



Syntactic Complexity Development in Research Articles Over 25 Years: A Comparative Case Study of a Native and a Non-Native Applied Linguist

Abbas Zare-ee ^{*}, Zahra Falah Kordi

*English Department, Faculty of Literature and Foreign Languages, University of Kashan,
Kashan, Iran*

Received: 2026/05/23

Accepted: 2026/08/23

Abstract: The development of syntactic complexity is a major concern in academic writing. Research on English for Research Publication Purposes (ERPP) has greatly neglected longitudinal investigation of the development of syntactic complexity. Guided by the skill acquisition theory, this study examined syntactic complexity over two decades in the articles published by a native and a non-native scholar, focusing on changes in 14 syntactic measures. The aim was to document change over time and an increase in writing experience. Applying a retrospective case study design, the researchers analyzed abstracts and conclusions of 75 articles written from 2000 to 2024, using the L2 Syntactic Complexity Analyzer (L2SCA), a web-based tool that calculates syntactic complexity. We analyzed the resulting numerical data using linear regression to identify trends in syntactic change. The findings revealed that the non-native writer's abstracts improved syntactically with time/experience, favoring longer and more sophisticated sentences. In abstracts, the native writer showed a higher association with time in sentence length than the non-native writer (annual increase of 0.25 for the native and 0.08 for the non-native writer). In conclusion sections, the non-native writer progressively simplified clauses, while the native writer demonstrated a higher association with time in sentence length (-0.4 for non-native and +0.4 for native). The findings underscore the role of publishing experience over time and first language background as variables related to syntactic development in ERPP based on the calculated regressions. Results also imply that support for the learning of syntactic sophistication should be an inseparable part of ERPP pedagogy.

Keywords: EFL Writing, English for Academic Purposes (EAP), English for Research Publication Purposes (ERPP), Retrospective Case Study, Syntactic Complexity.

* Corresponding Author.

Authors' Email Address:

¹ Abbas Zare-ee (abbaszarei72@yahoo.co.uk), ² Zahra Falah Kordi (zare-ee72@kashanu.ac.ir)



Introduction

Publishing research reports in English is a daunting task for many scholars, especially for non-native speakers (NNS). Successful publishing, however, is essential for them because it creates societal impact, fosters professional development, leads to academic promotion, and stimulates future research (Flowerdew, 2013). In today's academic world, productivity plays a major role in shaping careers and outlooks, and it can be considered a key to manifesting scholarly productivity (Habibie, 2022). Publishing research is a pathway to career growth, professional recognition, and making an impact in society. Scholars are judged, ranked, and even paid by the number of their publications, the place of their publications, and the number of citations. Their publications become food for thought in society in general and in academia in particular. The task is, however, very complicated to master.

The argument which informs the present research is that, as a major step in learning to write academically, the improvement of syntactic complexity and grammatical accuracy can pose problems for native and non-native speakers. First, the native speaker may have no problem with basic interpersonal communication skills, but will require training in what is usually referred to as cognitive academic language proficiency (Brown & Lee, 2025). In fact, in second language writing, proficiency in English is not considered the same as writing competence. At advanced levels, the distinction between native and non-native speakers becomes less relevant, as writing skills are developed through formal education rather than early language acquisition (Hyland, 2015). Second, according to Swales (2004), the difficulties typically experienced by NNES scholars are similar in many ways to those experienced by native speakers as they relate to higher-order abilities in language. As Flowerdew and Habibie (2021) explained, publishing is not just about fixing grammar or polishing sentences; it involves learning how to craft arguments, structure ideas, and follow the unwritten rules of academic publishing. Third, apart from language knowledge or system knowledge in which the native speaker may exhibit few problems, successful publication of research reports in English requires world knowledge, writing process knowledge, genre knowledge, and even research methodology knowledge. Fourth, based on arguments mounted by Hyland and Shaw (2016), writing in English is a second language for all, including its native speakers, so that the claims of linguistic injustice in publication become questionable. Finally, for both native and non-native writers of English, there are always opportunities for further growth and development in writing and especially in English for research publication purposes (ERPP). Becoming an advanced academic writer with extensive, influential publishing is challenging, and even experienced researchers struggle hard to increase their writing productivity.

Because of the complexities involved in writing for publication in English, attempts to help writers have expanded in recent decades, as evidenced by the development of ERPP as a branch of EAP that aims to address the concerns of professionals and (post)graduate students in publishing (Bitchener, 2018; Cargill & Burgess, 2008). ERPP is especially expanding in international academic contexts to help researchers who want to publish in international journals (Flowerdew & Habibie, 2021). ERPP emerged as a response to the growing dominance of English as lingua franca, to address the need for international researchers to write and share their work in English. Publishing difficulties often stem from language barriers or other non-linguistic factors, such as being based outside the major center of knowledge production (Canagarajah, 1996). Most ERPP research has addressed problems of scholars whose first language is not English, even though, as Hyland and Shaw (2016) contend, there is inadequate evidence to show that there is bias against non-native English speakers. ERPP is a specialized approach to teaching English and academic literacy that supports everyone writing for research publication, whether they are non-native English speakers or native English-speaking scholars. The genre analysis approach, aiming to understand how different groups of writers and individuals go about the process of writing for publication (Flowerdew & Habibie, 2021), is yet another growing area that attempts to clarify differences in published materials produced by native and non-native scholars. If challenges for both native and non-native writers are in many ways similar at advanced levels in ERPP, as argued (Canagarajah, 1996; Hyland, 2015; Swales, 2004), are developmental paths at advanced levels also similar? Comparative studies that explore how native and non-native writers develop in scholarly publishing can throw more light not only on the challenges but also on their possible solutions.

