A Comparative Analysis of Syntactic Complexity in Applied Linguistics Abstracts Written by Chinese Novice Writers and Native English Advanced Writers

Mingming Zhao, Tianshuang Ge*

Department of Foreign Languages, Dalian University of Technology, Dalian, China
Email: *tianshuangge@dlut.edu.cn

Abstract

The rhetorical structure of abstracts has been a widely discussed topic, as it can greatly enhance the abstract writing skills of second-language writers. This study aims to provide guidance on the syntactic features that L2 learners can employ, as well as suggest which features they should focus on in English academic writing. To achieve this, all samples were analyzed for rhetorical moves using Hyland's five-rhetorical move model. Additionally, all sentences were evaluated for syntactic complexity, considering measures such as global, clausal and phrasal complexity. The findings reveal that expert writers exhibit a more balanced use of syntactic complexity across moves, effectively fulfilling the rhetorical objectives of abstracts. On the other hand, MA students tend to rely excessively on embedded structures and dependent clauses in an attempt to increase complexity. The implications of these findings for academic writing research, pedagogy, and assessment are thoroughly discussed.

Keywords

Syntactic Complexity, Rhetorical Functions, Research Article Abstract, L2 Academic Writing

1. Introduction

The English abstract plays a crucial role for L2 learners, especially those pursuing master's and doctoral degrees. As their writing skills improve, the complexity of their tasks also increases. To overcome this challenge, L2 learners must become proficient in formal genres to successfully navigate their academic fields [1]. In recent years, there has been a growing emphasis on aligning linguistic

elements with rhetorical functions in academic writing. Understanding the rhetorical structure of abstracts is essential for novice and second-language writers to enhance their abstract writing skills. This understanding enables them to effectively incorporate and summarize information using advanced linguistic and rhetorical techniques. Aspiring academic writers can benefit from studying journals they wish to publish in, identifying differences, and emulating the writing of exceptional scholars.

However, despite the increasing interest in the relationship between syntactic complexity and function in L2 writing research, there is a lack of research on the rhetorical roles of complex sentence structures in abstract writing. Recent studies by Casal *et al.* [2] and Lu *et al.* [3] have addressed this gap and demonstrated a strong connection between syntactic complexity and the rhetorical purposes of research articles. These studies indicate that the use of complex sentence structures may vary depending on the rhetorical goals of the abstract.

In response to this gap, Lu *et al.* [4] [5] propose a framework that combines corpus- and genre-based approaches to academic writing research and pedagogy. This framework can assist novice writers in understanding the rhetorical functions of academic discourse and making appropriate language choices. By analyzing authentic abstracts and identifying the syntactic features associated with different rhetorical moves, L2 learners can develop a better understanding of how to effectively use complex sentence structures in their own abstract writing. This research has important implications for improving L2 learners' abstract writing skills and enhancing their overall academic writing proficiency.

This study aims to explore the relationship between syntactic complexity and rhetorical moves in abstract writing among Chinese novice writers and international advanced writers. By analyzing abstracts from international journal papers, the study seeks to provide a comprehensive understanding of how syntactic complexity is employed across different rhetorical moves. This research will offer Chinese graduates a global perspective on abstract writing and demonstrate how complex sentence structures can be used to achieve rhetorical goals. The study will specifically investigate how fourteen measures of syntactic complexity are related to the rhetorical moves outlined by Hyland [6] in the abstract sections of published research articles and Chinese master's theses. This research has the potential to enhance the abstract writing skills of Chinese graduates and contribute to their overall academic writing proficiency.

1.1. Syntactic Complexity in L2 Writing Quality

Syntactic complexity refers to the diversity, sophistication, and elaboration of syntactic structures used in language production [7] [8] [9]. Studies on L2 writing have found a positive relationship between higher language proficiency and greater syntactic complexity [7] [10]. However, it is important to consider that increased complexity does not always indicate higher proficiency or better language performance, as argued by Ortega [9]. Factors such as register, genre, and

task can influence syntactic complexity [8] [11]. For example, Pallotti [11] discovered that nonnative speakers tend to use longer and more complex syntactic structures in phone conversations, although in an unnatural or inappropriate manner compared to native speakers. Therefore, specific tasks and genres, particularly within the spoken genre, may require a lower level of syntactic complexity at a higher proficiency level.

Numerous studies in the field of second language (L2) research have explored the relationship between syntactic complexity and writing quality. These studies can be categorized as either cross-sectional or longitudinal. Longitudinal studies specifically investigate changes in syntactic complexity in L2 writing over a designated period of time [12] [13] [14]. For instance, Casanave [13] conducted an analysis of written works by Japanese English learners at an intermediate level and observed an increase in syntactic complexity, characterized by longer and more intricate clauses, over the course of three semesters.

Studies in the field of second language (L2) research have examined the correlation between syntactic complexity and proficiency using a cross-sectional design. While studies on native authors focus on whole sentence production, research on L2 authors has shifted towards analyzing phrase complexity and syntactic embedding. For instance, Larsen-Freeman [15] used T-unit indices to differentiate essays written by L2 learners at different proficiency levels, finding that the percentage and average length of error-free T-units were the best indicators of proficiency. Additionally, Ferris [16] discovered that high proficiency L2 writers exhibited greater usage of passives, nominalizations, conjuncts, and prepositions compared to low proficiency writers, indicating a positive correlation between writing quality and syntactic complexity.

Furthermore, research in applied linguistics has explored syntactic complexity across different disciplines. Studies by Lu *et al.* [4] [5] and Zhang and Lu [17] have shown that syntactic complexity significantly predicts L2 writing development. For example, Lei [18] investigated syntactic differences between Ph.D. and MA applied linguistics students and found that intermediate English writers used fewer dimensions of syntactic complexity compared to advanced learners, as evidenced by measures such as the average length of T-units, number of clauses, subordinate clauses, and compound noun structures in T-units.

In conclusion, research in L2 studies has extensively investigated the relationship between syntactic complexity and the quality of writing, taking into account proficiency levels, first language backgrounds, and disciplinary differences. These studies highlight the significance of syntactic complexity as an indicator of writing proficiency and provide valuable insights into the improvement of L2 writing skills.

However, the exact relationship between syntactic complexity and proficiency levels is still unclear, as not all studies show a positive correlation. This is because researchers use different measures to assess sentence complexity. Therefore, it is important to consider various factors such as the type of writing, cha-

racteristics of the learners, task requirements, and the overall context when analyzing syntactic complexity in L2 writing. Based on the previous findings on proficiency-related differences in different aspects of L2 writing, this study aims to examine potential variations in sentence complexity across different rhetorical functions in L2 writing, taking into account proficiency levels.

1.2. Multidimensional Measurements in Syntactic Complexity

In recent years, researchers have focused on finding reliable and valid measures to quantify the complexity of sentence structures in L2 writing. Traditionally, length-based measures such as the mean length of T-unit, mean length of sentence, or mean length of clause have been used to assess syntactic complexity [19]. However, relying solely on one dimension of syntactic complexity has faced criticism, and researchers have recognized the importance of exploring syntactic complexity using multiple dimensions [19] [20].

Wolfe-Quintero *et al.* [21] summarized measures of accuracy, fluency, and complexity employed in 39 studies on the development of L2 writing, recommending additional measures for further research. Norris and Ortega [19] proposed a multidimensional operationalization of syntactic complexity, incorporating sub-constructs at the global, clausal, and phrasal levels. Since then, numerous measures have been proposed to assess each of these dimensions, integrated into commonly used computational tools for analyzing syntactic complexity.

