

Syntactic Complexity in Novice and Expert L2 Academic Writing: The Role of Syntactic Elaboration and Diversity

Rajab Esfandiari*¹

Mohammad Ahmadi²

Received: 01/03/2023

Accepted: 11/12/2023

Abstract

Syntactic complexity (SC) is an important construct for gauging L2 writing proficiency. Previous studies, including Biber et al. (2016) and Dong et al. (2023), have largely focused on syntactic elaboration and disregarded syntactic diversity. This study investigates how academic writing proficiency is associated with SC, through an evaluation of both the variety and elaboration of sentence structures. To that end, a corpus of the abstract sections of MA theses, PhD dissertations, and empirical research articles in applied linguistics produced by Iranian English medium academic writers was compiled. Drawing on automated text analysis tools (Stanford Core NLP and L2 SC analyzer), we measured syntactic elaboration and diversity of the texts written by three distinct sets of academic authors. The results indicated that phrasal and global measures of syntactic elaboration significantly predicted academic writing proficiency in both novice and expert writers. The results of the second phase of the study indicated that the incorporation of syntactic diversity measure considerably increased the predictive power of the model. The pedagogical implications of the findings for academic writing instruction through consciousness-raising activities for both novice and expert writers are discussed.

Keywords: academic writing, expert and novice writing, syntactic diversity, syntactic elaboration, syntactic complexity

* Corresponding author

¹ Department of English Language Teaching, Faculty of Literature and Humanities, Imam Khomeini International University, Qazvin, Iran; esfandiari@hum.ikiu.ac.ir

² Department of English Language, Faculty of Humanities, Lorestan University, Khorramabad, Iran; Ahmadi.m@lu.ac.ir

[DOI: 10.22051/lghor.2023.43081.1780](https://doi.org/10.22051/lghor.2023.43081.1780)

Introduction

Indices related to the complexity of sentence structures have demonstrated their crucial utility for research purposes in numerous domains connected to language, including the acquisition of language by children, (e.g., Lu, 2009), language impairment (e.g., Leonard et al., 2007), language and aging (e.g., Kemper et al., 2003), and second language acquisition (e.g., Kyle, et al., 2021). More specifically, in second language acquisition contexts, syntactic complexity (SC) has been used to explore and categorize learners' proficiency levels (e.g., Li et al., 2023). The main rationale behind using SC for measuring L2 learners' writing proficiency is that, as L2 learners progress through more advanced levels of language production, their SC increases (Larsen-Freeman, 2009). Although writing skill cannot be solely measured by SC (Lu, 2011), it is recognized as one of the most significant factors, since it is one of the most prominent ones, and measures like embedding and subordination encourage the generation of complex ideas (Beers & Nagy, 2009).

Over the past decades, SC in academic writing has featured prominently in a number of important research strands. Most of the studies in this field have mainly concentrated on the progress of writing skills (e.g., Crossley & McNamara, 2010; Parkinson & Musgrave, 2014), disciplinary variations (e.g., Staples et al., 2016), L1 backgrounds (e.g., Ruan, 2018), and genre variations (e.g., Staples & Reppen, 2016). Although these studies have furthered our understanding of the role of SC in academic writing for measuring performance and proficiency, most measurement practices in academic writing research have solely focused on syntactic elaboration. However, this approach has overlooked the importance of syntactic diversity, which is an equally important dimension of SC (See Bulté & Housen, 2012). Accordingly, without syntactic diversity, we would have a “hantchy SC picture of L2 writing” (Bi & Jiang, 2020, p. 3). While some recent studies (e.g., Casal et al., 2021) have incorporated more measures into their investigations of the construct of SC, the potential influence of other relevant domains such as syntactic diversity in determining SC and SC development is still a desideratum (De Clercq & Housen, 2017). This is probably because the construct is poorly defined and used with different meanings (in a way that its multidimensionality is not fully recognized or reflected terminologically).

Despite the growing interest in SC in academic writing research, very few studies (e.g., Ansarifar, et al., 2018; Karami & Salahshoor, 2014; Mohamadi Zenouzagh, 2020; Valizadeh & Soltanpour, 2021) have investigated the relationship between SC and writing proficiency in Iran. However, these studies have primarily focused on measuring SC using a few selected features and have not considered the multidimensionality of SC, particularly syntactic diversity. Hence, the goal of this study is to address this issue by exploring the relationship between SC, encompassing both elaboration and variety, and academic writing proficiency among Iranian English medium academic writers. Through this effort, we aim to promote knowledge and comprehension of SC in the context of Iranian EFL writers and highlight the importance of considering syntactic diversity in measuring writing proficiency.

Literature Review

Measures of SC in Previous Literature

Different research studies have attempted to quantify the construct of SC by means of reliable metrics which could capture development, performance, and proficiency in various academic contexts. Kyle and Crossley (2018), for example, argued that, traditionally, SC has been characterized using broad measurements that gauge complexity either at the clause or sentence level (e.g., mean length of the clause/sentence). Bulté and Housen (2014) claimed that most research studies have employed measures of complexity with the underlying belief that language becomes more complex as it includes a larger number of linguistic elements. Generally, the measures commonly used for measuring SC are classified into five main groups, which are outlined in Table 1.