In spite of attempts to face linguistic challenges in scholarly publishing through research and instruction in ERPP (e.g., Flowerdew & Habibie, 2021), little is known about the comparative development of syntactic sophistication and complexity in scholarly publishing by native and non-native speakers. Increasing levels of sophistication, accuracy and complexity in writing have been cross-sectionally studied in relation to language knowledge and writing proficiency. However, longitudinal case studies examining syntactic complexity development in published research articles by native and non-native writers are missing. Little is known about how native and non-native writers modify the syntactic sophistication/complexity of their writing over time. Existing studies (e.g., Wang & Lu, 2026; Yoon & Polio, 2017) rarely provide longitudinal analyses that trace changes in syntactic measures over time or link these developments to writing experience. This lack of

chronological and contextual investigation limits understanding of how non-native scholars develop syntactic sophistication for publication. Furthermore, research examining the academic writing experiences of non-native scholars who have studied and worked in non-English-speaking countries is scarce (Curry & Lillis, 2004). The present study addressed the gap by examining the long-term development of syntactic complexity in the writings of a native and a non-native scholar, contributing to a more nuanced understanding of how experience and L1 background shape academic writing in ERPP. The study aimed to explore how syntactic complexity evolves in the writing of native and non-native English writers, with a specific focus on published research articles. One purpose was to explore the development of 14 syntactic measures in the writings of a non-native scholar and a native speaker over a long period of two decades. Another purpose was to explore the possible role of L1 by comparing the two writers and the effect of time experience on the development of these syntactic measures through longitudinal retrospection. This study sought to answer the following two research questions:

- 1) Do measures of syntactic complexity change over time in the published works of the native and the non-native scholar?
- 2) Are there any significant differences in changes in complexity measures between the native speaking scholar and the non-native EFL scholar?

Review of the Literature

Theoretical Framework

One ERPP tradition of research in the study of published articles is the study of syntactic complexity, which relates to the variety of language forms a scholar uses and how advanced or complex those forms are (Ortega, 2003). Over the past few decades, the importance of syntactic complexity in second language (L2) writing has been recognized, and the theoretical definition of syntactic complexity has expanded. In second language acquisition (SLA) research, syntactic complexity is typically treated as a dependent variable that reflects linguistic development and the learner's ability to use diverse grammatical resources effectively (Ortega, 2015). It should be noted that within the complexity, accuracy, and fluency (CAF) framework, complexity involves structural variety, accuracy concerns rule conformity, and fluency reflects ease of expression (Wolfe-Quintero et al., 1998).

Based on their review of studies addressing the construct of syntactic complexity, the measurement of complexity, and the association of syntactic complexity with writing quality in a second language, Lei et al. (2023) claimed that the literature suggests four assumptions.

First, syntactic complexity is a complex construct that cannot be measured by a single index. Second, task-related factors can influence syntactic complexity. Third, syntactic complexity can be a yardstick for the measurement of writing quality and proficiency in a second language. Finally, with the investment of more time in learning to write academically, syntactic complexity increases. [Biber et al. \(2011\)](#) argued that these studies employ a production perspective in which the expectation is that syntactic complexity increases as writers (both native and non-native speakers) become more proficient. The serious concern that they raised about these assumptions, however, is that there is little research evidence to show that T-unit-based measures of complexity are appropriate for the assessment of writing development.

Based on the related literature, the study of syntactic complexity in published articles in English should not address only non-native learners. [Habibie \(2016\)](#) writes that research on scholarly publication in English has mainly focused on non-native scholars and international graduate students, while the experiences of Anglophone novice scholars remain underexplored. He argues that little attention has been paid in related research to changes in the textual characteristics of native and non-native writers' scholarly production as they develop, probably with the assumption that they are scholars and they know it all in their status. The reality is that recent developments in academia have introduced strict systems of evaluation, making publication in high-ranking journals a major indicator of success and also a challenge for both native and non-native researchers and research institutions ([Flowerfew & Habibie, 2021](#)).

[Fazel's \(2018\)](#) longitudinal case study of publishing challenges experienced by two Canadian-born Anglophone doctoral students in language education clearly shows that articles by novice non-native scholars also deserve exploration. Their participants were both native speakers and had prior publication experience, but both encountered significant difficulties, indicating that it is novice writing status rather than native-speaker status that primarily shapes the challenges of writing for publication. Using a multidimensional model of genre knowledge (i.e., formal, process, rhetorical, and subject-matter knowledge), [Fazel \(2018\)](#) identified key challenges faced by the two natives. In terms of formal knowledge, both struggled with structuring articles. Their process-related challenges included navigating peer review, communicating with editors, selecting journals appropriate for their topic, avoiding submission to predatory journals, and having limited access to exemplars of submission and review genres (see also [Flowerdew & Li, 2020](#)). Their rhetorical challenges involved framing research to appeal to readers and disciplinary conventions. They also

experienced emotional pressure when they received critical feedback and rejection, demonstrating that even native-speaking novice scholars can find the publication process demoralizing (Hyland, 2015; Li, 2006). Being a non-Anglophone scholar adds extra challenges to those reported here (Mur-Dueñas, 2019).

This study aims to provide further research evidence on the development of syntactic complexity in academic writing. A comparative study of the textual characteristics of articles published by native and non-native speakers can present a clearer picture of the development of writing skills in ERPP.

Empirical Studies on Syntactic Complexity

The study of syntactic complexity in academic writing in English over the past few decades has focused on conceptualizing its nature, measuring it in different contexts, investigating factors that affect it, relating it to writing quality, and observing its inter- and intra-learner variability. For example, Ortega (2003) found that the mean length of sentence (MLS) and the mean length of T-unit (MLT) were reliable measures to distinguish students' writing in second language and foreign language settings. Such distinctions, however, may not necessarily reflect different developmental paths. Lu (2011) proposed 14 measures for syntactic complexity calculation. His measures are, of course, merely quantitative. As an example, studies that focused on influential factors, Yoon and Polio (2017) and, more recently, Wang and Lu (2026), found that the genre of writing affected syntactic complexity.

Istiharoh and Pranowo (2023) measured the development of students' syntactic complexity during an Essay Writing Course, using the L2SCA tool described in Table 1. Their work involved 20 university students, who were part of the researcher's class. Although students received guidance during the writing course, their syntactic development was not consistently progressive across all the measures, and the genre of the essay also affected their results. Mazgutova and Kormos (2015) investigated syntactic complexity development during a short but intensive 60-hour, 4-week writing course for advanced learners of English who received intensive writing practice and detailed written feedback. Their findings revealed significant progress in syntactic complexity among the younger and less proficient international student group. Vyatkina et al. (2015) also tested the assumption that linguistic development enables college learners to gradually grow in written syntactic complexity. By the end of their two years, the researchers observed the learners' evolving use of grammatical structures in their written work. These studies demonstrate that linguistic development can

occur over varying time frames and learners can access a broader and more sophisticated range of grammatical resources over time.