For example, the L2 Syntactic Complexity Analyzer [22] provides 14 holistic measures across five categories, including the length of production units (e.g., mean length of sentence), overall sentence complexity (e.g., clauses per sentence), coordination (e.g., T-units per sentence), subordination (e.g., dependent clauses per clause), and phrasal complexity (e.g., complex nominals per clause). According to Yang and Lu [10], mean length of sentence and mean length of T-unit are seen as global complexity measures, while the other six measures are seen as local-level complexity measures.

In conclusion, researchers have moved beyond relying solely on length-based measures and have embraced a multidimensional approach to assess syntactic complexity in L2 writing. By considering multiple dimensions at the global, clausal, and phrasal levels, researchers can gain a more comprehensive understanding of syntactic complexity and its relationship to L2 writing development.

While Wolfe-Quintero *et al.* [21] focused on clause and verb-related structures, Biber *et al.* [23] emphasized the role of noun phrases and noun modifiers. They questioned the effectiveness of traditional measures like T-unit and subordination-based indices and suggested the need for more detailed indices to distinguish proficiency levels. He found that everyday conversations mainly involve finite clauses as adverbials and verb complements, while academic writing predominantly uses phrasal structures. This aligns with Lu *et al.*'s [3] study on L2 Chinese writing, which found that fine-grained indices of phrasal complexity are better predictors of writing quality in narrative and argumentative essays com-

pared to larger-grained indices.

Moreover, researchers, including Biber *et al.* [20], advocate for incorporating measures of phrasal complexity into a comprehensive understanding of syntactic complexity. Based on this widely accepted multidimensional concept, this study highlights the significance of phrasal complexity as a key indicator in writing.

Based on Lu [8] and Yang *et al.* [10], Table 1 presents our conceptualization of the multi-dimensional construct of syntactic complexity. This study aims to investigate fourteen syntactic complexity indices, which encompass two global indices, seven clausal indices, and five phrasal indices. The inclusion of this comprehensive set of measures was motivated by its representation of the various measures commonly employed in L2 writing research. While some measures may have stronger predictive power for L2 proficiency or exhibit partial redundancy, the decision to include a larger set of measures was made to outweigh concerns regarding redundancy. This approach enables the examination of previous claims and the identification of measures that genuinely reflect L2 proficiency or writing quality, while encompassing all significant dimensions of syntactic complexity.

1.3. Connecting Rhetorical and Formal Linguistic Features of Academic Writing

ESP scholars have conducted research on the rhetorical structures of academic genres, using function-first discourse analysis frameworks to identify and analyze the rhetorical moves and steps of a genre. Corpus methodologies have also been employed to examine linguistic features associated with genre practices. These corpus-based studies focus on identifying linguistic forms automatically and interpreting them in a decontextualized manner.

EAP writing research emphasizes the complex nature of discursive practices, considering writers' intentions, choices, and the expectations of discourse community members. Various methodologies are used, but analyzing rhetorical moves and steps and linguistic patterns in corpora is particularly common.

Quantitative analysis of the relationship between linguistic constructs and the quality of academic writing has been a focus of corpus-based EAP writing research. Computational tools, such as those measuring sentential, clausal, and phrasal complexity, have been used to automate syntactic complexity analysis. Recent advancements in NLP have enhanced researchers' ability to identify linguistic features in large writing corpora, enabling a deeper understanding of writing skill development and measures of writer competence. Table 2 provides an overview of NLP tools used for computing complexity features.

Although corpus-based research has provided valuable insights into the quantitative relationship between linguistic features and writing quality, a limitation is the tendency to focus solely on linguistic features without considering their rhetorical functions. Writing quality is not solely determined by the presence and frequency of linguistic features, but also by their effective use in achieving rhetorical functions.

Table 1. A multidimensional construct of syntactic complexity metrics.

Measures	Code	Definition
Global level		
Mean length of sentence	MLS	# of words/# of sentences
Mean length of T-unit	MLT	# of words/# of T-units
Clausal level		
Mean length of clause	MLC	# of words/# of clauses
Dependent clauses per T-unit	DC/T	# of dependent clauses/# of T-units
T-units per sentence	T/S	# of T-units/# of sentences
Dependent clauses per clause	DC/C	# of dependent clauses/# of clauses
Clauses per sentence	C/S	# of clauses/# of sentences
Clauses per T-unit	C/T	# of clauses/# of T-unit
Complex T-units per T-unit	CT/T	# of complex T-units/# of T-units
Phrasal level		
Coordinate phrases per clause	CP/C	# of coordinate phrases/# of clauses
Complex nominals per clause	CN/C	# of complex nominals/# of clauses
Coordinate phrases per T-unit	CP/T	# of coordinate phrases/# of T-units
Verb phrases per T-unit	VP/T	# of verb phrases/# of T-units
Complex nominals per T-unit	CN/T	# of complex nominals/# of T-units

Table 2. NLP tools for measurements of syntactic complexity.

Tools	Metrics	Accuracy	Requirements	Reference	
The Biber Tagger	Grammatical complexity metrics	89%	Not public available	Biber *et al.*	1999
Coh-Metrix	108 metrics of Cohesion, syntactic complexity, lexical diversity, word information	90%	Online interface (Chrome or Firefox browser)	McNamara *et al.*	2014
L2 Syntactic Complexity Analyzer (L2SCA)	14 metrics of syntactic complexity	F-score of 0.83 for complex nominal	Online interface; MAC, LINUX or UNIX systems with Java	Lu	2010
Tools for Automatic Analysis of Syntactic Sophisation and Complexity (TAASSC)	178 indies of syntactic complexity	91%	Windows, MAC or Python	Kyle	2016

To address this limitation, integrated corpus and genre analytic investigations have emerged to examine the rhetorical and linguistic dimensions of text together. For example, Cortes [24] analyzed a corpus of research article introductions and identified lexical bundles aligned with different rhetorical moves and steps. Similarly, Omidian *et al.* [25] classified multi-word expressions in research article abstracts based on their communicative functions. However, both studies assigned rhetorical move tags to chunks without considering the larger context, which could be problematic.

In contrast, Durrant and Mathews-Aydınlı [26] manually annotated a corpus of graduate student writing to identify rhetorical moves and steps, then associated formulaic forms with each rhetorical function. Le and Harrington [27] also identified word clusters in the Discussion sections of applied linguistics research articles and analyzed their discourse functions within a larger context.

While these studies have bridged the "function-form gap" in genre research, they have primarily focused on lexical and phraseological features. Further research should investigate the rhetorical functions of complex syntactic structures, as they directly impact writing quality and are particularly relevant for academic specialists. By considering the rhetorical functions of both linguistic and syntactic features, a more comprehensive understanding of genre knowledge and development can be achieved.

In recent years, there has been a growing recognition of the importance of integrating formal and rhetorical knowledge in proficient academic writing. This emphasis on aligning linguistic elements with rhetorical functions has led to a number of studies exploring the relationship between syntactic complexity and rhetorical functions in research article introductions. These studies have highlighted the significance of syntactic complexity in differentiating between different rhetorical functions.

Furthermore, these studies have also identified differences in syntactic complexity between emerging and expert writers, suggesting that syntactic complexity plays a crucial role in distinguishing writing proficiency levels. Additionally, specific rhetorical functions have been found to be associated with complex sentence patterns, as evidenced by the works of Lu *et al.* [3], Yin *et al.* [28], and Jiang and Kang [29].

