Syntactic Complexity and Lexical Diversity in L1/L2 Writing of EFL Learners

Fazel Alaei

Department of English Language & Literature, University of Tehran, Iran
fazelalaei@gmail.com

Shiva Kaivanpanah (Corresponding Author)

Department of English Language & Literature, University of Tehran, Iran
shkaivan@ut.ac.ir

Maedeh Mohammadi

Department of English Language & Literature, University of Tehran, Iran
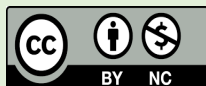ad88mohammadi@gmail.com

Abstract

The present study examined the difference between L1 and L2 lexical diversity of argumentative and narrative writings of L2 learners, and the contribution of syntactic complexity and lexical diversity to the writing quality in the L2 argumentative and narrative writings of EFL learners. To this end, 46 pre-intermediate and intermediate Iranian EFL learners from four intact classes wrote one argumentative and one narrative essay in L2, and one argumentative and one narrative essay in L1 on different topics. Paired-samples *t*-tests revealed that lexical diversity surfaced more in the L1 writing of the learners. Multiple linear regressions indicated that among five measures of syntactic complexity, mean length of T-unit and clauses per T-unit better predict the quality of argumentative writing. In addition, complex nominals per clause are better predictors of narrative writing quality. Simple linear regressions showed that lexical diversity is a significant predictor of L2 writing in both genres. Based on the findings, writing instructors are advised to provide L2 learners with explicit instruction on the use of diverse vocabulary and different syntactic structures in order to help them improve the quality of their writing.

Introduction

Communicating messages and views requires writers to express their ideas linguistically. The linguistic choices writers make influence the perception of writing quality (Crossley & McNamara, 2012); higher-rated texts contain syntactically complex structures (Casal & Lee, 2019; Ortega, 2015) and a diverse range of vocabulary (Ha, 2019; Kim, 2014). Syntactic complexity (hereafter SC) is a predictor of writing quality and makes it possible for writers to produce more expressive language using certain grammatical forms (Beers & Nagy, 2011). SC measures the degree of varied and sophisticated language structures produced by learners (e.g., Beers & Nagy, 2009; Casal & Lee, 2019; Kim & Crossley, 2018; Lu, 2011; Ortega, 2003). Another predictor of writing quality is lexical diversity (hereafter LD), which measures vocabulary knowledge and range (e.g., Choi, 2017; Chon & Lee, 2015; Wu, Dixon, Sun, & Zhang, 2019; Yu, 2009).

The investigation into SC and LD and their relationships with writing quality is not new. These relationships are complex and many early studies in this area have failed to show consistent patterns (Beers & Nagy, 2009; Wang & Slater, 2016). Given the need for distinct types of grammatical structures and vocabulary in different writing genre types (Beers & Nagy, 2009; Olinghouse & Wilson, 2013), we cannot presume that the relationship between SC and LD with writing quality remains unchanged across different genres. L2 writers need to recognize the discourse forms required for different genres and deploy certain types of grammatical forms and vocabulary (Beers & Nagy, 2009; Olinghouse & Wilson, 2013) to serve different communicative functions (Berman, 2008; Berman & Nir-Sagiv, 2007).

Many studies have examined the relationship between writing quality and SC (e.g., Beers & Nagy, 2009; Casal & Lee, 2019; Martínez, 2018; Yang, Lu, & Weigle, 2015) and writing quality and LD (González, 2017; McNamara, Louwerse, McCarthy, & Graesser, 2010; Yu, 2009) in either L1 (Beers & Nagy, 2009; Olinghouse & Leaird, 2009) or L2 writing (Engber, 1995; Ortega, 2003; Yang, Lu, & Weigle, 2015; Yu, 2009). However, to the best of the researchers' knowledge, no study has examined how genre mediates the effects of SC and LD in L1 and L2 writings of L2 writers. Therefore, the present study addresses these concerns.

1. Literature review

1.1. Syntactic complexity

In a sizeable number of studies in second language research, complexity, accuracy, and fluency (hereafter CAF) are used to describe second language performance (Crookes, 1989; Ellis, 2009; Ellis & Barkhuizen, 2005; Skehan, 1998). In second language acquisition research, learner language can be analyzed employing only one of the components of CAF which "implies the major stages of change in the underlying L2 system" (Housen, Kuiken, & Vedder, 2012, p. 3). The first component of CAF is complexity, which will be explored as a predictor of writing quality in the present study. Ellis and Barkhuizen (2005,139) define complexity as "the extent to which learners produce elaborated language". Housen and Kuiken (2009) consider accuracy as the most uncomplicated aspect of CAF and define it as the language which is error-free.

Fluency as another dimension of CAF refers to writing quickly and smoothly, which may not necessarily be grammatical (Hammerly, 1991).

In L2 writing, SC is defined with regard to the range and the level of sophistication of the structures emerging in the written language (Lu, 2011; Ortega, 2003). Three major measures (frequency, ratio, and index) are deployed to examine SC. Frequency refers to the rate at which specific syntactic structures occur, ratio is the number of a production unit divided by the whole number of another unit, and index refers to particular measures employed to calculate numerical data, which include the coordination index, the complexity formula and complexity index.