Lu and Ai (2015) investigated how first language (L1) background influences syntactic complexity in college-level L2 English argumentative writing. Analyzing 1,400 essays from non-native speakers (NNS) across seven L1 groups (German, Bulgarian, French, Russian, Tswana, Japanese, Chinese; 200 each), compared to 200 native-speaker (NS) essays, they applied the introduced 14 measures (via Lu's L2SCA analyzer). For all NNS as one group, significant differences from NS were seen in only three measures, with NNS generally lower in unit length and phrasal sophistication. However, for each L1 background alone, they found greater differences from the native speakers, implying that L1 background is a significant variable in the picture. Similarly, Flowerdew (2015) conducted a comparative analysis of research articles published by native and non-native scholars. He reported significant differences in syntactic complexity between native and non-native English speakers. Native speakers tended to use a broader range of syntactic structures, while non-native writers demonstrated slower growth in complexity over time.

More recently, Lei et al. (2023) explored the development of syntactic complexity in English compositions written by 1081 intermediate Chinese university students. Tracking their syntactic development over two years revealed that measures related to sentence length, coordination, and phrasal sophistication increased significantly over time, while the use of clausal subordination decreased. In another longitudinal case study, Polat et al. (2020) explored how Turkish learners of English develop written syntactic complexity over time. They analyzed 852 writing samples produced by 284 EFL learners across three semesters. The results demonstrated that learners at the lower proficiency levels primarily relied on simple structures like phrasal coordination, while learners at the intermediate level used more varied and complex constructions.

The Current Study

The studies reviewed above indicate that syntactic complexity varies over time. However, they do not reveal whether or not this variation is experienced by native and non-native writers alike. There have been many attempts in L2 education in general and in ERPP in particular to explore how syntactic complexity develops over time in academic writing. These studies mainly ignore developments in NS writing and mainly use it as a point of reference and comparison. Most related studies track the development of syntax over short time periods of a month to two years. The present work added methodological strength to the ongoing

discussion by adopting a retrospective case study design that looked at the articles published by a native and a non-native scholar between 2000 and 2024, which can potentially present a better picture of change over time. It also attempted to fill the gap by looking at established scholars' works in both native and non-native writing for publication in applied linguistics.

This work was delimited to the exploration of complexity, not as perceived in psycholinguistics or language typology, but as grammatically operationalized in applied linguistics and English language teaching. In spite of using extensive textual data, the study lacked a qualitative lens and was limited only to quantitative measures of syntactic complexity summarized in Table 1.

Table 1. L2SCA Measures of Syntactic Complexity Based on [Lu \(2011\)](#)

No	Code	Measure	Definition
1.	MLS	Mean Length of Sentence	The average number of words per sentence in the text
2.	MLT	Mean Length of T-Unit	The average number of words per T-unit
3.	MLC	Mean Length of Clause	The average number of words per clause
4.	C/S	Clauses per Sentence	The ratio of total number of clauses to total number of sentences
5.	C/T	Clauses per T-Unit	The ratio of total number of clauses to total number of T-units
6.	DC/C	Dependent Clauses per Clause	The ratio of dependent (subordinate) clauses to total clauses
7.	DC/T	Dependent Clauses per T-Unit	The ratio of dependent clauses to total T-units
8.	T/S	T-Units per Sentence	The ratio of total number of T-units to total number of sentences
9.	CT/T	Complex T-Units per T-Unit	The ratio of complex T-units (that is, T-units containing at least one dependent clause) to total T-units
10.	CP/C	Coordination Phrases per Clause	The ratio of coordinate phrases (phrases joined by coordinating conjunctions) to the number of clauses
11.	CP/T	Coordination Phrases per T-Unit	The ratio of coordinate phrases to number of T-units
12.	CN/C	Complex Nominals per Clause	The ratio of complex nominals (e.g., noun phrases with modifiers, post-modifiers, or embedded structures) to the number of clauses
13.	CN/T	Complex Nominals per T-Unit	The ratio of complex nominals to the number of T-units
14.	VP/T	Verb Phrases per T-Unit	The ratio of verb phrases to the number of T-units

As depicted in Table 1, Lu (2011) described syntactic complexity as the variety of sentence structures and the level of their advancement and sophistication. He proposed fourteen measures of syntactic complexity which form the units of analysis in many studies of text grammatical characteristics, including the current study.

Methodology

Design

The study used a retrospective case study design to explore the development over time of syntactic measures in two scholars' writing for publication in English, a native English-speaking scholar and a non-native EFL scholar. Quantitative text qualities were measured and analyzed chronologically. A retrospective longitudinal case study is a type of study that works on answering questions about the past during a single investigation point by gathering data over time (Dörnyei, 2007). The study measured, recorded, and analyzed syntactic complexity indices in English research articles published by both a native and a non-native scholar over a span of 25 years, from 2000 to 2024. Guided by Yin's (2009) classification of case study designs, a retrospective case study design was used with two cases, each representing a different first language (L1) background. The units of focused case analysis are depicted in Table 2.

Table 2. Units for Case Study Analysis of Scholar A and B's Publishing History

Focused unit of analysis 1: time and experience		Focused unit of analysis 2: Syntactic complexity measures
2000-2004	Novice	Mean length of clause
		Mean length of sentence
		Mean length of *T-unit
2010-2014		Clauses per T-unit
		Complex T-units per T-unit
		Dependent clauses per clause
2005-2009		Dependent clauses per T-unit
		Coordination phrases per clause
		Coordination phrases per T-units
2015-2019		T-Units per sentence
		Complex nominals per clause
		Complex nominals per T-unit
2020-2024	Experienced	Verb phrases per T-unit
		Clauses per sentence

As reflected in Table 2, both holistic and focused analyses were integrated. For each case, writing experience over time was considered. Then, for each writing sample, more detailed syntactic analysis was carried out. On a holistic level, the participants' publishing history was examined, providing a broad view of their academic writing trajectory. At a more focused level, syntactic complexity development was analyzed over time and experience, using 14 specific syntactic complexity measures as subunits of analysis based on the measures given by [Lu \(2011\)](#).

Corpus Data Collection

In the current study, the abstracts and conclusions of research articles published by two cases, a non-native and a native English-speaking scholar, were examined. The two cases were chosen purposively from among experienced linguists who published in their field of research continuously over a long time without notable interruption or disciplinary change. Reza (a pseudonym), as a non-native scholar, and Tom (also a pseudonym), as the native scholar, were the participants in the research. Both cases are established applied linguists who have written prolifically and published research over the 2000-2024 period examined in this study. Reza and Tom share many similar experiences in writing for publication in English, and they were targeted because of their common platform of reference, making the comparison possible. Both are full professors of applied linguistics. Both are actively engaged in research and scholarly publication on second language education. They both publish in (Q1-Q2) indexed journals in language and linguistics, and they are the main and corresponding author in their selected works, accessible to professionals through their institutions.