Building upon these previous studies, the present research aims to examine complex syntactic structures in sentences that serve various rhetorical functions. However, it is important to note that the corpus for this study consists of abstracts written by Chinese master's students and international experts. By investigating the relationship between syntactic complexity and the level of move realization, the author seeks to further enhance our understanding of how syntactic complexity contributes to effective academic writing.

2. Method

2.1. Corpus Design

A comparative analysis was conducted to examine the disparities between abstracts authored by advanced academic writers and graduate L2 writers in the field of applied linguistics. A corpus of 200 pertinent abstracts from the period spanning 2017 to 2020 was compiled from two distinct sources. The first source encompassed abstracts composed by Chinese graduate writers and was obtained from the China National Knowledge Infrastructure (CNKI) website. The second source comprised abstracts written by advanced academic writers and published in esteemed international journals in applied linguistics, such as English for Spe-

cific Purposes, Journal of English for Academic Purposes, Journal of Second Language Writing, and System.

To ensure the reliability and relevance of the collected data, each abstract was meticulously examined and verified for its close alignment with the domain of applied linguistics. The texts were stored in plain TXT files during the data collection phase. Moreover, an Excel file was created to store the original data, which included details such as author names, journal titles, word counts, and collection dates. This systematic organization of data aimed to facilitate the efficient retrieval of the original files and associated information throughout the study.

Although the texts in different sections varied substantially in length, this variation was not expected to affect the comparisons made in our analysis, for the reason that the complexity structure indices were all calculated as mean length of production units or ratios of the frequency of one type of complexity structure to that of another. Based on the framework, two corpora are built as shown in **Table 3**. The first corpus consisted of master's theses written by Chinese graduates majoring in applied linguistics. The average length of each essay in this corpus was 181.45 tokens. The second corpus was composed of research articles from four international journals, with an average length of 489.28 tokens.

2.2. Rhetorical Move Annotation

To investigate potential variations in the syntactic complexity of research article abstracts, the author conducted a move analysis of each sample using Hyland's five-rhetorical move model (see **Table 4**). According to Hyland [30], each move in his five-rhetorical model serves a specific communicative purpose.

During annotation, both the author and supervisor independently coded samples in both corpora using rhetorical chunks as the unit of analysis. Sentences were then annotated to examine the relationship between rhetorical functions and syntactic complexity. In cases where multiple moves occurred within a single sentence, the sentence was annotated with a list of those moves, with the

Table 3. Descriptive statistics of the two corpora.

Years	IWA			CWA		
	Data Sources	Text N.	Words N.	Data Sources	Text N.	Words N.
2017-2021	English for Specific Purposes	25	4660	Hunan University	25	12,935
	Journal of English for Academic Purposes	25	4010	Dalian University of Technology	25	10,886
	Journal of second language writing	25	5145	Harbin Institute of Technology	25	12,283
	System	25	4330	East China Normal University	25	12,824
	Total	100	18,145	Total	100	48,928

Table 4. Hyland's (2000) five-rhetorical move model for the rhetorical structure of RA abstracts.

Labels	Moves	Function
M1	Introduction	Establishes context of the paper and motivates the research or discussion.
M2	Purpose	Indicates purpose, thesis or hypothesis, outlines the intention of the paper.
M3	Methods	Provides information on design, procedures, assumptions, approach, data, etc.
M4	Results	States main findings or results, the argument, or what was accomplished.
M5	Conclusion	Interprets or extends results beyond scope of paper, draws inferences, points to applications or wider implications

move realized by the main clause considered as the sentence's main function. This approach allowed for efficient computation of syntactic complexity indices, as sentences are fundamental units for such analysis. Additionally, it enabled the examination of how the use of complex structures in a sentence by research article writers may be influenced by the presence of multiple rhetorical moves. Inter-rater reliability, assessed with Krippendorff's alpha, yielded a score of 0.9, with disagreements resolved through discussion.

A coding scheme was refined during the pilot annotation phase, considering frequency and numbers. The data collection and preparation processes involved labor-intensive manual coding. After move coding, samples were evaluated for syntactic complexity using L2SCA variables in the TAASSC tool, generating numeric values for 14 holistic measures. Empty TXT files were removed based on the frequency table.

To guide the corpus annotation process, a coding scheme was refined during the pilot annotation phase, as illustrated in Table 5. Both frequency and numbers were considered during the analysis. The data collection and preparation processes required significant manual coding, which was labor-intensive and time-consuming. After manually annotating the move coding, each sample in the dataset was assessed for syntactic complexity using L2SCA variables in the TAASSC tool developed by Kyle [31]. This tool provided numeric values for 14 holistic measures of syntactic complexity. Empty TXT files were then removed based on the frequency table.

2.3. Syntactic Complexity Measurements

To assess the syntactic complexity of each sample, fourteen different measures representing interconnected sub-constructs were used. These measures were outlined in the Introduction and are shown in Table 1.

At the global level, complexity was measured by overall sentence complexity (MLS) and overall T-unit complexity (MLT). The increase in T-unit length reflects

Table 5. Descriptive statistics of each corpora.

	Frequency		N. of Sentences	N. of Words	N. of Sentences	N. of Words
	CWA	IWA	IWA	IWA	CWA	CWA
M1	97/100	58/100	107	2755	447	10,188
M2	88/100	86/100	89	2499	214	5641
M3	99/100	99/100	185	4685	410	9944
M4	99/100	100/100	220	5913	720	18,523
M5	73/100	84/100	100	2293	178	4632

an increase in sentence length. At the clausal level, the elaboration of clause was measured by calculating the average number of words per clause (MLC). The degree of clausal subordination was assessed by calculating the average number of dependent clauses per T-unit (DC/T), dependent clauses per clause (DC/C), clauses per T-unit (C/T), clauses per sentence (C/S), and complex clauses per T-unit (CT/T). Clausal coordination was assessed by calculating the number of T-units per sentence (T/S). Both finite clauses and non-finite elements were considered in this analysis. Measures of phrasal coordination (CP/C, CP/T) and noun phrase-noun phrase complexity (VP/T, CN/T, CN/C) were included to represent phrasal syntactic complexity.

The L2 syntactic complexity analyzer (L2SCA) developed by Lu [8] was used to analyze the samples. This analyzer takes a written English text as input and generates frequency counts of nine linguistic units in the text, including words, sentences, clauses, dependent clauses, T-units, complex T-units, coordinate phrases, complex nominals, and verb phrases. It then generates 14 indices of syntactic complexity for the text, following Lu's [8] [22] definitions. In summary, the measures used in this analysis encompass five dimensions of syntactic complexity: length of production unit, amount of subordination, amount of coordination, degree of phrasal sophistication, and overall sentence complexity. These measures capture the three levels of the construct.

2.4. Research Questions

The present study aims to address the gaps identified in the previous sections and explore the complexity features associated with different rhetorical functions in native advanced writers and Chinese novice writers of applied linguistics. To achieve this, the study seeks to answer three research questions:

1) How does the syntactic complexity vary across different rhetorical functions in applied linguistics abstracts written by native English advanced writers?

2) How does the syntactic complexity vary across different rhetorical functions in applied linguistics abstracts written by Chinese novice writers?

3) What are the significant differences in syntactic complexity between applied linguistics abstracts written by Chinese novice writers ad native English

advanced writers when considering different rhetorical functions?