Table 1

Lu's (2011) Classification of SC Measures

Measure

Type 1: Length of production

Mean length of clause (MLC)

Mean length of sentence (MLS)

Mean length of T-unit (MLTU)

Type 2: Sentence complexity

Clauses per sentence (C/S)

Type 3: Subordination

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)

Type 4: Coordination

Coordinate clauses per clause (CP/C)

Coordinate clauses per T-unit (CP/T)

T-units per sentence (T/S)

Type 5: Particular structures

Complex nominals per clause (CN/C)

Complex nominals per T-unit (CN/T)

Verb phrases per T-unit (VP/T)"

Note. Adapted from “A corpus-based evaluation of SC measures as indices of college-level ESL writers' language development” by X. Lu, 2011, *TESOL Quarterly*, 45(1), 42.

A large proportion of studies in previous literature on SC used one or two of the measures listed in Table 1 targeting complexity at clausal or sentential level and, to a lesser extent, at phrasal level. In the words of De Clercq and Housen (2017), this line of research is called a “reductionist approach” to investigating L2 complexity, as it provides singular focus on SC. While some recent studies have attempted to address these criticisms (e.g., Kyle & Crossly, 2018; Naserpour et al., 2020; Yoon & Polio, 2017), little is known about the incorporation of syntactic diversity measures. Present research is, accordingly, concerned with the examination of SC in advanced academic writing through the lenses of both syntactic elaboration and syntactic diversity. If SC is defined as the degree of variety, sophistication, and elaboration of grammatical forms used in production (Norris & Ortega, 2009), the study of syntactic elaboration alone is, by no means, an accurate indication of syntactic growth and development. It is now widely recognized that syntactic elaboration on its own should not automatically be taken as a sign of increased proficiency. Thus, syntactic diversity measures (the topic we deal with below) “may further refine the descriptive toolset used in complexity research by considering the variation of syntactic structures in relation to proficiency” (De Clercq & Housen 2017, p. 4).

Syntactic Diversity Measures

Since previous studies largely focused on syntactic elaboration and overlooked syntactic diversity, it remains to be seen which syntactic diversity metrics can be reliably used across proficiency levels in different research contexts. L2 writing researchers have employed a few syntactic diversity metrics such as syntactic diversity index (SDI) (Bi & Jiang, 2020), verb phrase construction (Verspoor et al., 2012), and so forth in previous literature. However, these measures need to be used with caution as they may not always be a reliable index of diversity in different research contexts. For example, Bi and Jiang (2020) noted that SDI can only capture syntactic diversity of particular constructions, such as clause types and ignore other important language features.

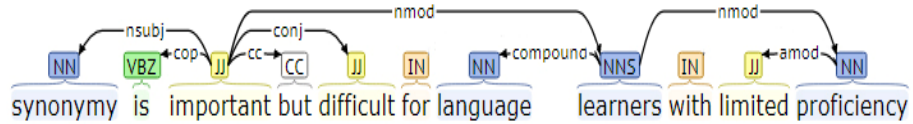
One of the most recent measures of syntactic diversity proposed by Bi and Jiang (2020) is concerned with dependency relations. The term “dependency grammar” is not used to describe a specific theory of grammar; rather, it is a way of describing grammatical relations between two grammatical units. In contrast to constituency parsing, which concerns itself with groups of words, dependency parsing is concerned with the binary relation, that is, “a grammatical relation that holds between a governor and a dependent” (De Marneffe & Manning, 2008, p. 2). Dependency parsing is based on the assumption that words have grammatical functions in relation to other words in a sentence (De Marneffe & Nivre, 2019). Dependency parsing is sometimes preferable to constituency parsing because it makes “some syntactic analyses (e.g., identifying arguments of a verbs) much more convenient than constituency parsed representations” (Kyle, 2016, p. 42). As Chen and Manning (2014) put it, “The accuracy of POS tags is reported to be around 97% with training data” (p. 745). Figure 2 shows how dependency parsing works on a sample taken from the corpus of the present study.

Dependency parsed structures are used for analyzing grammatical relations among words or phrases with respect to their syntactic functions (e.g., direct object). The relationships are analyzed in terms of head words and the words that modify those heads. Unlike constituency parsing which is based on the hierarchical investigation of phrase structures, dependency parsing is concerned with governors and dependents, which makes it much more convenient for such analyses as identifying the arguments of verbs (Kyle, 2016). Figure 1 displays dependency

parsing of a sentence borrowed from the texts of the present research.

Figure 1

Visual Depiction of Dependency Relations



SC and Academic Writing Proficiency

One of the important functions of the construct of SC and its proposed measures in the previous studies within the genre of academic writing was characterizing development and discriminating proficiency differences. A considerable body of investigations has centered on how SC is connected to the quality of writing as operationalized by the scores assigned by human raters (e.g., Casal & Lee, 2019). Nevertheless, the challenge with this line of research is that evaluating the texts by means of manual analysis is a time-consuming method for comparing samples within an existing corpus, particularly when dealing with large corpora, and demands a considerably high level of proficiency on the part of coders. Another method is to contrast the written material created by novice and expert writers. Novice texts refer to “unpublished pieces of writing that have been written in educational or training settings” (Scott & Tribble, 2006, p. 133), while expert texts refer to “pieces of writing that has been published” (Römer & Arbor, 2009, p. 148). Thus, it is unnecessary to manually scrutinize the texts to assess the writers' skill levels.