For research materials, the articles by both scholars covering the years 2000 to 2024 were collected from Google Scholar and related journal databases made accessible through our university. 75 academic research articles published in well-known, reliable journals by the two scholars were selected as qualified for inclusion because the case was the main author; the article was a research report, not a position or theoretical paper; it was in English and on issues in second language education; it was published between 2000-2024; and it was published in an indexed journal in the language and linguistics category. These inclusion criteria were considered in order to minimize the possible effect of extraneous variables such as co-authorship, shifts in article type, changes in journal quality, and disciplinary variations in writing. The researchers discussed the selection criteria, clarified ambiguities, and agreed on the selection of all texts. The selected final corpus included 40 articles by the non-native

scholar and 35 by the native scholar. The selection was proportionate to the total number of articles published by each case in a time period and the total number of words was also considered to make the comparison justifiable. For each article, the abstract and conclusion were selected for analysis. If an article didn't have a separate conclusion section, the discussion and conclusion part was used instead. The materials for the analyses were saved as text files. This involved removing irrelevant publication data from the PDF files and including only the main text of the abstract and the main text of the conclusion in separate .txt files for analysis. The files were analyzed one by one using the online L2 Syntactic Complexity Analyzer (L2SCA). The program yielded statistics on different measures of complexity. These were controlled and recorded in Excel files for further analyses. Table 3 depicts the details of the research materials.

Table 3. Abstract and Conclusion Word Counts by Language

Writer	Number of Articles	L1	section	Number of words	Total words
Native	35	English	abstract	6,688	28,817
			conclusion	22,129	
Non-native	40	Persian	abstract	7,872	28,498
			conclusion	20,626	

The procedures involved several steps. The first step was selecting and saving the different sections of each research article intended for analysis. The second step was entering individual abstracts and conclusions by the NS and NNS scholars into the web-based tool L2SCA for analysis. Each text was entered individually, with a word limit of 1,000 per submission. In the third step, the tool analyzed each text sample based on the 14 predefined measures of syntactic complexity. Within seconds, L2SCA produced detailed statistical data for each article section, quantifying the syntactic features for the texts. This process was repeated systematically for all 75 article samples. In the next step, the researcher compiled all the syntactic complexity statistics, tabulated the findings, and imported the data into SPSS software (Version 24) for further analysis and interpretation. For each article, the year was recorded along with the complexity markers to track how the use of syntactic structures may have evolved. By analyzing the numerical data generated for each article, patterns of development and change were carefully observed. Finally, the measures for the non-native scholar were compared with those of the native English-speaking scholar.

Results

Descriptive data on syntactic complexity measures in the academic texts written by both native and non-native scholars illustrated improvement over time in some of the complexity measures. As Table 4 depicts, in the texts produced by the non-native scholar over time, the average length of the 40 abstracts was 211.9 words. The lowest average was observed for "dependent clauses per clause", with a value of 0.37, while the highest averages were for "mean length of sentence" at 24.7, and "mean length of T-units" at 24.6. The rest of the syntactic complexity measures fell within the range between these two extremes. In the conclusion sections of the same 40 articles, the average word count was 515.65, much longer than the abstracts, as expected. Again here, dependent clauses per clause had the lowest mean (0.42) while the highest average was for the "mean length of sentence", reaching 26.59. The average values of the remaining features fell between these low and high points.

Table 4. Descriptive Statistics of Syntactic Complexity Measures for the Non-Native English Writer Over 25 Years

N	Measure of complexity	40 Abstracts		40 Conclusions	
		Mean	STD	Mean	STD
1.	W	211.9	47.12	515.65	200.3
2.	MLS	24.728	4.34	26.59298	5.06
3.	MLT	24.56257	5.12	25.15794	4.77
4.	MLC	16.61966	3.78	15.35091	2.50
5.	C/S	1.521488	0.26	1.784928	0.36
6.	VP/T	2.242748	0.47	2.527283	0.53
7.	C/T	1.502445	0.27	1.68938	0.35
8.	DC/C	0.37029	0.11	0.41835	0.10
9.	DC/T	0.580818	0.25	0.70508	0.25
10.	T/S	1.016983	0.09	1.083943	0.16
11.	CT/T	0.476548	0.20	0.513473	0.15
12.	CP/T	0.88758	0.39	0.743648	0.28
13.	CP/C	0.596605	0.27	0.45574	0.18
14.	CN/T	3.677138	0.99	3.763613	0.78
15.	CN/C	2.4859	0.70	2.296373	0.40

Note. W: number of words, MLS: mean length of sentence, MLT: mean length of T-units, MLC: mean length of clause, C/S: clause per sentence, VP/T: verb phrases per T-unit, C/T: clauses per T-unit, DC/C: dependent clauses per clause, DC/T: dependent clauses per T-unit, T/S: T-units per sentence, CT/T: complex T-units per T-unit, CP/T: Coordination phrases per T-units, CP/C: Coordination phrases per clause, CN/T: Complex nominals per T-unit, CN/C: Complex nominals per clause./ STD: standard deviation

In the comparison of descriptive data for abstracts and conclusions written by Reza, the non-native English-speaking scholar, long sentences and long T-units were observed over the years. *MLS* (Mean Length of Sentences) showed the highest standard deviation, meaning that sentence lengths vary a lot over time. Sometimes the scholar wrote shorter sentences, and other times longer ones. At the same time, he used fewer dependent clauses and coordination phrases than other grammatical elements. The measure *DC/C* demonstrated a very low standard deviation, meaning that this feature of writing stayed more consistent over the years. These findings suggest that Reza's style balances length and simplicity; his sentences may be long, but they are not very complex in structure, especially at the clause level. This may show a developmental stage in academic writing for the experienced non-native speaker.

The analysis of the native scholar's writing (Table 5) followed the same steps for analyzing the abstract and conclusion sections of articles written by Tom, using the L2SCA tool.