The first and second research questions aim to analyze the language used to achieve rhetorical functions in abstracts by examining fourteen measures of syntactic complexity in Chinese master's theses and native English advanced writer's papers in applied linguistics. This analysis expands the research on the "function-form gap" beyond lexical and phraseological perspectives to include syntactic domains.

The third research question investigates the differences in syntactic complexity in abstract writing written by Chinese novice writers and native English advanced writers. Additionally, the study examines the rhetorical functions of complex sentences in each measure and the distribution of text dedicated to each rhetorical move in the two corpora. By identifying these differences, the study provides insights into the construction of complex sentences for rhetorical purposes and sheds light on the potential benefits of such structures for novice writers in applied linguistics research articles. This information can be utilized in teaching rhetorical and formal conventions of academic writing and can serve as a resource of complex sentences with rhetorical annotations for pedagogical activities.

2.5. Analytical Procedures

Once the corpora have been prepared, the first step is to address the first two research questions by automatically identifying 14 complexity metrics using TAASSC. Next, the normality of each metric in the two corpora is examined using the Kolmogorov-Smirnov Test. Subsequently, the Kruskal-Wallis one-way ANOVA test and the ANOM test are conducted to investigate any significant differences and the probability of occurrence of syntactic complexity across moves in the abstracts of the two corpora.

In relation to the third research question, the syntactic complexity that significantly differs between the two groups for each move is identified and displayed using the Mann-Whitney U test. Finally, based on the results and findings obtained from these analyses, the aims of our study can be determined.

3. Results and Discussion

3.1. Syntactic Complexity in Rhetorical Moves of IWA Abstracts

Regarding the first research question, as shown in Table 6, significant differences were found among the five rhetorical moves in terms of the fourteen syntactic complexity indices, except for MLS and MLT, as indicated by the results of Kruskal-Wallis one-way analysis of variance (ANOVA) conducted by SPSS (V. 27). The descriptive statistics of syntactic complexity in various rhetorical moves in IWA are also presented in Tables 7-9.

To further investigate the syntactic complexity differences in sentences representing different moves, this study also employed graphical analysis of means (ANOM) to identify and visually display moves that significantly differed from

Table 6. Significant difference among five-rhetorical moves in 14 syntactic complexity indices in IWA.

Syntactic complexity measures		IWA Asym. Sig. (2-sided test)
Global level		
	MLS	0.210
	MLT	0.093
Clausal level		
	MLC	**0.001***
	DC/T	**0.000***
	DC/C	**0.001***
	T/S	**0.001***
	C/T	**0.000***
	CT/T	**0.001***
	C/S	**0.001***
Phrasal level		
	CP/C	**0.000***
	CN/C	**0.001***
	CN/T	**0.014***
	VP/T	**0.000***
	CP/T	**0.001***

Note: * indicates that the measure has significant differences in five rhetorical moves.

Table 7. Descriptive statistics of global complexity between different moves in IWA.

Moves	M1		M2		M3		M4		M5	
Text N.	59		85		99		100		85	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MLS	27.24	7.33	28.57	10.74	26.80*	9.67	29.64**	10.95	26.93	13.15
MLT	25.93	8.24	26.99	11.78	25.17	10.70	27.23**	10.75	23.38*	9.32

Note: * indicates that the measure has the fewest numbers or shortest length in that move; ** indicates that the measure has the most numbers or longest length in that move.

Table 8. Descriptive statistics of clausal complexity between different moves in IWA.

Moves	M1		M2		M3		M4		M5	
Text N.	59		85		99		100		85	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MLC	17.93	7.975	21.27	10.16	21.66	10.59	17.25*	6.43	23.06**	9.38
T/S	1.08	0.22	0.99	0.24	1.03	0.27	1.13**	0.30	0.54*	0.76
C/S	1.69	0.60	1.41	0.69	1.23*	0.61	1.87	0.76	16.76**	8.22
DC/C	0.29	0.24	0.20	0.28	0.16*	0.25	0.33	0.22	1.54**	0.76
DC/T	0.59	0.59	0.39	0.63	0.27	0.50	0.68**	0.60	0.23*	0.27
C/T	1.58	0.59	1.36	0.72	1.16*	0.61	1.69	0.66	2.34**	1.25
CT/T	0.45	0.40	0.31	0.46	0.20*	0.34	0.50	0.37	1.02**	0.16

Note: * indicates that the measure has the fewest numbers or shortest length in that move; ** indicates that the measure has the most numbers or longest length in that move.

Table 9. Descriptive statistics of phrasal complexity between different moves in IWA.

Moves	M1		M2		M3		M4		M5	
Text N.	59		85		99		100		85	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
CP/C	0.49*	0.56	0.79	0.89	0.92**	0.79	0.70	0.65	0.69	0.73
CP/T	0.70	0.66	0.95	1.11	1.05**	0.80	1.04	0.84	0.39*	0.47
CN/C	3.10	1.41	3.56	1.84	3.26	1.78	2.95*	1.37	3.97**	1.82
CN/T	4.52	1.86	4.54	2.15	3.75	1.78	4.66**	2.10	0.54*	0.62
VP/T	2.30	0.93	2.31	1.44	1.89	0.99	2.42**	1.14	1.58*	0.83

Note: * indicates that the measure has the fewest numbers or shortest length in that move; ** indicates that the measure has the most numbers or longest length in that move.

the group for each measure. The results of ANOM were consistent with the findings of ANOVA (see Tables 7-12). The results of the mean analysis show that among the empirical abstracts in international journals of applied linguistics, under the condition of Asymptotic Sig. (2-sided test) < 0.05: In the global level, as shown in Table 10, for MLS and MLT, the results of ANOM did not indicate any significant differences across the different moves.

As for the clausal dimension, the non-parametric Kruskal Wallis tests are conducted to determine whether there are significant differences between clausal complexity measures (MLC, DC/T, T/S, DC/C, C/S, C/T, and CT/T) respectively and rhetorical moves. Table 8 indicates that sentences realizing five rhetorical moves differed significantly in all clausal complexity measures. Regarding the clausal elaboration measured by MLC, sentences relaying the function M5 (Mean = 23.06) are significantly longer ($p < 0.05$) compared to the overall mean, while sentences relaying the function M4 (Mean = 17.25) are significantly shorter than the overall mean.

In terms of clausal coordination, sentences realizing M4 exhibit a significantly higher ratio of T/S (Mean = 1.13, SD = 0.30) compared to the overall mean, whereas sentences representing M5 (Mean = 0.54) demonstrate a significantly lower ratio of T/S than the overall mean. Moving on to the level of subordination at the clausal level, it is observed that sentences realizing M5 function contain a significantly higher number of clauses per sentence (C/S) (Mean = 16.76) than the average. This means that M5 sentences tend to have more clauses within each sentence.

In contrast, sentences realizing M1, M2, M3, and M4 contain significantly fewer C/S compared to the average. Specifically, sentences relaying the function of M3 have the fewest C/S (Mean = 1.23) among them.

Furthermore, sentences representing M5 also display a significantly higher number of clauses per T-unit (C/T) (Mean = 2.34, SD = 1.25), complex T-units per T-unit (CT/T) (Mean = 1.02, SD = 0.16), and dependent clauses per clause (DC/C) (Mean = 1.54, SD = 0.76) than the overall mean.

Conversely, sentences realizing M3 exhibit significantly lower values for C/T (Mean = 1.16), CT/T (Mean = 0.2), and DC/C (Mean = 0.16) compared to the

Table 10. Rhetorical moves that differ significantly in the global complexity.