Some prior research studies have centered on examining sentence complexity in the writing of novice and/or expert authors. Ansarifar et al. (2018), for example, examined phrasal complexity in academic writing. A dataset was built that included the abstract sections of Master's theses, PhD dissertations, and empirical research articles (RAs) in the field of applied linguistics. Their contention was that the student writers in both MA and PhD groups were considered to possess equivalent levels of proficiency since they had passed the same entrance examination and received education based on the same official curriculum. However, they concluded that PhD level students are more experienced in academic

writing because (1) they are chosen based on a strict selection procedure which consists of a nationwide assessment test for admission and an appraisal of their research experience and (2) they complete four extra years of intensive courses compared to MA students. Parkinson and Musgrave (2014) investigated the complexity of phrases in texts written by students of English for academic purposes, in a manner that is comparable to the present study. They compiled a corpus of texts created by students who were preparing for graduate studies (lower-proficiency group) and those who had already enrolled in graduate programs (higher-proficiency group). The findings confirmed the writing development phases suggested by Biber et al. (2011), according to which less proficient writers tend to rely heavily on features like attributive adjectives that are typical of early writing stages. In addition, the higher-proficiency group used modifying features at a frequency that was similar to that found in published academic writing.

While there are some studies conducted in the Iranian context with the aim of analyzing how writing skill is associated with the complexity of sentence structures (e.g., Azadnia et al., 2019; Esfandiari & Ahmadi, 2021; Shadloo et al., 2017; Yazdani, 2018), the investigation of syntactic diversity is still an under-researched area. Azadnia et al. (2019), for example, conducted a study that examined the resemblances and variations between written works generated by college-level English students and those created by individuals who are native speakers of English. They discovered that two out of the four metrics, i.e., the mean number of modifiers and sentence syntax similarity, were identified as distinguishing factors between texts written in a person's first language (L1) and those written in their second language (L2). However, "left embeddedness and minimal edit distance" (p. 235) were found to be similar in both sets of texts. Moreover, Shadloo et al. (2017) examined the predictive power of SC measures in the argumentative writings of the learners from different proficiency levels. The findings indicate that phrasal features cannot be used as reliable measures for predicting SC development in the genre of argumentative essays. As a result, although there is an increasing amount of literature concerning SC in writing, incorporating studies from Iran underscores the significance of examining this topic within specific language contexts.

The current state of research on SC in L2 writing proficiency among

Iranian learners highlights the need for further investigation into the role of both comprehensive measures of SC and syntactic diversity, which has yet to be explored in this specific context. Esfandiari and Ahmadi (2021), for example, conducted a corpus-based study to explore the correlation between metrics of sentence complexity and academic writing proficiency among professional and student writers. Their results showed that phrasal measures of SC were significantly associated with writing proficiency, particularly for the student group. In another study, Yazdani (2018) investigated the SC of Iranian learners' written and spoken English. The study revealed that the learners' written texts exhibited higher levels of SC than their spoken language. Building on previous research, our study aims to investigate the contributions of both syntactic diversity and syntactic elaboration to L2 writing proficiency among Iranian learners.

In order to conduct a more multidimensional investigation of SC in academic writing, the present study intended to examine SC in academic register by focusing on two important dimensions of the construct simultaneously (i.e., syntactic elaboration and syntactic diversity). Accordingly, we specifically focused on two academic contexts of student-written and professional prose to represent novice and expert academic writing. We, therefore, sought to answer the following two research questions.

1. Do syntactic elaboration measures predict the writing proficiency of expert and novice academic writers?
2. Does the incorporation of a diversity measure enhance the predictive power of SC measures in determining the writing abilities of novice and expert academic writers?

Method

Construction of the Corpus

In order to conduct quantitative analyses in the present study, we compiled two corpora of expert academic writing and novice academic writing. The former consisted of the abstract sections of RAs in applied linguistics published in Iranian journals (Table 2). The latter encompassed abstract sections extracted at random from theses and dissertations available in databases of the most prominent state-run universities in Iran that offer Master and PhD degrees in applied linguistics. All the

texts were published between 2019 and 2021. We chose recently-published texts because they reflect current trends in academic language use. To eliminate extraneous variables that are not related to the concept being studied, such as rhetorical organization of the texts, we only chose empirical research studies and discarded other types.

Table 2

Overview of Iranian Journals Included in the Corpus

Journal	Years of Publication
Iranian Journal of Applied Language Studies	2009-ongoing
Journal of Teaching Language Skills	2009-ongoing
Journal of English Language Teaching and Learning	2010-ongoing
Journal of Language and Translation	2010-ongoing
Journal of Research in Applied Linguistics	2010-ongoing
Issues in Language Teaching	2012-ongoing
Applied Research on English Language	2012-ongoing
Iranian Journal of Language Teaching Research	2013-ongoing
Iranian Journal of English for Academic Purposes	2015-ongoing

The rationale behind selecting abstracts was twofold: primarily, their length is manageable, thereby allowing for a more extensive examination of a greater quantity of texts for cross comparisons than other parts (e.g., Introduction, Methods, Results, and Discussion). Second, they are the initial sections that editors and journal gatekeepers face, and they may decide whether or not to keep reading the rest of the manuscripts based on the abstracts. Thus, RA abstracts, as a major part-genre, play an important role in academic writing as they “foreground important claims, minimize methodology and background statements, and pack information into visuals” (Hyland, 2000, p. 86). The selection of texts followed a stratified random sampling approach, which involved dividing the texts into groups based on different attributes of the texts.