Table 5. Descriptive Statistics of Syntactic Complexity Measures for the Native Writer Over 25 Years

N	Measure of complexity	35 ABSTRACTS		35 CONCLUSIONS	
		Mean	STD	Mean	STD
1.	W	191.0857	99.37	632.2571	375.9
2.	MLS	31.34055	8.01	37.37894	17.61
3.	MLT	29.42304	7.43	33.5259	17.57
4.	MLC	19.22287	6.69	18.83393	13.45
5.	C/S	1.708131	0.47	2.349974	0.89
6.	VP/T	2.412791	0.77	3.096569	1.21
7.	C/T	1.605091	0.44	2.078994	0.81
8.	DC/C	0.442706	0.15	0.504746	0.19
9.	DC/T	0.761837	0.41	0.973331	0.50
10.	T/S	1.069077	0.13	1.262191	0.36
11.	CT/T	0.526743	0.25	0.621097	0.24
12.	CP/T	0.854091	0.45	0.723449	0.31
13.	CP/C	0.562197	0.32	0.388831	0.14
14.	CN/T	4.328306	1.21	3.996129	1.57
15.	CN/C	2.769417	0.69	2.171117	0.69

By mean number of words, the NNS tended to write longer abstracts than the NS (average of about 191.08 words for NS and 211.9 for the NNS). On the other hand, the NS wrote longer conclusions with an average of 632.3 words per conclusion (515.7 for the NNS). Overall, the data shows that over 25 years, Tom preferred writing longer sentences (high MLS) over the years, especially in his conclusions. At the same time, he used fewer dependent and coordinating phrases per clause (low DC/C and CP/C), and these features stayed quite consistent across his writing. This suggests that while Tom liked to write in long lengths, he avoided making the text too complex by using many complex structures.

The general observation was that, over the span of 20 to 25 years of academic writing, both Reza and Tom showed a strong preference for longer sentences, as indicated by their high Mean Length of Sentence (MLS). However, they used different individual writing habits. Reza wrote increasingly longer and more varied sentences; his sentence length changed a lot, suggesting a flexible and dynamic style. Despite using long sentences, he did not use a lot of dependent clauses and coordinating phrases. Tom also preferred long sentences, especially in conclusions. Like Reza, he used a few dependent and coordinating phrases. Although Tom seemed to use more complex structures, likely due to his comfort with the language, he avoided making the grammar too dense and kept the writing less complex, as reflected in the negative values reported for the conclusion. It is essential to note that having a high average value for a particular measure does not necessarily indicate significant progress or development in using that feature over time. In other words, a consistently high mean does not automatically equate with a high rate of improvement.

To accurately determine whether there was any significant change over time and to measure the amount of change, a series of separate linear regression analyses for each complexity measure were performed to reveal trends and change patterns that simple averages alone might not manifest. In regression equations and scatterplots, the annual rate of change was also calculated and recorded. The data presented in Table 6 reflect the changes in 14 different syntactic measures across 25 years of academic writing by Tom, focusing on two key sections of his research articles: abstracts and conclusions. These measures help indicate the complexity and development of his academic style over time.

Table 6. Changes in Syntactic Complexity Measures over Time for the Native English Writer

N	Measure of complexity	Abstracts		Conclusions	
		R Square (%)	Annual change (Times)	R Square (%)	Annual change (Times)
1.	MLS	0.043	0.25	0.0	0.045
2.	MLT	0.021	0.16	0.1	0.08
3.	MLC	0.006	0.08	0.02	- 0.09
4.	C/S	0.017	0.007	0.7	0.01
5.	VP/T	0.004	0.007	0.1	0.005
6.	C/T	0.063	0.009	1.0	0.01
7.	DC/C	1.3	- 0.009	1.2	0.003
8.	DC/T	0.2	0.003	3.1	0.013
9.	T/S	2.7	0.003	0.7	-0.004
10.	CT/T	1.4	0.004	1.6	0.004
11.	CP/T	0.0	0.001	0.0	0.0006
12.	CP/C	0.0	- 0.0001	0.9	- 0.002
13.	CN/T	0.1	- 0.006	0.0	0.004
14.	CN/C	0.8	- 0.009	1.8	- 0.01

Note: For all reported inferential measures, the p-values were less than 0.05, indicating that the findings were significant.

In the abstract sections, most measures show an upward trend. The biggest improvements were in Mean Length of Sentence (MLS), which went up by 0.25, and Mean Length of T-units (MLT), which increased by 0.16 each year. This means Tom's abstract writing became more advanced, with longer sentences and T-units (Figure 1). Mean Length of Clause (MLC) also grew by 0.08, showing more detailed clauses. Other measures, such as Clauses per Sentence (C/S), Verb Phrases per T-unit (VP/T), Clauses per T-unit (C/T), and Dependent Clauses per T-unit (DC/T), also showed minor improvements, by small amounts (0.001 to 0.009). However, a few measures declined: Coordination Phrases per Clause (CP/C), Complex Nominals per T-unit (CN/T). The largest decrease was in Complex Nominals per Clause (CN/C) and Dependent Clauses per Clause (DC/C), both down by - 0.009. This might show a small decrease in some types of clause complexity. Changes in conclusions were less noticeable, as seen in Table 6. For example, MLT in conclusions

increased by only 0.08 times per year (Figure 2). Overall, the results show that for Tom, most measures of syntactic complexity in academic writing improved over time in the abstracts. His writing became more advanced and complex, with longer sentences and T-units.

