nature, whose research fields involve Linguistics, Psychology, Education, Computer science, and etc. That is to say; it should be noted that the results of research on written syntactic complexity may vary across distinct research methods and purposes.

2. Depth of Research Contents:

Studies on written syntactic complexity have made great progress in their research depth at the micro-level. For instance, along with the development of computational linguistics, the analysis of written syntactic complexity measures underwent a transition from manual analysis to automatic analysis, making the process of analyzing more efficient and reliable. While in terms of syntactic complexity measures, many researchers embarked on a new exploration on the validity of these measures, namely, whether these previous measures can flexibly reflect the differences in syntactic complexity under different contexts.

Therefore, in light of the current trend of written syntactic complexity research, some notes need to be considered in subsequent studies.

First, with respect to SCMs (syntactic complexity measures), it should be noted that considering the fact that a consistent definition for these SCMs is much needed, and the total number of SCMs is large, it is difficult to compare the results across studies or to identify consistent patterns in the use of SCMs. In addition, whether the SCMs are sensitive enough to gauge language learners' writing quality and language proficiency remains to be seen.

Second, in terms of the factors involved in the complexity of syntactic structure, existing research shows that syntactic complexity is affected by many factors, such as language proficiency, different SCMs, task design, writing genre, writing modality, and formality of task. However, to date, there are still some potential factors affecting the measures of syntactic complexity, thus resulting in inadequate control of experimental variables in most studies. Therefore, it is necessary to further clarify the factors that affect the complexity of syntactic structures and examine these factors' independent and overlapping effects.

To sum, the research on written syntactic complexity has attracted much attention and faces many challenges. We hope that such a study may bring new changes in writing instruction and assessment in the classroom through a scientometric analysis of syntactic complexity.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