Table 3*Details Providing a Description of the Three Datasets*

Datasets	Quantity of texts	Average length of abstracts	Total number of words
MA	240	238.84	57321
PhD	144	398.94	57448
EW	327	175.57	57412

SC Metrics

To measure SC of the written materials generated by expert and novice academic writers in the current study, we employed the indices introduced by Lu (2011), who categorized them based on the functions they served (See Table 1). To choose the measures of SC, we set the criteria of multidimensionality and distinctness (Norris & Ortega, 2009). In essence, each metric was required to encompass a specific aspect of SC without duplicating or overlapping with others. Thus, considering Lu's (2011) categorization of SC indices which included five different kinds of complexity metrics, we selected one measurement index from the second, third and fifth categories, and two indices from the first category. We chose two measures from the first category because, as Norris and Ortega (2009) noted, despite their superficial similarity, MLC does not capture complexity at the clausal level just the same as MLTU; rather, MLC gauges the complexity of phrases "as it can only increase through the addition of pre- or post-modification within a phrase or through nominalization" (p. 238).

Automated Text Analysis Tools

SC analyses were carried out by means of Syntactic Complexity Analyzer (SCA), which is an accessible free of charge web-based program, designed to automate 14 SC measures (See Lu, 2010). The reason why we employed SCA was that the tool allows for multidimensional investigation of SC (each measure is said to tap into the complexification of one particular dimension) and batch processing. In contrast, some computational tools, such as Biber Tagger, are not freely available. Additionally, SCA has a high reliability index which is between 0.83 and 1.00 in terms of its agreement with human annotators (Lu & Ai, 2015). With regard to text segmentation, SCA draws on Stanford parser as the first stage, which is followed by tokenization.

Following Bi and Jiang (2020), in order to calculate syntactic diversity of the texts, the count of grammatical relations in the text (tokens) and the total types of grammatical relations (types) were considered. Since TTR is sensitive to text length (Crossley & McNamara, 2010), we corrected the formula in this study in order to avoid this problem. To this end, we drew on a new measure called MSTTR-50 (mean 50-segmental TTR of dependency relations) (See Bi & Jiang, 2020), where each text was divided into different parts.

$$MSTTR-50 = \frac{1}{n} \sum_{i=1}^n TTR_i$$

The reason why 50 was chosen was that “there were approximately 50 dependency relation types in English according to the Stanford typed dependencies manual” (p. 4). As a result, the value of syntactic diversity was assessed through the aforementioned “equation where n equals the number of segments with 50 dependency relations” (p. 4). It is worth noting that the texts were automatically annotated through Stanford Core NLP Version 3.9.2., a useful tool for natural language processing in Java and Python environments. According to Miestamo et al. (2021), dependency relations are directed links between pairs of words in a sentence, where the link represents a syntactic association between the head and modifier. An example of a dependency relation would be the relation between the word "book" and its modifier "the" in the sentence "The book is on the table." In this case, "the" is a dependent of "book" and the dependency relation between them would be labeled as a determiner-noun relation.

Statistical Analyses

To analyze the relationship between SC of the texts and academic writing proficiency, we used ordinal logistic regression (OLR) by employing five measures of SC as independent variables to predict the academic writing proficiency (dependent variable). This was operationalized on academic writing texts written by MA and PhD students, and on published texts of expert writers. OLR is a type of regression analysis used to predict the ordinal level data such as proficiency levels where the predictor variable can be either continuous or categorical (Perry, 2020).

There are different ways for performing ordinal logistic regression using different statistical packages. Probably the most convenient statistical package for

performing OLR is SPSS, where the analysis can be carried out through three different routes (i.e., Ordinal Regression, Multinomial Logistic Regression, and Generalized Linear Model). The major problem with the first route is that there is not a clear consensus in the literature about how to interpret the information such as pseudo-R square (Pituch & Stevens, 2015). In addition, it does not produce information on the Odd Ratio (OR), which is the probability of a case falling at the next higher level. The second route is through MLR, which is rarely recommended for ordered data. The third route is through GLM, which provides us with more detailed, precise depiction of the associations between the independent variables and the dependent variable (Zheng & Agresti, 2000) (i.e., academic writers' proficiency levels in the present study). SPSS and other statistical software programs utilize a model known as proportional odds (PO). In sum, GLM offers "the advantage of using the original scale, being numerically simple to interpret regardless of the choice of probability distribution for the GLM, and it is thus comparable in numerical value across GLMs with deferent links and choices of probability distribution" (Zheng & Agresti, 2000, p. 1780). Accordingly, we employed GLM to find the relationship between the predictor variables and outcome variables in the present study.

Results

Research Question 1

Table 4 summarizes the descriptive statistics of five SC measures used in the current study. Expert, PhD and MA texts had the mean length of 30.09, 28.95, and 26.45 words per sentence, 28.74, 26.71, and 22.14 words per T-unit, and 18.78, 15.81, and 14.02 words per clause, respectively. In addition, there were 1.81, 1.60, and 1.65 clauses per sentence, 0.59, 0.46, and 0.51 dependent clauses per T-unit, and 4.18, 3.51, and 2.73 complex nominals per T-unit in expert, PhD, and MA texts, respectively.