Measure	Significantly high	Significantly low
MLS	None	None
MLT	None	None

Table 11. Rhetorical moves that differ significantly in the clausal complexity.

Measure	Significantly high	Significantly low
MLC	M5	M4
T/S	M4	M5
C/S	M5	M1, M2, M3, M4
DC/C	M5	M1, M2, M3, M4
DC/T	M4	M5
C/T	M5	M2, M3
CT/T	M5	M2, M3

Table 12. Rhetorical moves that differ significantly in the phrasal complexity.

Measure	Significantly high	Significantly low
CP/C	M3	M1
CP/T	M3,M4	M5
CN/C	M5	M4
CN/T	M1, M2, M4	M5
VP/T	M4	M5

overall mean. Sentences representing M2 and M3 also contain significantly fewer C/T (Mean = 1.36) and CT/T (Mean = 1.16) than the overall mean, while sentences representing M1, M2, M3, and M4 have significantly fewer DC/C compared to the overall mean.

Additionally, sentences realizing M4 (Mean = 0.68, SD = 0.60) contain a significantly higher ratio of dependent clauses per T-unit (DC/T) compared to the overall mean, while sentences realizing M5 (Mean = 0.23, SD = 0.27) have a lower DC/T ratio than the overall mean.

Regarding the phrasal level, for sentences representing M3, it is found that they contain significantly more coordinate phrases per clause (CP/C) (Mean = 0.92, SD = 0.79) and coordinate phrases per T-unit (CP/T) (Mean = 1.05, SD = 1.04) compared to the overall mean. This means that M3 sentences tend to have more coordination structures within each clause and within each T-unit.

On the other hand, sentences representing M1 contain significantly fewer CP/C (Mean = 0.49, SD = 0.56) than the overall mean which suggests that these sentences have fewer coordinate structures within each clause. Similarly, sentences representing M5 contain significantly fewer CP/T (Mean = 0.39, SD = 0.47) compared to the overall mean.

Moving on to the complex noun clauses, it is observed that sentences representing M5 contain a significantly higher number of complex noun clauses per clause (CN/C) (Mean = 3.97, SD = 1.82) compared to the average, whereas sentences representing M4 exhibit significantly fewer CN/C (Mean = 2.95, SD = 1.37) compared to the average.

Additionally, sentences representing M1 (Mean = 4.52), M2 (Mean = 4.54), and M4 (Mean = 4.66) contain significantly more complex noun phrases per T-unit (CN/T) compared to the average, while sentences representing M5 (Mean = 0.54, SD = 0.62) contain significantly fewer CN/T compared to the average.

In terms of phrasal complexity measured by verb phrases per T-unit (VP/T), sentences in M4 (Mean = 2.42, SD = 1.14) noticeably contain significantly more verb phrases than the average, whereas sentences representing M5 (Mean = 1.58, SD = 0.83) contain significantly fewer VP/T compared to the average.

3.2. Syntactic Complexity in Rhetorical Moves of CWA Abstracts

In relation to the second research question, as indicated in Table 13, the statistical analysis revealed significant differences between the rhetorical moves for all 14 measures in CWA. However, the results of the ANOM analysis did not align with the findings of the ANOVA.

Table 13. Significant difference among five-rhetorical moves in 14 syntactic complexity indices.

Syntactic complexity measures	CWA Asymptotic Sig. (2-sided test)
Global level	
MLS	0.001*
MLT	0.037*
Clausal level	
MLC	0.001*
DC/T	0.001*
C/S	0.001*
T/S	0.000*
C/T	0.001*
CT/T	0.001*
DC/C	0.021*
Phrasal level	
CP/C	0.001*
CN/C	0.023*
CN/T	0.001*
VP/T	0.001*
CP/T	0.001*

Note: * indicates that the measure has significant differences in five rhetorical moves.

As shown in **Table 14** and from the results of ANOM (see **Tables 15-19**), at the global level, sentences that fulfill M4 functions (M = 31.59) are significantly

Table 14. Descriptive statistics of global complexity between different moves in CWA. N = 449.

Moves	M1		M2		M3		M4		M5	
N	96		83		99		98		73	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MLS	24.50*	6.29	26.63	11.67	28.97	11.15	31.59**	17.42	30.30	12.43
MLT	25.76*	8.96	28.51	12.14	27.30	15.18	30.43	18.31	30.87**	13.87

Note: * indicates that the measure has the fewest numbers or shortest length in that move; ** indicates that the measure has the most numbers or longest length in that move.

Table 15. Rhetorical moves that differ significantly in the global complexity.

Measure	Significantly high	Significantly low
MLS	M4	M1
MLT	None	None

Table 16. Descriptive statistics of clausal complexity between different moves in CWA. N = 449.

Moves	M1		M2		M3		M4		M5	
N	96		83		99		98		73	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MLC	16.37	4.40	19.58	9.32	21.45**	12.08	15.41*	3.68	20.61	9.70
T/S	1.00	0.25	0.96*	0.25	1.06	0.34	1.08**	0.53	1.03	0.25
C/S	1.57	0.49	1.48	0.62	1.26*	0.48	2.13**	1.26	1.58	0.91
C/T	1.63	0.52	1.61	0.74	1.18*	0.47	2.06**	1.29	1.62	0.98
CT/T	0.38	0.29	0.32	0.37	0.22*	0.27	0.51**	0.25	0.34	0.40
DC/C	0.33	0.20	0.32	0.27	0.15*	0.19	0.41**	0.16	0.27	0.26
DC/T	0.62	0.49	0.64	0.72	0.25*	0.38	0.89**	0.77	0.63	0.88

Note: * indicates that the measure has the fewest numbers or shortest length in that move; ** indicates that the measure has the most numbers or longest length in that move.

Table 17. Rhetorical moves that differ significantly in the clausal complexity.

Measure	Significantly high	Significantly low
MLC	M3	M1, M4
DC/T	None	None
T/S	None	None
DC/C	M4	M3
C/S	M4	M3
C/T	M4	M3
CT/T	M4	M3

Table 18. Descriptive statistics of phrasal complexity between different moves in CWA. N = 449.

Moves	M1		M2		M3		M4		M5	
N	96		83		99		98		73	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
CP/C	0.59*	0.36	0.71	0.72	1.03**	0.95	0.60	0.32	0.94	0.84
CP/T	0.93*	0.65	0.96	0.83	1.21	1.13	1.12	0.78	1.35**	1.15
CN/C	2.85	0.99	3.20	1.54	3.19	2.06	2.45*	0.70	3.33**	1.60
CN/T	4.50	1.93	4.69	2.20	3.84*	2.59	4.85	2.80	5.19**	3.03
VP/T	2.27	0.81	2.60	1.12	1.82*	1.05	2.86**	1.63	2.57	1.37

Note: * indicates that the measure has the fewest numbers or shortest length in that move; ** indicates that the measure has the most numbers or longest length in that move.

Table 19. Rhetorical moves that differ significantly in the phrasal complexity.

Measure	Significantly high	Significantly low
CP/C	M3	M1, M4
CN/C	None	M4
CP/T	None	None
VP/T	M4	M3
CN/T	None	M3

longer in terms of mean length of sentences (MLS) compared to the overall mean. Conversely, sentences that relay M1 functions (M = 24.50) are significantly shorter compared to the average in MLS. Although there are significant differences in global complexity measured by mean length of text (MLT) across the five moves, none of the moves display significantly higher or lower MLT than the average.