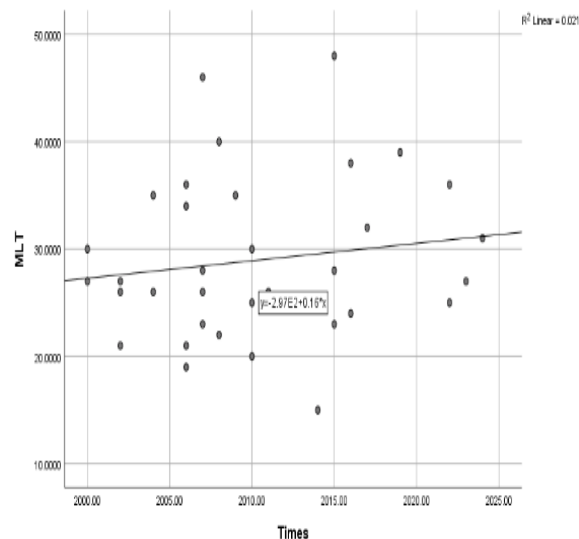


Figure 1. Regression Graph for MLT in NS Abstracts

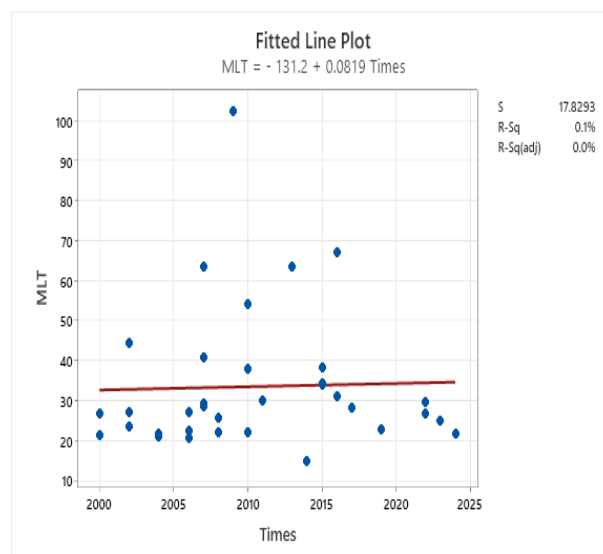


Figure 2. Regression Graph for MLT in NS Conclusions

Table 7 presents the annual change in syntactic measures in the articles produced by the non-native scholar over 20 years. (He did not have any publications qualified for inclusion in the first five years in the abstracts; all 14 measures showed some level of improvement, even if only slightly. The biggest improvement was in the Mean Length of T-units (MLT), which increased by 0.13 times per year. The smallest improvement was in the Clause per Sentence

(C/S) measure, growing by just 0.0003 per year. Notably, none of the measures showed any decline during this time in the abstracts, but many of the 14 measures in the conclusion section declined over time, as indicated by the negative values in Table 7.

Table 7. Improvement Rates of Syntactic Complexity Measures for Non-native Writers

N	Measure of complexity	Abstracts		Conclusions	
		R Square (%)	Annual change (Times)	R Square (%)	Annual change (Times)
1.	MLS	0.007	+ 0.08	0.2	- 0.04
2.	MLT	0.016	+ 0.13	0.2	- 0.04
3.	MLC	4.06	+ 0.02	7.2	+ 0.1
4.	C/S	4.7	+ 0.0003	4.5	+ 0.01
5.	VP/T	0.014	+ 0.01	3.0	- 0.019
6.	C/T	2.3	+ 0.008	4.1	- 0.01
7.	DC/C	0.7	+ 0.0019	4.9	- 0.004
8.	DC/T	1.5	+ 0.006	9.7	- 0.01
9.	T/S	0.5	+ 0.001	0.3	+ 0.001
10.	CT/T	1.0	+ 0.004	8.2	- 0.009
11.	CP/T	2.5	+ 0.01	0.1	+ 0.001
12.	CP/C	1.3	+ 0.006	2.5	+ 0.005
13.	CN/T	4.2	+ 0.04	1.9	- 0.02
14.	CN/C	1.5	+ 0.01	1.2	+ 0.009

Note: For all reported inferential measures, the p-values were less than 0.05, indicating that the findings were significant.

Overall, the results suggest that in the abstract section of his research articles, Reza gradually increased complexity over time. Notably, he increased the length of T-units by 0.13 times per year (Figure 3). In Reza's conclusion sections, the length of T-units decreased annually by 0.04 times (Figure 4). At the same time, there was a decline in his use of long sentences, verb phrases, clauses, dependent clauses, and complex structures within T-units, showing a shift in how he built and balanced his sentences over the years.

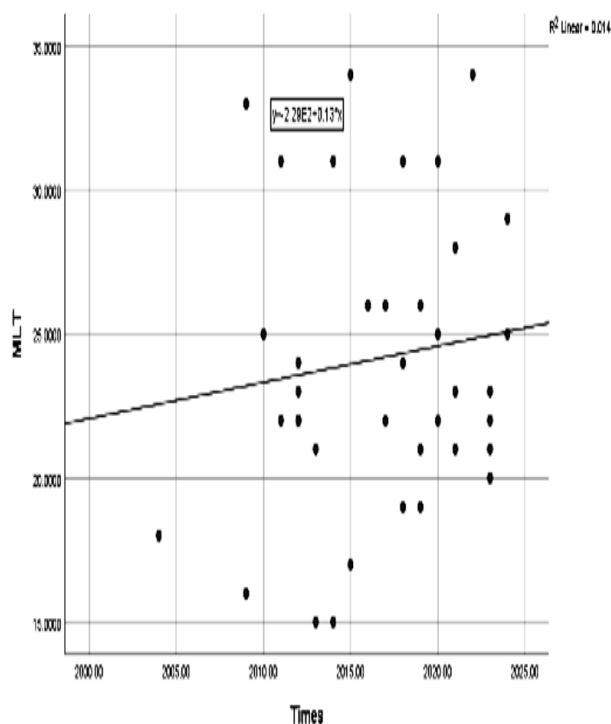


Figure 3. Regression Graph for MLT in NNS Abstracts

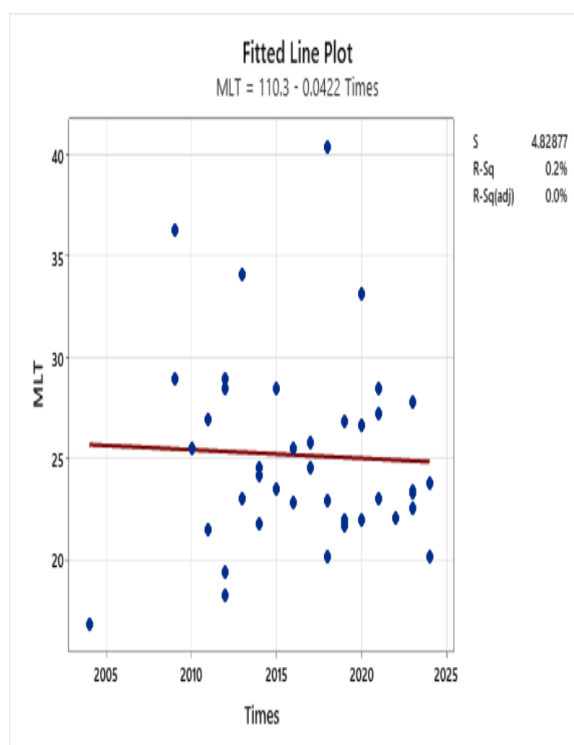


Figure 4. Regression Graph for MLT in NNS Conclusions

As shown in Table 8, Tom showed more prominent improvement in the abstracts. His writing became more mature and complex, especially through the use of longer sentences and

T-units with fewer complex nominals. In Tom's conclusions, many of the complexity measures declined over time. Overall, Reza's development was more pronounced in length-level elements and other complexity features in abstracts. Reza also showed some clause-level complexity reduction in conclusions. Tom's writing, on the other hand, showed greater gains for length-level elements in abstracts and little clause-level reduction in conclusions.