Table 4*Descriptive Statistics of Complexity Measures Across Proficiency Levels*

Measure	Group	Number of texts	Mean	Std. Deviation
Mean length of T-unit (MLTU)	EW	240	28.74	4.38
	PhD	144	26.71	4.23
	MA	327	22.14	5.65
Mean length of clause (MLC)	EW	240	18.78	4.39
	PhD	144	15.81	5.12
	MA	327	14.02	3.78
Clauses per sentences (C/S)	EW	240	1.81	0.34
	PhD	144	1.60	0.23
	MA	327	1.65	0.26
Dependent clauses per T-unit (DC/T)	EW	240	0.59	0.37
	PhD	144	0.56	0.31
	MA	327	0.51	0.28
Complex nominals per T-unit (CN/T)	EW	240	4.18	1.13
	PhD	144	3.51	0.99
	MA	327	2.73	1.12

Preliminary analyses were conducted to confirm that there were no violations of the assumptions. The variables and the values were exported to SPSS (Version 25) for statistical analyses. Table 5 displays the findings, which indicate that the Likelihood Ratio Chi-Square test revealed a significant improvement in fit of the full model compared to the null model (which had no predictors), with a $\chi^2(5)$ value of 44.112 and $p < 0.003$. This indicates that using the predictors in the present study to predict the outcome variable yields significantly more precise predictions than a frequency-based prediction that does not consider predictor variables.

Table 5*Omnibus Test of SC Measures*

Likelihood Ratio Chi-Square	df	Sig.
44.11	5.00	0.00

Table 6 shows that, out of the five measures of SC analyzed, three of them (MLTU, MLC, CN/T) had a statistically significant impact on the model. The EXP(B) values demonstrated that MLC had the strongest influence on academic writing proficiency levels, with an odds ratio of 12.13. This indicates that, for every one-unit increase in MLC, the chances of being in a higher category of writing proficiency increase by a factor of 12.13. The second most influential factor was CN/T, with an EXP(B) value of 3.41, followed by MLTU with an EXP(B) value of 2.37.

Table 6*Tests of Model Effects and Parameter Estimates of Phrasal Complexity Measures*

Predictor variables	Type III				
	B	Exp(B)	Wald Chi-Square	df	Sig.
MLC	3.23	12.13	13.98	1	0.00
CN/T	1.12	3.41	6.04	1	0.00
MLTU	0.91	2.37	3.18	1	0.01

Drawing upon the obtained data from Table 6, we can conclude that the syntactic elaboration measures employed in the present research, particularly MLC, CN/T, and MLTU, can predict the writing proficiency of expert and novice academic writers to a significant extent. The exact extent of their predictive power varies for each group, but MLC consistently emerges as the strongest predictor for all groups.

Research Question 2

For the second research question, five syntactic elaboration and one syntactic diversity measures were entered into the combined model including MLTU, MLC, CN/T, DC/T, C/S, and MSTTR-50. Table 7 shows that our combined

model was a remarkable improvement in fit over syntactic elaboration model (no diversity measures included) [$\chi^2(5) = 53.214, p < 0.001$]. The dependent variable, therefore, in terms of the measures of syntactic elaboration and diversity is more accurate than syntactic elaboration alone. The strongest predictor variable in the combined model was MSTTR-50 with an EXP(B) value of 13.98, which implies that, for every additional point scored on the measure of syntactic diversity, the likelihood of being in a higher proficiency level of writing increases by a factor of 13.98.

Table 7

Omnibus Test of the Combined Model

Likelihood Ratio Chi-Square	df	Sig.
53.21	6.00	0.00

The incorporation of a measure of syntactic diversity (i.e., MSTTR-50) significantly improved the predictive power of the model that included both syntactic elaboration and diversity measures in predicting the writing proficiency of expert and novice academic writers. As shown in Table 7, the Likelihood Ratio Chi-Square test value of 53.214 (df = 6.00, $p < 0.001$) confirms the significant impact of the inclusion of a measure of syntactic diversity (MSTTR-50) on the predictive power of the combined model. MSTTR-50 emerged as the strongest predictor in the combined model, suggesting that syntactic variety is a crucial element to consider in addition to syntactic elaboration measures when predicting writing proficiency across all three groups (MA students, PhD students, and expert writers).

Discussion

We set out to explore Syntactic Complexity (SC) in academic writing by analyzing the texts produced by expert and novice writers. In this part, we present and discuss the findings and compare them with those of previous research.

In response to our first research question, we found that three measures of SC, namely MLTU, MLC, and CN/T, were statistically significant predictors of academic writing proficiency. According to Norris and Ortega (2009), MLTU is a measure of global SC, while MLC measures complexity at the phrasal level. However, global measures like MLTU do not provide information about the specific

modifications that contribute to the complexity of the structures (Kreyer & Schaub, 2018). Therefore, global measures are sometimes referred to as “omnibus measures” (Biber et al., 2020) since they amalgamate multiple structural units into a single measure.

Regarding the two measures (clauses per sentences and dependent clauses per T-unit), which did not make a significant contribution to the model, there are two potential reasons. First, it is possible that they were not as strongly correlated with writing expertise as the other three measures. Second, they might have been highly correlated with other measures, leading to multicollinearity. We checked multicollinearity and found that it was not the reason why two of the measures did not make a significant contribution to the model. Instead, these measures were not strongly correlated with writing expertise compared to the other measures that did make a significant contribution.

However, the present study identified two length-based measures of SC that significantly predicted academic writing proficiency. This finding supports the conventional belief that longer structural units lead to more syntactically complex structures (Ortega, 2003). Nonetheless, measures solely based on length may not offer insights into the particular modifications made within the structures, making it difficult to interpret the scores. (Kyle & Crossly, 2018). Yet they represent the overall dimension of complexity and can measure long-term changes that other measures may not capture (Norris & Ortega, 2009).