As shown above, at the clausal level, sentences that fulfill M3 functions (Mean = 21.45) are significantly longer in terms of mean length per clause (MLC) compared to the overall mean. On the other hand, sentences that serve M1 (Mean = 16.37) and M4 (Mean = 15.41) functions are significantly shorter than the overall mean.

Regarding the number of subordinate structures at the clausal level, sentences that realize M4 functions contain a significantly higher number of clauses per T-unit (C/T), complex T-units per T-unit (CT/T), dependent clauses per clause (DC/C), dependent clauses per T-unit (DC/T), and clauses per sentence (C/S) compared to the overall mean.

Conversely, sentences that fulfill M3 functions have significantly fewer C/T, CT/T, DC/C, DC/T, and C/S than the overall mean. As for dependent clauses per T-unit (DC/T) and T-units per sentence (T/S), the ANOM analysis results show no rhetorical move that significantly differs from the overall mean of DC/T

and T/S at the clausal level.

At the phrasal level, sentences relaying the function of M3 (Mean = 1.03) contain significantly more coordinate phrases per clause (CP/C) than the overall mean. On the other hand, sentences that serve the functions of M1 (Mean = 0.59) and M4 (Mean = 0.60) have significantly fewer coordinate phrases compared to the overall average.

In terms of verb phrases per T-unit (VP/T), sentences that fulfill the function of M4 (Mean = 2.86) have more verb phrase structures compared to the average. However, sentences that fulfill the function of M4 (Mean = 2.45) contain significantly fewer complex nominals per clause (CN/C) compared to the average. Sentences that realize M3 (Mean = 3.84) also have fewer complex nominals per T-unit (CN/T) compared to the overall average. It's important to note that none of these indicators show significant differences between rhetorical moves. The results of the non-parametric Kruskal-Wallis tests did not support the findings of the ANOM analysis.

To summarize, the analysis revealed significant differences between the rhetorical moves for various measures in the two corpora. The expert group shows greater variability in the use of subordinate structures compared to the novice group. In terms of phrasal complexity, none of the rhetorical moves exceeds the average values for the CN/C and CN/T indicators within the group. This suggests that the novice group has a similar frequency of use in phrasal complexity, which is significantly different from the expert group. Furthermore, while most of the differences between rhetorical moves in both groups are significant in only one indicator compared to the average, there are also instances where moves differ in two or more indicators. This indicates that expert writers tend to use more or less complex sentences to achieve different rhetorical goals.

3.3. Differences of Syntactic Complexity in Each Move of Abstracts between the Two Corpora

Based on the results above, to analyze the differences between IWA and CWA in terms of syntactic complexity measures across five rhetorical moves, a Mann-Whitney U test will be conducted on the data of both groups, as shown in Table 20. The notable differences in syntactic complexity between IWA and CWA across the five rhetorical moves are presented.

In sentences that fulfil the function of M1, there are significant variations in the number of words per sentence (MLS), the amount of clausal coordination (T/S), and the amount of phrasal coordination (CP/C, CP/T) between the two groups.

For sentences that serve the function of M2, there is no significant difference in global syntactic complexity. However, there are significant differences in the amount of clausal subordination (C/T, DC/C, DC/T) and the degree of phrasal sophistication (VP/T) between the two groups.

Sentences that realize M3 do not exhibit any significant differences in each index between the two groups.

Table 20. Significant differences in syntactic complexity between IWA and CWA across the five rhetorical moves.

Measures IWA-CWA	M1	M2	M3	M4	M5
Global level					
MLS	**0.010***	0.157	0.275	0.786	0.001*
MLT	0.654	0.570	0.989	0.204	0.001*
Clausal level					
MLC	0.962	0.166	0.956	0.127	0.002*
DC/T	0.431	0.001*	0.107	0.016*	0.297
DC/C	0.458	**0.001***	0.174	0.072	0.270
T/S	0.029*	0.095	0.113	**0.012***	0.783
C/T	0.369	**0.008***	0.08	**0.005***	0.507
CT/T	0.390	0.181	0.079	0.781	0.858
C/S	0.365	0.241	0.075	0.142	0.471
Phrasal level					
CP/C	0.007*	0.845	0.637	0.994	0.001*
CN/C	0.563	0.188	0.192	0.014*	0.043*
CP/T	0.010*	0.401	0.627	0.383	**0.000***
VP/T	0.806	0.028*	0.454	0.008*	0.267
CN/T	0.953	0.712	0.167	0.944	0.071

Note: * indicates that the measure has significant differences in five rhetorical moves.

In sentences that fulfill M4, there are significant differences in the amount of clausal subordination (C/T, DC/T), clausal coordination (T/S), and the degree of phrasal sophistication (CN/C, VP/T).

For sentences that realize M5, there are significant differences in each index when considering global complexity. However, at the clausal level, the only significant difference is in the elaboration (MLC) between the two groups. In terms of the amount of phrasal coordination (CP/C, CP/T) and the degree of phrasal sophistication (CN/C), there exists a significant difference between the two groups.

3.3.1. Global Complexity and Rhetorical Goals

A manual review of long sentences associated with moves that contained significantly larger or smaller proportions of long sentences revealed insights into the relationship between length of sentence and these rhetorical aims. As illustrated in Example 1, international expert writers often synthesized or established connections between specific items of previous research, thus resulting in many longer sentences. Example 2 demonstrates a distinct tendency in announcing the present research by stressing the significance of the research. For another, Chinese novice authors often included descriptions of the specific suggestion items

or limitations to their conclusion statements, such as the 'just pay heeds to' clause below, and sometimes separated their conclusion statements into stages reflecting theoretical and practical significance, thus lengthening the sentence as shown in Examples 4 and 5. In contrast, in realizing the rhetorical function of M5, the experts often used short and direct clauses that connected a label with its findings, as demonstrated in Example 3.

EX.1. Interpreting research findings in doctoral thesis discussions is a demanding rhetorical task for writers, as it requires them to both make propositions of their own findings and engage with previous scholarship by evaluating others' findings in a way that their academic discourse community finds acceptable. (M1, 45, IWA)

EX.2. Building a good interpersonal relationship with the public can accelerate the resolution of the crisis and facilitate the process of reversing negative public opinions. (M1, 24, CWA)

EX.3. These findings contribute to provide some insight on students' learning trajectory and can inform appropriate educational interventions. (M5, 17, IWA)

EX.4. Based on the above findings, this study puts forward the following suggestions for the teaching of verb-noun collocations for Chinese college students: 1) Raising the awareness of collocation; 2) Introducing dictionaries and corpora into classroom teaching; 3) Increasing metacognition of English collocation production and use. (M5, 45, CWA)

EX.5. Additionally, the study just pays heeds to the development of students' group IC, which is a cut-in point in future studies for further conducting relevant research of individuals' IC. (M5, 29, CWA)

3.3.2. Clausal Complexity and Rhetorical Goals

At the clausal level, differences in clausal complexity were observed between IWA and CWA for the five rhetorical moves. Higher T/S values were found for sentences conveying M1 and M4 in IWA than CWA. While in CWA, higher DC/C and DC/T values were found for sentences conveying M2 and M4. Additionally, CWA had higher clausal coordination values (C/T) than IWA for sentences conveying M2 and M4. These findings highlight the importance of considering clausal complexity and coordination to improve writing logic and academic style.