- [1] Ansarifard, A., Shahriari, H., & Pishghadam, R. (2018). Phrasal complexity in academic writing: A comparison of abstracts written by graduate students and expert writers in applied linguistics. *Journal of English for Academic Purposes*, 31, 58–71. <https://doi.org/10.1016/j.jeap.2017.12.008>
- [2] Aryadoust, V., & Ang, B. H. (2021). Exploring the frontiers of eye-tracking research in language studies: a novel co-citation scientometric review. *Computer Assisted Language Learning*, 34(7), 898–933. <https://doi.org/10.1080/09588221.2019.1647251>
- [3] Barrot, J., & Gabinete, M. K. (2021). Complexity, accuracy, and fluency in the argumentative writing of ESL and EFL learners. *IRAL - International Review of Applied Linguistics in Language Teaching*, 59(2), 209–232. <https://doi.org/10.1515/iral-2017-0012>
- [4] Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5–35. <https://doi.org/10.5054/tq.2011.244483>
- [5] Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21–46). Amsterdam: John Benjamins.
- [6] Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65. <https://doi.org/10.1016/j.jslw.2014.09.005>
- [7] Casal, J. E., Lu, X., Qiu, X., Wang, Y., & Zhang, G. (2021). Syntactic complexity across academic research article part-genres: A cross-disciplinary perspective. *Journal of English for Academic Purposes*, 52, 100996. <https://doi.org/10.1016/j.jeap.2021.100996>
- [8] Chen, C. (2004). Searching for intellectual turning points: Progressive Knowledge Domain Visualization. *Proceedings of the National Academy of Sciences*, 101(Supplement 1), 5303–5310. <https://doi.org/10.1073/pnas.0307513100>
- [9] Chen, C. (2009). Towards an explanatory and computational theory of scientific discovery. *Journal of Informetrics*, 3(3), 191–209. <https://doi.org/10.1016/j.joi.2009.03.004>
- [10] Chen, C. (2010). The structure and dynamics of co-citation clusters: A multiple-perspective co-citation analysis. *Journal of the American Society for Information Science and Technology*, 61(7), 1386–1409. <https://doi.org/10.1002/asi.21309>
- [11] Chen, C. (2016). *CiteSpace: A practical guide for scientific mapping literature*. Hauppauge, NY: Nova Publishers.
- [12] Frantz, R. S., Starr, L. E., & Bailey, A. L. (2015). Syntactic Complexity as an Aspect of Text Complexity. *Educational Researcher*, 44(7), 387–393. <https://doi.org/10.3102/0013189X15603980>
- [13] Golparvar, S. E., & Rashidi, F. (2021). The effect of task complexity on integrated writing performance: The case of multiple-text source-based writing. *System*, 99, 102524. <https://doi.org/10.1016/j.system.2021.102524>
- [14] Grant, L., & Ginther, A. (2000). Using Computer-Tagged Linguistic Features to Describe L2 Writing Differences. *Journal of Second Language Writing*, 9(2), 123–145. [https://doi.org/10.1016/S1060-3743\(00\)00019-9](https://doi.org/10.1016/S1060-3743(00)00019-9)
- [15] Guo, H. (2017). *Measurement of Emerging Research Field Identification: Theory, Indicators and Examples*. Beijing: Science Press.
- [16] Jackson, D. O., & Suethanapornkul, S. (2013). The Cognition Hypothesis: A Synthesis and Meta-Analysis of Research on Second Language Task Complexity. *Language Learning*, 63(2), 330–367. <https://doi.org/10.1111/lang.12008>
- [17] Jagaiah, T., Olinghouse, N. G., & Kearns, D. M. (2020). Syntactic complexity measures: variation by genre, grade level, students' writing abilities, and writing quality. *Reading and Writing*, 33(10), 2577–2638. <https://doi.org/10.1007/s11145-020-10057-x>
- [18] Kim, J., & Nam, H. (2019). How do textual features of L2 argumentative essays differ across proficiency levels? A multidimensional cross-sectional study. *Reading and Writing*, 32(9), 2251–2279. <https://doi.org/10.1007/s11145-019-09947-6>
- [19] Kuiken, F., Vedder, I., Housen, A., & de Clercq, B. (2019). Variation in syntactic complexity: Introduction. *International Journal of Applied Linguistics (United Kingdom)*, 29(2), 161–170. <https://doi.org/10.1111/ijal.12255>
- [20] Kyle, K., & Crossley, S. (2017). Assessing syntactic sophistication in L2 writing: A usage-based approach. *Language Testing*, 34(4), 513–535. <https://doi.org/10.1177/0265532217712554>
- [21] Kyle, K., Crossley, S., & Verspoor, M. (2021). Measuring longitudinal writing development using indices of syntactic complexity and sophistication. *Studies in Second Language Acquisition*, 43(4), 781–812. <https://doi.org/10.1017/S0272263120000546>
- [22] Larsson, T., & Kaatari, H. (2020). Syntactic complexity across registers: Investigating (in)formality in second-language writing. *Journal of English for Academic Purposes*, 45, 100850. <https://doi.org/10.1016/j.jeap.2020.100850>
- [23] Lee, J. (2020). Task closure and task complexity effects on L2 written performance. *Journal of Second Language Writing*, 50(October), 100777. <https://doi.org/10.1016/j.jslw.2020.100777>
- [24] Li, X., Ma, E., & Qu, H. (2017). Knowledge mapping of hospitality research a visual analysis using CiteSpace. *International Journal of Hospitality Management*, 60, 77–93.
- [25] Lim, M. H., & Aryadoust, V. (2021). A scientometric review of research trends in computer-assisted language learning (1977–2020). *Computer Assisted Language Learning*, 0(0), 1–26. <https://doi.org/10.1080/09588221.2021.1892768>
- [26] Liu, L., & Li, L. (2016). Noun phrase complexity in EFL academic writing: A corpus-based study of postgraduate academic writing. *Journal of Asia TEFL*, 13(1), 48–65. <https://doi.org/10.18823/asiatefl.2016.13.1.4.48>
- [27] Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474–496. <https://doi.org/10.1075/ijcl.15.4.02lu>
- [28] Lu, X. (2017). Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment.