The results obtained from this study, which introduces MLTU as a significant predictor of academic writing proficiency, are consistent with the findings of some previous research, such as Crossley and McNamara's (2014) work, but are in contrast to those of other studies, such as Youn's (2014). On the whole, the previous literature implies that texts generated by more skilled L2 writers typically exhibit longer T-units (Kyle & Crossly, 2018). However, differences in participant proficiency levels, academic genres, and other methodological issues may account for the mixed results across studies. For instance, Youn's (2014) study involved undergraduate academic writers with different L1 backgrounds (e.g., Japanese, Korean, and Chinese), whereas the present study focused on graduate and expert academic writers who shared the same L1 (Farsi).

The finding that both global and specific measures of SC were able to

differentiate between the proficiency levels of academic writers suggests that different stages of writing development may require different measures for assessing SC. This idea is consistent with the dynamic systems theory (DST) (Larsen-Freeman, 2006) approach to learning, which posits that language development is not a discrete, stage-like process, but rather a dynamic, adaptive system where particular structures may be progressive or regressive at different times. Thus, given the ever-changing nature of learners' syntactic abilities at different acquisition stages, language learners may use different syntactic devices (such as complexity through coordination, subordination, phrasal elaboration, etc.) to complexify their texts. As Bi and Jiang (2020) noted, "results of previous studies targeting advanced college-level EFL learners may not be applicable to less proficient EFL learners" (p. 2).

The results of this investigation demonstrated that phrasal indices of SC were able to predict writing proficiency in both novice and professional academic writers. This part of results is consistent with the assertion made by Biber et al. (2011) that skilled academic writers tend to incorporate a larger number of noun phrase modifiers into their writing. However, this improvement in phrasal measures did not come at the expense of global measures (MLTU), indicating that there is a trade-off between phrasal and global metrics. Biber et al. also suggested that complex noun phrases are more frequent in scholarly writing, while clausal structures are more common in conversations. Similarly, Staples et al. (2016) argued that phrasal constructions continue to evolve throughout the university years even for native English speakers. These results indicate that the complexity features used to assess advanced academic writing need to be different from those used to assess other registers that are typically developed without formal instruction (e.g., spoken register).

The finding that subordination-based measures of SC did not distinguish academic writers in terms of writing proficiency does not necessarily imply that advanced academic writing is not complex in terms of these linguistic features. Instead, it suggests that academic writers do not use these features with greater or lesser frequency (Casal & Lee, 2019).

The present study's findings regarding phrasal features are consistent with earlier research on L2 writing development. For example, Crossley and McNamara (2014) found that, over a semester, L2 learners produced texts aligned with the

features of academic writing, including greater use of nouns and phrasal complexity. Similarly, Staples et al. (2016) argued that phrasal features become more prevalent as academic level increases, while clausal features are more common in lower-level texts. According to Staples and Reppen (2016), phrasal features represent an important component of scholarly writing associated with higher proficiency in both L1 and L2 writing.

An interesting finding of the present study was that the syntactic diversity measure (MSTTRDR-50) was the strongest predictor of academic writing proficiency in the combined model. Furthermore, including the diversity measure improved the predictability of the model by 9.10% compared to the previous model that did not include it. These results suggest that syntactic diversity is a crucial predictor of academic writing proficiency in the L2 context.

Conclusion and Implications

The present study investigated the relationship between various measures of SC and writing proficiency among young EFL learners. The findings indicated that global and phrasal measures (MLTU, MLC, and CN/T) had a significant unique contribution to the prediction of writing proficiency. Furthermore, the insertion of syntactic diversity measure (MSTTR-50) improved the prediction of proficiency, emphasizing the significance of assessing SC from multiple dimensions by using various methods of elaboration. The findings suggest that global and specific measures of SC can complement each other in assessing academic writing proficiency, and new length-insensitive indices capturing complexity as syntactic diversity can provide a clearer picture of the construct of SC in academic writing.

The implications of the study can be meaningful in the instruction of academic writing, as the findings suggest that L2 writing classes should provide clear guidance on producing varied and complex sentence structures. Classroom activities should assist new writers in recognizing the widespread utilization of elaborated/diversified features in academic writing, identifying the contexts where these features can be used, and using these features in their own writing. The present study has some limitations, such as investigating only one particular part-genre and using only one syntactic diversity measure, which should be considered in future research.

References

- Ansarifar, A., Shahriari, H., & Pishghadam, R. (2018). Phrasal complexity in academic writing: A comparison of abstracts written by graduate students and expert writers in applied linguistics. *Journal of English for Academic Purposes*, 31(1), 58-71. <https://doi.org/10.1016/j.jeap.2017.12.008>
- Azadnia, M., Lotfi, A., & Biria, R. (2019). A study of SC via Coh-metrix: Similarities and differences of Ph. D. dissertations written by Iranian university students and English native speakers. *Research in English Language Pedagogy*, 7(2), 232-254. <https://doi.org/10.30486/relp.2018.663453>
- Beers, S. F., & Nagy, W. E. (2009). SC as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and Writing*, 22(2), 185-200. <https://doi.org/10.1007/s11145-007-9107-5>
- Bi, P., & Jiang, J. (2020). SC in assessing young adolescent EFL learners' writings: Syntactic elaboration and diversity. *System*, 91. <https://doi.org/10.1016/j.system.2020.102248>
- Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press.
- 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>
- Biber, D., Gray, B., & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics*, 37(5), 639-667. <https://doi.org/10.1093/applin/amu059>
- Biber, D., Gray, B., Staples, S., & Egbert, J. (2020). Investigating grammatical complexity in L2 English writing research: Linguistic description versus predictive measurement. *Journal of English for Academic Purposes*, 46, 1-15. <https://doi.org/10.1016/j.jeap.2020.100869>
- Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21-46). John Benjamins Publishing Company.
- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26(1), 42-65. <https://doi.org/10.1016/j.jslw.2014.09.005>
- Bulté, B., & Housen, A. (2018). SC in L2 writing: Individual pathways and emerging group trends. *International Journal of Applied Linguistics*, 28(1), 147-164. <https://doi.org/10.1111/ijal.12196>