This result indicated expert writers deliberately attached the coordinate clauses to relay the function of the introduction parts and results parts. For another, the Chinese novice writers attached more subordination to realize the purpose parts and method parts. Examples 6, 7, 8, 9 illustrated these differences. In Example 6, the expert writers summarize the results parts from the research by utilizing various coordinate clauses, like "doctors and students; adjuncts, interpersonal, existential and pronoun themes, and marked themes." While in Chinese novice writers' abstract writing, the subordinate structures are utilized while delivering the results parts. As shown in Example 7, the first sentence contained three dependent clauses, one adverbial clause of manner and two adverbial

clauses of purpose. In sentences relaying the function of method parts, the Chinese novice writers also employed multiplied dependent clauses to make the purpose statements more convincing and more concrete as shown in Example 9. The use of nonfinite clauses in statements which announce the present study are observably associated with "research verb to verb" constructions, as evidenced by Example 8. In addition, these two examples illustrate the different importance of nonfinite gerunds, either connected with of-phrases, (e.g., construction of; the goal of) or other prepositions (e.g., with; by) to include methodological, theoretical, or implication-based information in propositionally dense purpose statements.

EX.6.

The findings show that doctors and students differ significantly in their use of specific theme types indicating their different understanding and use of the oral case presentation as well as their social position in the professional field. Indexing students' novice status are differences in the use of conjunctive adjuncts, interpersonal, existential and pronoun themes, and marked themes as overt signaling of the presentation structure. (M4, IWA)

EX.7.

The author of this paper has found that the interpretation devices like amplification, replacement and omission were mainly utilized by the interpreters to mediate the hedges so as to realize the purposes of interpreting in the Premiers' Press Conferences. The appropriate use of the hedges by the interpreters could not only boost the efficacy of expression instead of weakening the accuracy of diplomatic language, but also facilitate the cross-cultural communication and realize the delivery of Chinese voice and stance. Meanwhile, the proper use of the hedges by the interpreters could present the courtesy, highlight the diplomatic protocol and avoid the embarrassment, conflict and potential risk in the diplomatic setting. (M4, CWA)

EX.8.

This paper investigates the academic writing challenges encountered by L2 postgraduate students in Engineering and the strategies they developed (or, from the perspective of faculty, should develop) to address these issues. (M2, IWA)

EX.9.

This dissertation is dedicated to exploring the linguistic features and pragmatic functions of reported speech used in courtroom discourse by means of both quantitative and qualitative methods. Concretely speaking, the research is intended to investigate how reported speech is used in the trial, what features are showed when language users in the court make linguistic choices in the selection of reported voices, construction of reported messages and expression of reporting attitudes with the goal of achieving their communicative purposes, and what pragmatic functions are fulfilled in the trial by making different linguistic choices in reported speech. (M2, CWA)

3.3.3. Phrasal Complexity and Rhetorical Goals

At the phrasal level, differences in phrasal complexity were evident between

IWA and CWA across the five rhetorical moves. IWA exhibited higher values of complex nominals per clause (CN/C) in sentences conveying M5 and sentences relaying the function of M4, indicating the use of more complex nominal structures. On the other hand, CWA had higher CP/C values in sentences conveying M1 and M5, suggesting a greater use of coordinated phrases. Moreover, CWA had higher values of CP/T, indicating a higher frequency of coordinated phrases per T-unit, for both M1 and M5; the same situation goes with the verb phrases per T-unit measure when relaying the function of M2 and M4. These findings emphasize the significance of phrasal complexity and coordination in achieving specific rhetorical goals.

EX.10.

These findings contribute to provide some insight on students' learning trajectory and can inform appropriate educational interventions. (M5, IWA)

EX.11.

The research values of the present study lies in that, theoretically, it manifests the feasibility and practicality of the analytical framework in current research which combines the Linguistic Adaptation Theory with the Principle of Goal Direction, and it provides some references to further studies. Practically, based on the major findings, some suggestions are proposed for judicial practice. For example, witnesses should be encouraged to appear in the court so that the use of reported speech voiced from them can be reduced, the quotations of law articles are supposed to be reported more in order to enhance legal awareness. (M5, CWA)

Overall, these findings suggest that improving writing logic and academic style requires attention to phrasal complexity and coordination. It is important to use coordinated phrases and complex nominal structures appropriately and effectively in different rhetorical moves. By paying attention to these aspects, the quality and academic level of writing can be enhanced, making it more logical, academic, and fluent.

4. Conclusions

According to the results above, the following conclusions can be drawn: there are significant differences in overall sentence complexity between IWA and CWA. In terms of subordinate clause complexity, IWA exhibits different characteristics when conveying different rhetorical devices, while CWA also shows significant differences when conveying different rhetorical devices. There are group differences in subordinate clause coordination between different rhetorical devices at the sentence level. Overall, IWA and CWA exhibit different characteristics when conveying different rhetorical devices. At the phrase level, there are significant differences in the complexity of coordinating phrases (CP/C) and complex noun phrases (CN/C) in sentences conveying M5 functionality.

Firstly, these results indicate that improving academic writing requires attention to the complexity and coherence of sentences and the use of rhetorical de-

vices. In the writing process, efforts should be made to achieve a balance in sentence length and avoid overly long or short sentences. Additionally, attention should be paid to the complexity and coordination of subordinate clauses to ensure logical and coherent sentence structures. For different rhetorical devices, appropriate sentence structures and phrase complexity should be chosen based on the goals. By paying attention to these aspects, writing fluency and academic level can be improved.

Secondly, as for Chinese postgraduates, the disproportionate focus on subordinate clauses and "over-use" of complex phrases might deserve more attention. On the one hand, these issues are closely related to academic writing norms, writer identity construction, and international academic discourse status, all of which influence the students' academic writing practice. On the other hand, these academic writing skills are expected to be acquired by students at the graduate level when dealing with academic studies. Still, many of them fail to deploy linguistic strategies to construct compact writing tactically. The underlying notion is that academic writing for Chinese writers is supposed to move beyond blindly or mechanically embedding multiple modifiers, but attempt to raise awareness on shift to an integrated combination of noun phrase features concerning more advanced stages. Specifically, as advanced writers demonstrate movement toward more complex phrasal structure, Chinese postgraduates are supposed to push the inherent notion of academic writing away from the over-use of simple nominal structures at early stages (e.g., premodifying nouns, and of-phrases), and moving more toward producing complex non-clausal phrases and lengthy post-modifying nouns at later stages.

Our results also show that advanced writers will vary their choices in the use of complex syntactic structures depending on their rhetorical goals. Results pertaining to our third research question revealed significant variation across the rhetorical moves in the degree of syntactic complexity assessed using almost full indices considered in the two sub-corpora, indicating that different rhetorical functions may entail greater or lesser use of different complex structures. Furthermore, while most rhetorical moves in those two corpora differed significantly from the group in only one measure, a few moves differed in two or more measures, indicating expert writers' tendency to employ especially more or less complex sentences to realize those functions.

What's more, the specific comparison show that Chinese graduate writers are less rich and flexible in syntactic complexity by achieving different rhetorical functions than international advanced writers, as reflected in their usage preference for simpler linguistic forms to perform communicative goals and the inability to express themselves flexibly and accurately through "paradigmatic". This poses a new challenge for the teaching of English writing for academic purposes, and it is worthwhile for teachers to consider what ways they can overcome the bottleneck in EAP. In addition, this research will further investigate the differences in the frequencies and proportions of each type of syntactic complexity in two sub-corpora, so as to present a thorough understanding of graduate L2

writers and advanced writers in EAP writing.