Table 8. Comparison of Improvement Rates of Syntactic Complexity Measures for the Native and Non-native Writers

N	Measure of complexity	Abstract		Conclusion	
		Native speaker	Non-Native speaker	Native speaker	Non-Native speaker
1.	MLS	0.25	+ 0.08	0.04	- 0.04
2.	MLT	0.16	+ 0.13	0.08	- 0.04
3.	MLC	0.08	+ 0.02	- 0.09	+ 0.1
4.	C/S	0.007	+ 0.0003	0.01	+ 0.01
5.	VP/T	0.007	+ 0.01	0.005	- 0.01
6.	C/T	0.009	+ 0.008	0.01	- 0.01
7.	DC/C	- 0.009	+ 0.001	0.003	- 0.004
8.	DC/T	0.003	+ 0.006	0.01	- 0.01
9.	T/S	0.003	+ 0.001	0.004	+ 0.001
10.	CT/T	0.004	+ 0.004	0.004	- 0.009
11.	CP/T	0.001	+ 0.01	0.0006	+ 0.001
12.	CP/C	- 0.0001	+ 0.006	- 0.002	+ 0.005
13.	CN/T	- 0.006	+ 0.04	0.004	- 0.02
14.	CN/C	- 0.009	+ 0.01	- 0.01	+ 0.009

Note: For all reported inferential measures, the p-values were less than 0.05, indicating that the findings were significant.

Discussion

While the researchers express caution in interpreting the results, they believe that the real textual data and the computerized measurements that we used to explore syntactic complexity offer some insights about syntactic complexity development in writing in English for

research publication purposes. Our quantitative results indicate that, over time, the NS and NNS scholars' abstracts improved in most syntactic complexity measures calculated by the L2SCA program. Associations with time revealed that the NNS scholar progressively wrote abstracts with higher complexity measures, but the NS scholar improved only in some targeted measures in his abstracts. The NS scholar's abstracts demonstrated significant changes in Mean Length of Sentence (MLS) and Mean Length of T-unit (MLT), with smaller improvements in other measures. The NNS scholar's abstracts showed improvements over time in all 14 measures, with the largest increase in MLT and the smallest in C/S. These results provide indirect evidence that supports skill acquisition theory. The theory claims that acquisition of L2 skills (including ERPP writing skills) follows declarative, procedural, and automatic stages of development ([Suzuki, 2022](#)). Our cases, who were learned scholars, acquired declarative knowledge of writing for scholarly publication but continued during their careers to develop many features of their writing.

Annual decreases were also observed in syntactic measures in conclusions written noticeably by the NS writer, indicating that syntactic maturity in academic writing does not always imply an increase in these complexity measures. The passage of time and increased academic writing experience, as indicated by a cumulative number of published articles, were related to syntactic developmental paths observed in our study; however, the researchers acknowledge that the authors may have modified their writing in response to changing disciplinary practices, journal expectations, co-authorship effects, reviewer expectations/comments, or other related factors. The study showed that syntactic complexity changes over time at least in some of its identifying features. A deeper analysis of this issue requires careful qualitative analysis of the texts, a limitation of the current study. Our findings partially align with [Biber et al.'s \(2011\)](#) observation that academic writing evolves from clausal elaboration (e.g., longer T-units and clauses) toward greater phrasal complexity (e.g., more complex nominal structures) as writers gain experience. Our NNS scholars' consistent improvement in abstracts shows that the development of syntactic sophistication in academic abstract writing is an important feature to consider in ERPP training. The finding that our cases simplified structures in conclusions may indicate the use of a textual strategy to increase clarity for readers. Using less complex clause embedding is a strategy that is used by writers to boost clarity ([Hyland, 2006](#)).

[Menke and Strawbridge's \(2019\)](#) findings align with the results of the study in highlighting the role of individual differences in L2 writing development. Just as the Spanish learners showed unique growth patterns across measures of complexity, in our study, Reza

and Tom followed different developmental trajectories: Reza steadily building syntactic complexity as an EFL writer, and Tom refining strong structures as a native speaker. These studies emphasize learners' starting points and individual learning paths as shapers of how syntactic skills may evolve over time. Interestingly, both scholars also made certain choices to simplify writing. The observation that not all indicators of syntactic complexity improved over time indicates that a syntactically more complex text is not necessarily an academically more appropriate one in ERPP. Syntactic sophistication should be at the service of the successful creation and communication of academic wisdom in a learned community. [Polat et al. \(2020\)](#) found that Turkish EFL learners significantly increased the syntactic complexity of their texts over time. Pre-intermediate learners used simpler structures, but intermediate learners employed more varied constructions. This is corroborated by our findings: The NNS scholar wrote more complex sentences over time. The findings also align with [Lei et al. \(2023\)](#), who showed that syntactic complexity develops over time in the English writing of 1,081 intermediate Chinese university students across four semesters. The findings also strengthen the claims by [Lu and Ai \(2015\)](#) that L1 background can affect the path of syntactic development.

Conclusion

The major finding of the study was that measures of syntactic complexity do change and that patterns of association with time are different for the NS and NNS scholars. The study's findings carry important implications for ERPP pedagogy and EFL academic writing instruction, because they show that over time, it is not only writing process knowledge and/or disciplinary genre knowledge that changes, but also syntactic knowledge and its use. A decline in some complexity measures of conclusions written by the NNS and of both the abstracts and conclusions written by the NS can imply that syntactic knowledge in ERPP does not improve linearly over time.

This study also suffered from some limitations. Even though the number of texts was notable, only two scholars were studied, decreasing the generalizability of the findings. Moreover, we employed [Lu's \(2010\)](#) syntactic complexity measures that suffer from the limitation of decontextualized quantitative analyses, even though they can be used as tools to monitor progress in consecutive writing samples in ERPP instruction. In a published academic text, unique conditions of text production in ERPP, including disciplinary conventions, journal and editorial expectations, co-authorship influences, and revisions in the publication process, are possible influential factors that were ignored in our study. Future

studies on syntactic development and its role in mastering ERPP need to examine larger corpora, consider possible cross-disciplinary variations, and include more variables in regression equations to show how much of the variation in writing for publication is syntactic. Further studies can also qualitatively analyze academic texts to track progress in writing proficiency and suggest what novice writers can learn from advanced scholars. The analysis was confined only to abstracts and conclusions, potentially missing patterns in other sections of research articles. The study also looked only at the development of syntactic complexity, ignoring fluency, accuracy, and lexical variations. Future research can be designed to address these limitations.