- Casal, J. E., & Lee, J. J. (2019). SC and writing quality in assessed first-year L2 writing. *Journal of Second Language Writing*, 44, 51-62. <https://doi.org/10.1016/j.jslw.2019.03.005>
- Casal, J. E., Lu, X., Qiu, X., Wang, Y., & Zhang, G. (2021). SC across academic research article part-genres: A cross-disciplinary perspective. *Journal of English for Academic Purposes*, 52. <https://doi.org/10.1016/j.jeap.2021.100996>
- Casanave, C. P. (1994). Language development in students' journals. *Journal of Second Language Writing*, 3(3), 179-201. [https://doi.org/10.1016/1060-3743\(94\)90016-7](https://doi.org/10.1016/1060-3743(94)90016-7)
- Chen, D., & Manning, C. D. (2014). A fast and accurate dependency parser using neural networks. In Y. Marton (Ed.), *The 2014 conference on empirical methods in natural language processing* (pp. 740-750). Association for computational linguistics.
- Connor-Linton, J., & Amoroso, L. (Eds.). (2014). *Measured language: Quantitative studies of acquisition, assessment, and variation*. Georgetown University Press.
- Crossley, S. A., & McNamara, D. S. (2010). Cohesion, coherence, and expert evaluations of writing proficiency. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 32, No. 32).
- Crossley, S. A., & McNamara, D. S. (2014). Does writing development equal writing quality? A computational investigation of SC in L2 learners. *Journal of Second Language Writing*, 26(1), 66-79. <https://doi.org/10.1016/j.jslw.2014.09.006>
- De Clercq, B., & Housen, A. (2017). A cross-linguistic perspective on SC in L2 development: Syntactic elaboration and diversity. *The Modern Language Journal*, 101(2), 315-334. <https://doi.org/10.1111/modl.12396>
- De Marneffe, M. C., & Manning, C. D. (2008). *Stanford typed dependencies manual* (pp. 338-345). Technical report, Stanford University.
- De Marneffe, M. C., & Nivre, J. (2019). Dependency grammar. *Annual Review of Linguistics*, 5, 197-218. <https://doi.org/10.1146/annurev-linguistics-011718-011842>
- Dong, J., Wang, H., & Buckingham, L. (2023). Mapping out the disciplinary variation of SC in student academic writing. *System*, 113. <https://doi.org/10.1016/j.system.2022.102974>
- Esfandiari, R., & Ahmadi, M. (2021). Syntactic complexity measures and academic writing proficiency: A corpus-based study of professional and students' prose. *Journal of Asia TEFL*, 18(3), 745-763. <http://dx.doi.org/10.18823/asiatefl.2021.18.3.1.745>
- Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. Longman.
- Karami, M., & Salahshoor, F. (2014). The relative significance of lexical richness and SC as predictors of academic reading performance. *International Journal of Research*

- Studies in Language Learning*, 3(2), 17-28. <http://doi.org/10.5861/ijrsl.2013.477>
- Kemper, S., Herman, R. E., Lian, C. H. T., & Liu, Y. (2003). Age differences in sentence production. *Journal of Gerontology: Psychological Sciences*, 58(5), 260-268. <http://doi.org/10.1093/geronb/58.5.p260>
- Kreyer, R., & Schaub, S. (2018). The development of phrasal complexity in German intermediate learners of English. *International Journal of Learner Corpus Research*, 4(1), 82-111. <https://doi.org/10.1075/ijlcr.16011.kre>
- Kuiken, F., & Vedder, I. (2012). Syntactic complexity, lexical variation and accuracy as a function of task complexity and proficiency level in L2 writing and speaking. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 143-170). John Benjamins.
- Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of SC and usage-based indices of SC sophistication* [Unpublished doctoral dissertation]. Georgia State University. doi: <https://doi.org/10.57709/8501051>
- Kyle, K., & Crossley, S. A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *The Modern Language Journal*, 102(2), 333-349. <https://doi.org/10.1111/modl.12468>
- Kyle, K., Crossley, S., & Verspoor, M. (2021). Measuring longitudinal writing development using indices of SC and sophistication. *Studies in Second Language Acquisition*, 43(4), 781-812. <https://doi.org/10.1017/S0272263120000546>
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590-619. <https://doi.org/10.1093/applin/aml029>
- Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 579-589. <https://doi.org/10.1093/applin/amp043>
- Larsson, T., Reppen, R., & Dixon, T. (2022). A phraseological study of highlighting strategies in novice and expert writing. *Journal of English for Academic Purposes*, 60. <https://doi.org/10.1016/j.jeap.2022.101179>
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40(3), 387-417. <https://doi.org/10.1111/j.1467-1770.1990.tb00669.x>
- Leonard, L. B., Weismer, S. E., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research*, 50(2), 408-428.
- Li, Y., Lin, S., Liu, Y., & Lu, X. (2023). The predictive powers of fine-grained syntactic