Acknowledgements

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Samraj, B. (2005) An Exploration of a Genre Set: Research Article Abstracts and Introductions in Two Disciplines. *English for Specific Purposes*, **24**, 141-156.
https://doi.org/10.1016/j.esp.2002.10.001

[2] Casal, J.E., Lu, X. and Qiu, X. (2021) Syntactic Complexity across Academic Research Article Part-Genres: A Cross-Disciplinary Perspective. *Journal of English for Academic Purposes*, **52**, Article ID: 100996.
https://doi.org/10.1016/j.jeap.2021.100996

[3] Lu, X., Casal, J.E. and Liu, Y. (2020) The Rhetorical Functions of Syntactically Complex Sentences in Social Science Research Article Introductions. *Journal of English for Academic Purposes*, **44**, Article ID: 100832.
https://doi.org/10.1016/j.jeap.2019.100832

[4] Lu, X., Casal, J.E. and Liu, Y. (2021) Towards the Synergy of Genre- and Corpus-Based Approaches to Academic Writing Research and Pedagogy. *International Journal of Computer-Assisted Language Learning and Teaching*, **11**, 59-71.
https://doi.org/10.4018/IJCALLT.2021010104

[5] Lu, X. and Casal, J.E. (2021) "Maybe Complicated Is a Better Word": Second Language English Graduate Student Responses to Syntactic Complexity in a Genre-Based Academic Writing Course. *International Journal of English for Academic Purposes: Research and Practice*, **1**, 95-114. https://doi.org/10.3828/ijeap.2021.7

[6] Hyland, K. (2004) Disciplinary Discourses: Social Interactions in Academic Writing. University of Michigan Press, Ann Arbor.

[7] Crossley, S.A. and McNamara, D.S. (2014) Does Writing Development Equal Writing Quality? A Computational Investigation of Syntactic Complexity in L2 Learners. *Journal of Second Language Writing*, **26**, 66-79.
https://doi.org/10.1016/j.jslw.2014.09.006

[8] Lu, X. (2011) A Corpus-Based Evaluation of Syntactic Complexity Measures as Indices of College-Level ESL Writers' Language Development. *TESOL Quarterly*, **45**, 36-62. https://doi.org/10.5054/tq.2011.240859

[9] 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

[10] Yang, W., Lu, X. and Weigle, S.A. (2015) Different Topics, Different Discourse: Relationships among Writing Topic, Measures of Syntactic Complexity, and Judgments of Writing Quality. *Journal of Second Language Writing*, **28**, 53-67.
https://doi.org/10.1016/j.jslw.2015.02.002

[11] Pallotti, G. (2009) Caf: Defining, Refining and Differentiating Constructs. *Applied*

Linguistics, **30**, 590-601. https://doi.org/10.1093/applin/amp045

[12] Hunt, K.W. (1965) Grammatical Structures Written at Three Grade Levels. NCTE Research Report No. 3. National Council of Teachers of English, Champaign.

[13] Casanave, C. (1994) Language Development in Students' Journals. *Journal of Second Language Writing*, **3**, 179-201. https://doi.org/10.1016/1060-3743(94)90016-7

[14] Stockwell, G. and Harrington, M. (2003) The Incidental Development of L2 Proficiency in NSNNS Email Interactions. *CALICO Journal*, **20**, 337-359. https://doi.org/10.1558/cj.v20i2.337-359

[15] Larsen-Freeman, D. (1978) An ESL Index of Development. *TESOL Quarterly*, **12**, 439-448. https://doi.org/10.2307/3586142

[16] Ferris, D.R. (1994) Lexical and Syntactic Features of ESL Writing by Students at Different Levels of L2 Proficiency. *TESOL Quarterly*, **28**, 414-420. https://doi.org/10.2307/3587446

[17] Zhang, X. and Lu, X. (2022) Revisiting the Predictive Power of Traditional vs. Fine-Grained Syntactic Complexity Indices for L2 Writing Quality: The Case of Two Genres. *Assessing Writing*, **51**, Article ID: 100597. https://doi.org/10.1016/j.asw.2021.100597

[18] Lei, L. (2017) Research on Syntactic Complexity in Academic Writing by Chinese English Learners. *Journal of PLA University of Foreign Languages*, **40**, 1-10+159.

[19] Norris, J.M. and Ortega, L. (2009) Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity. *Applied Linguistics*, **30**, 555-578. https://doi.org/10.1093/applin/amp044

[20] Biber, D., Gray, B. and Poonpon, K. (2011) Should We Use Characteristics of Conversation to Measure Grammatical Complexity in L2 Writing Development? *TESOL Quarterly*, **45**, 5-35. https://doi.org/10.5054/tq.2011.244483

[21] Wolfe-Quintero, K., Inagaki, S. and Kim, H.Y. (1998) Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity. University of Hawaii Press, Honolulu.

[22] Lu, X. (2010) Automatic Analysis of Syntactic Complexity in Second Language Writing. *International Journal of Corpus Linguistics*, **15**, 474-496. https://doi.org/10.1075/ijcl.15.4.02lu

[23] Biber, D., Gray, B., Staples, S. and Egbert, J. (2020) Investigating Grammatical Complexity in L2 English Writing Research: Linguistic Description versus Predictive Measurement. *Journal of English for Academic Purposes*, **46**, Article ID: 100869. https://doi.org/10.1016/j.jeap.2020.100869

[24] Cortes, V. (2013) The Purpose of This Study Is to: Connecting Lexical Bundles and Moves in Research Article Introductions. *Journal of English for Academic Purposes*, **12**, 33-43. https://doi.org/10.1016/j.jeap.2012.11.002

[25] Omidian, T., Shahriari, H. and Siyanova-Chanturia, A. (2018) A Cross-Disciplinary Investigation of Multi-Word Expressions in the Moves of Research Article Abstracts. *Journal of English for Academic Purposes*, **36**, 1-14. https://doi.org/10.1016/j.jeap.2018.08.002

[26] Durrant, P. and Mathews-Aydınlı, J. (2011) Identifying Formulaic Language in EAP Writing: A Computational Approach. *English for Specific Purposes*, **30**, 222-235. https://doi.org/10.1016/j.esp.2010.05.002

[27] Le, T.N.P. and Harrington, M. (2015) Phraseology Used to Comment on Results in the Discussion Section of Applied Linguistics Quantitative Research Articles. *English for Specific Purposes*, **39**, 45-61. https://doi.org/10.1016/j.esp.2015.03.003

[28] Yin, S., Gao, Y. and Lu, X. (2021) Syntactic Complexity of Research Article Part-Genres: Differences between Emerging and Expert International Publication Writers. *System*, **97**, Article ID: 102427. https://doi.org/10.1016/j.system.2020.102427

[29] Jiang, L. and Kang, M.C. (2021) Investigating Syntactic Complexity across Rhetorical Moves of Abstracts from International Journals. *Journal of Shenyang University*, **23**, 569-585.

[30] Hyland, K. (2000) Disciplinary Discourses: Social Interactions in Academic Writing. *English for Specific Purposes*, **19**, 51-70.

[31] Kyle, K. (2016) Measuring Syntactic Development in L2 Writing: Fine Grained Indices of Syntactic Complexity and Usage-Based Indices of Syntactic Sophistication. Doctoral Dissertation, Georgia State University, Atlanta.