- Language Testing*, 34(4), 493–511. <https://doi.org/10.1177/0265532217710675>
- [29] Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing*, 29, 16–27. <https://doi.org/10.1016/j.jslw.2015.06.003>
- [30] Lu, X., Casal, J. E., & Liu, Y. (2020). The rhetorical functions of syntactically complex sentences in social science research article introductions. *Journal of English for Academic Purposes*, 44, 100832. <https://doi.org/10.1016/j.jeap.2019.100832>
- [31] Lu, X., Casal, J. E., Liu, Y., Kisselev, O., & Yoon, J. (2021). The relationship between syntactic complexity and rhetorical move-steps in research article introductions: Variation among four social science and engineering disciplines. *Journal of English for Academic Purposes*, 52, 101006. <https://doi.org/10.1016/j.jeap.2021.101006>
- [32] Lu, X., Yoon, J., Kisselev, O., Casal, J. E., Liu, Y., Deng, J., & Nie, R. (2021). Rhetorical and phraseological features of research article introductions: Variation among five social science disciplines. *System*, 100, 102543. <https://doi.org/10.1016/j.system.2021.102543>
- [33] Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, 29, 3–15. <https://doi.org/10.1016/j.jslw.2015.06.004>
- [34] McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2012). Automated evaluation of text and discourse with Coh-Metrix. In Automated Evaluation of Text and Discourse with Coh-Metrix. Cambridge University Press. <https://doi.org/10.1017/CBO9780511894664>
- [35] Menke, M. R., & Strawbridge, T. (2019). The writing of Spanish majors: A longitudinal analysis of syntactic complexity. *Journal of Second Language Writing*, 46, 100665. <https://doi.org/10.1016/j.jslw.2019.100665>
- [36] Mohamadi Zenouzagh, Z. (2020). Syntactic complexity in individual, collaborative and E-collaborative EFL writing: mediating role of writing modality, L1 and sustained development in focus. *Educational Technology Research and Development*, 68(6), 2939–2970. <https://doi.org/10.1007/s11423-020-09818-w>
- [37] Nasser, M. (2021). Is postgraduate English academic writing more clausal or phrasal? Syntactic complexification at the crossroads of genre, proficiency, and statistical modelling. *Journal of English for Academic Purposes*, 49, 100940. <https://doi.org/10.1016/j.jeap.2020.100940>
- [38] Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–578. <https://doi.org/10.1093/applin/amp044>
- [39] Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24, 492–518. <https://doi.org/10.1093/applin/24.4.492>
- [40] Qin, W., & Uccelli, P. (2020). Beyond linguistic complexity: Assessing register flexibility in EFL writing across contexts. *Assessing Writing*, 45. <https://doi.org/10.1016/j.asw.2020.100465>
- [41] Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. In *IRAL - International Review of Applied Linguistics in Language Teachings*, 43, 1–32. <https://doi.org/10.1515/iral.2005.43.1.1>
- [42] Saricaoglu, A., Bilki, Z., & Plakans, L. (2021). Syntactic complexity in learner-generated research paper introductions: Rhetorical functions and move/step realization level. *Journal of English for Academic Purposes*, 53, 101037. <https://doi.org/10.1016/j.jeap.2021.101037>
- [43] Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic Writing Development at the University Level. *Written Communication*, 33(2), 149–183. <https://doi.org/10.1177/0741088316631527>
- [44] Wang, H., Yan, X., & Guo, H. (2019). Visualizing the knowledge domain of embodied language cognition: A bibliometric review. *Digital Scholarship in the Humanities*, 34(1), 21–31. <https://doi.org/10.1093/llc/fqy010>
- [45] Yoon, H.-J., & Polio, C. (2017). The Linguistic Development of Students of English as a Second Language in Two Written Genres. *TESOL Quarterly*, 51(2), 275–301. <https://doi.org/10.1002/tesq.296>
- [46] Zeng, X. (2021). Comparing Syntactic Complexity: Research Papers Written by Undergraduate Researchers and Published Research Papers. *Journal of Education*, 002205742110164. <https://doi.org/10.1177/00220574211016435>
- [47] Zhan, J., Sun, Q., & Zhang, L. J. (2021). Effects of manipulating writing task complexity on learners' performance in completing vocabulary and syntactic tasks. *Language Teaching Research*. <https://doi.org/10.1177/13621688211024360>