- complexity indices for letter writing proficiency and their relationship to pragmatic appropriateness. *Assessing Writing*, 56, 100707.
<https://doi.org/10.1016/j.asw.2023.100707>
- Liu, L., & Li, L. (2016). Noun phrase complexity in EFL academic writing: A corpus-based study of postgraduate academic writing. *The Journal of Asia TEFL*, 13(1), 48-65.
<http://dx.doi.org/10.18823/asiatefl.2016.13.1.4.48>
- Liu, X., & Koirala, H. (2012). Ordinal regression analysis: Using generalized ordinal logistic regression models to estimate educational data. *Journal of Modern Applied Statistical Methods*, 11(1), 242-254.
- Long, J. S., & Freese, J. (2006). *Regression models for categorical dependent variables using Stata* (Vol. 7). Stata press.
- Lu, X. (2009). Automatic measurement of syntactic complexity in child language acquisition. *International Journal of Corpus Linguistics*, 14(1), 3-28.
<https://doi.org/10.1075/ijcl.14.1.02lu>
- 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(1), 16-27. <https://doi.org/10.1016/j.jslw.2015.06.003>
- Mansourizadeh, K., & Ahmad, U. K. (2011). Citation practices among non-native expert and novice scientific writers. *Journal of English for Academic Purposes*, 10(3), 152-161. <https://doi.org/10.1016/j.jeap.2011.03.004>
- Miestamo, M., Sinnemäki, K., & Karlsson, F. (Eds.). (2021). *The handbook of dependency linguistics*. John Wiley & Sons.
- Mohamadi Zenouzagh, Z. (2020). Syntactic complexity in individual, collaborative and E-collaborative EFL writing: mediating role of writing modality, L1 and sustained development in focus. *Educational Technology Research and Development*, 68(6), 2939-2970. <https://doi.org/10.1007/s11423-020-09818-w>
- Naserpour, A., Zarei, A., & Esfandiari, R. (2020). The effects of task orientation and involvement load on learning collocations. *Teaching English as a Second Language Quarterly*, 39(1), 71-114. <https://doi.org/10.22099/jtls.2020.37032.2809>
- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555-578.
<https://doi.org/10.1093/applin/amp044>
- 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>
- Osborne, J. W. (2017). *Regression and linear modeling: Best practices and modern methods*. SAGE Publishing.
- Parkinson, J., & Musgrave, J. (2014). Development of noun phrase complexity in the writing of English for Academic Purposes students. *Journal of English for Academic Purposes*, 14(1), 48-59. <https://doi.org/10.1016/j.jeap.2013.12.001>
- Perry, S. (2020). Ordinal logistic regression models and statistical software: What you need to know. *Cornell statistical consulting unit*. https://cscu.cornell.edu/wp-content/uploads/91_ordlogistic.pdf
- Pituch, K. A., & Stevens, J. P. (2015). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS*. Routledge.
- Römer, U., & Arbor, A. (2009). English in academia: Does nativeness matter. *Anglistik: International Journal of English Studies*, 20(2), 89-100.
- Ruan, Z. (2018). Structural compression in academic writing: An English-Chinese comparison study of complex noun phrases in research article abstracts. *Journal of English for Academic Purposes*, 36(1), 37-47. <https://doi.org/10.1016/j.jeap.2018.09.001>
- Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education* (Vol. 22). John Benjamins Publishing.
- Shadloo, F., Shahriari Ahmadi, H., & Ghonsooly, B. (2017). Exploring grammatical complexity across proficiency levels in argumentative essays by Iranian EFL learners. *Language and Translation Studies (LTS)*, 49(4), 61-75. <https://doi.org/10.22067/lts.v49i4.61672>
- Skehan, P. (1998). Task-based instruction. *Annual Review of Applied Linguistics*, 18, 268-286. <https://doi.org/10.1017/S0267190500003585>
- Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexico-grammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing*, 32, 17-35. <https://doi.org/10.1016/j.jslw.2016.02.002>
- Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication*, 33(2), 149-183. <https://doi.org/10.1177/0741088316631527>
- Valizadeh, M., & Soltanpour, F. (2021). Focused direct corrective feedback: Effects on the elementary English learners' written SC. *Eurasian Journal of Applied Linguistics*, 7(1), 132-150. <https://doi.org/10.32601/ejal.911207>

Verspoor, M., Schmid, M. S., & Xu, X. (2012). A dynamic usage-based perspective on L2 writing. *Journal of Second Language Writing*, 21(3), 239-263.

<https://doi.org/10.1016/j.jslw.2012.03.007>

Wolfe-Quintero, K., Inagaki, S., & Kim, H. (1998). *Second language development in writing: Measures of fluency, accuracy and complexity*. University of Hawaii at Manoa.

Yazdani, S. (2018). Syntactic complexity in Iranian learners' English writing and speaking. *Journal on English as a Foreign Language*, 8(1), 75-96.

<http://e-journal.iain-palangkaraya.ac.id/index.php/jefl>

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>

Youn, S. J. (2014). Measuring syntactic complexity in L2 pragmatic production: Investigating relationships among pragmatics, grammar, and proficiency. *System*, 42, 270-287.

<https://doi.org/10.1016/j.system.2013.12.008>

Zheng, B., & Agresti, A. (2000). Summarizing the predictive power of a generalized linear model. *Statistics in Medicine*, 19(13), 1771-1781.

[https://doi.org/10.1002/1097-0258\(20000715\)19:13<1771:AID-SIM485>3.0.CO;2-P](https://doi.org/10.1002/1097-0258(20000715)19:13<1771:AID-SIM485>3.0.CO;2-P)



©2020 Alzahra University, Tehran, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0 license) (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)