Our comparison of the part-genres of abstract and conclusion for the NNS scholars implies that the structural aspects of abstract writing may have received more attention than those of the conclusion in writing ERPP training. Changes in the published works of our established NNS scholar and NS scholar over time imply that novice researchers and advanced EFL learners need help in writing syntactically more complex and sophisticated abstracts as part of training in writing for publication. Scholars also need to learn how to refine phrasal complexity to maintain sophistication and prevent oversimplification, particularly in writing conclusions. The study underscores the value of differentiated ERPP training programs for NS and NNS novice scholars with different publication experiences, enabling native speakers to polish stylistic nuance and EFL writers to overcome structural challenges. Workshops on writing for publication need to consider these differences in their design and implementation. For many learners of ERPP, practicing the overall structure and genre moves and functions of scholarly publication may not be enough, and mastery of advanced and sophisticated sentence-level syntax may be required.

Acknowledgements

No funding was received for this research. We are very grateful to the anonymous reviewers of the journal who significantly contributed to the improvement of this work.

References

- 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>
- Bitchener, J. (2018). *A guide to supervising non-native English writers of theses and dissertations*. Routledge.

- Brown, H. D., & Lee, H. (2025). *Principles of language learning and teaching: A course in second language acquisition*. Routledge.
- Canagarajah, A. S. (1996). "Nondiscursive" requirements in academic publishing, material resources of periphery scholars, and the politics of knowledge production. *Written Communication*, 13(4), 435–472. <https://doi.org/10.1177/0741088396013004001>
- Cargill, M., & Burgess, S. (2008). Introduction to the special issue: English for research publication purposes. *Journal of English for Academic Purposes*, 7(2), 75–76. <https://doi.org/10.1016/j.jeap.2008.02.006>
- Curry, M. J., & Lillis, T. (2004). Multilingual scholars and the imperative to publish in English: Negotiating interests, demands, and rewards. *TESOL Quarterly*, 38(4), 663–688. <https://doi.org/10.2307/3588284>
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.
- Fazel, I. (2018). Writing for publication as a native speaker: The experiences of two Anglophone novice scholars. In P. Habibie & K. Hyland (Eds.), *Novice writers and scholarly publication: Authors, mentors, gatekeepers* (pp. 79–95). Springer International Publishing.
- Flowerdew, J. (2013). English for research publication purposes. In B. Paltridge & S. Starfield (Eds.), *The handbook of English for specific purposes* (pp. 301–321). Wiley/Blackwell. <https://doi.org/10.1002/9781118339855.ch16>
- Flowerdew, L. (2015). Using corpus-based research and online academic corpora to inform writing of the discussion section of a thesis. *Journal of English for Academic Purposes*, 20, 58–68.
- Flowerdew, J., & Habibie, P. (2021). *Introducing English for research publication purposes*. Routledge.
- Flowerdew, J., & Li, Y. (2020). Teaching English for research publication purposes (ERPP): A review of language teachers' pedagogical initiatives. *English for Specific Purposes*, 59, 29–41. <https://doi.org/10.1016/j.esp.2020.03.002>
- Habibie, P. (2016). Writing for scholarly publication in a Canadian higher education context: A case study. In C. M. Badenhorst & C. Guerin (Eds.), *Research literacies and writing pedagogies for masters and doctoral writers* (pp. 51–67). Brill. https://doi.org/10.1163/9789004304338_004

- Habibie, P. (2022). Early-career scholars and scholarship: A social justice perspective. *Annual Review of Applied Linguistics*, 42, 55-63. <https://doi.org/10.1017/S0267190521000192>
- Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Routledge.
- Hyland, K. (2015). *Academic publishing: Issues and challenges in the construction of knowledge*. Oxford University Press.
- Hyland, K., & Shaw, P. (2016). Introduction. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 1–13). OUP.
- Istiharoh, A. F., & Pranowo, D. D. (2023). The development of syntactic complexity in university students' essays. *Englisia: Journal of Language, Education, and Humanities*, 11(1), 1-18. <https://doi.org/10.22373/ej.v11i1.19129>
- Lei, L., Wen, J., & Yang, X. (2023). A large-scale longitudinal study of syntactic complexity development in EFL writing: A mixed-effects model approach. *Journal of Second Language Writing*, 59, Article 100962. <https://doi.org/10.1016/j.jslw.2022.100962>
- 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>
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45(1), 36–62. <https://doi.org/10.5054/tq.2011.240859>
- 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>
- 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>
- Menke, M. R., & Strawbridge, T. (2019). The writing of Spanish majors: A longitudinal analysis of syntactic complexity. *Journal of Spanish Language Teaching*, 6(2), 100–117. <https://doi.org/10.1080/23311983.2019.1679955>
- Mur-Dueñas, P. (2019). The experience of a NNEs outer circle novice scholar in scholarly publication. In P. Habibie & K. Hyland (Eds.), *Novice writers and scholarly publication. Authors, mentors, gatekeepers* (pp. 97–115). Palgrave Macmillan.

- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492–518. <https://doi.org/10.1093/applin/24.4.492>
- Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of Second Language Writing*, 29, 82–94. <https://doi.org/10.1016/j.jslw.2015.07.008>
- Polat, N., Mahalingappa, L., & Mancilla, R. L. (2020). Longitudinal growth trajectories of written syntactic complexity: The case of Turkish learners in an intensive English program. *Applied Linguistics*, 41(5), 688–711. <https://doi.org/10.1093/applin/amz034>
- Suzuki, Y. (2022). Automatization and practice. In A. Godfroid & H. Hopp (Eds.), *The Routledge handbook of second language acquisition and psycholinguistics* (pp. 308–321). Routledge.
- Swales, J. (2004). *Research genres: Explorations and applications*. Cambridge University Press.
- Vyatkina, N., Hirschmann, H., & Golcher, F. (2015). Syntactic modification at early stages of L2 German writing development: A longitudinal learner corpus study. *Journal of Second Language Writing*, 71, Article 101282. <https://doi.org/10.1016/j.jslw.2015.07.001>
- Wang, Z., & Lu, X. (2026). Development of syntactic complexity in Chinese EFL learners across two proficiency levels and two written genres. *Journal of Second Language Writing*, 71, 101282. <https://doi.org/10.1016/j.jslw.2026.101282>
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. University of Hawai‘i Press.
- Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Sage.
- 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>