Exploring the relationship between SC and writing quality can be important to L2 writing teachers as it may offer new insights into how L2 writers can improve their writing quality. Therefore, researchers have always been looking for best measures of SC which can predict writing quality. In an early study, Wolfe-Quintero, Inagaki, and Kim (1998) reported that dependent clauses per clause and mean number of clauses per T-unit are the best measures of SC. In another study, Taguchi, Crawford, and Wetzel (2013) reported that noun phrase modification contributes to writing quality. Casal and Lee (2019) found a significant correlation between writing quality and SC measures including complex nominals per clause and the mean length of T-units in the essays of college students. More recently, Bi and Jiang (2020) noted that mean length of sentence, complex nominals per clause, and clauses per T-unit are the best predictors of writing quality in the narrative essays of Chinese EFL learners at beginner and intermediate levels.

Research shows that the relationship between SC and writing quality is mediated by genre (Crowhurst, 1980). Beers and Nagy (2009) partly support the conclusion that the relationship between SC and writing quality is influenced by genre as each genre may demand the use of certain types of grammatical forms by writers. A similar conclusion is made by Berman and Nir-Sagiv (2007) and Ravid and Berman (2010) who noted that genre mediates the relationship between SC and writing quality.

Genres are regarded as socially constructed language use which are distinguished by their specific communicative purposes (Halliday & Hasan, 1985). Genres can be categorized into narratives and non-narratives among which is argumentative. In order to attain the specific communicative and social goals, each genre tends to be different in language features (Beers & Nagy, 2011) affecting the relationship between SC and writing quality. Schultz (1991) argues that the most demanding genre is argumentative; this conclusion is supported by Lu (2011) who reported that argumentative compositions have higher syntactic complexity compared to narrative compositions. Research on L1 writing has shown that argumentative texts elicit more syntactically and lexically complex structures (Beers & Nagy, 2009; Berman & Nir-Sagiv, 2007) than narrative texts. Despite the interest in L2 writing, few studies have compared L2 writing performance in argumentative and narrative genres. Therefore, this study examines the SC measures and LD in argumentative and narrative genres, the most used genres in our context.

1.2. Measures of Syntactic Complexity

Measures of SC can indicate the "variation" and "sophistication" (Lu, 2011) of structures. These measures include global complexity (e.g., mean length of sentence and mean length of T-unit), complexity by subordination (e.g., mean number of clauses per T-unit), and complexity via subclausal or phrasal elaboration (e.g., mean length of clause, and complex nominals per clause) (Norris & Ortega, 2009). In this study, five measures of SC were used to predict L2 writing quality in argumentative and narrative genres. The first SC measure adopted is mean length of T-unit which is the most commonly used measure of SC and a common indicator of syntactic development (Nippold, Hesketh, & Duthie, 2005). The second SC measure utilized in the current study is clauses per T-unit. Clauses per T-unit can have predictive power at intermediate level of proficiency (Norris & Ortega, 2009); for intermediate levels, clausal subordination (including clauses per T-unit) can be used to discriminate between L2 writers (Ortega, 2003). Although research shows that mean length of T-unit and clauses per T-unit are predictors of writing quality (Bulté & Housen, 2014; Yang et al., 2015), they are criticized because measures based on T-unit do not indicate the difference between clausal and phrasal dimensions of SC, and measures from phrasal dimensions must be deployed (Biber et al., 2013); thus, mean length of clause and complex nominals per clause are utilized as the third and fourth measures of SC in the present study. These measures are strong predictors of L2 writing and have discriminative power (e.g., Bulté & Housen, 2014; Lu, 2011; Yang et al., 2015). Deploying mean length of clause in this study can show us the level of complexity within clauses (Beers & Nagy, 2009) because complex writing conveys meaning in a concise manner by summarizing information from multiple clauses and expressing it in a single clause, which leads to the reduction of the number of total words and an increase in the number of words per clause. The last selected SC measure is mean length of sentence which according to Bi and Jiang (2020) is the most reliable indicator of writing quality for learners at low to intermediate levels of proficiency.

1.3. Lexical Diversity

Lexical knowledge has been recognized to be essential for academic success (Daller, van Hout, & Treffers-Daller, 2003). Recognizing its significance, researchers deploy different measurements to assess lexical richness. One of the measures to gauge the lexical richness of texts is LD (Read, 2000) which is defined as the range of vocabulary in a learner's written or spoken language. Measuring the level of LD requires a comparison between different words (types) and the total number of words (tokens). Type-token ratio (TTR) has traditionally been used to measure LD (e.g., Lieven, 1978); however, it is sensitive to text length (Nation & Webb, 2011) because the longer the text is, the fewer the chances are for new types to surface, and as a result TTR values are lower. Hence, addressing this weakness of TTR, Malvern, Richards, Chipere, and Durán (2004) introduced the D measure defined as "a robust measure of lexical diversity which is not a function of sample size in the way TTR and its simple transformations are" (Malvern et al., 2004, 60). D permits the calculation of LD of short texts; the higher the D, the higher the lexical diversity of a text (Durán, Malvern, Richards, & Chipere, 2004). The approach developed by Malvern et al. (2004) resulted in the following equation with token count (N) and the parameter D which is the lexical diversity index and is best measured by

computer. Discussing the rationale and mathematical derivation of D is beyond the scope of this study. In brief, after creating a group of ideal curves of TTR versus tokens (N), the equation obtains the ideal curve best matching the actual curve of TTR versus tokens (N) for the language sample which is the parameter D (Malvern et al., 2004).

$$TTR = \frac{D}{N}\left[\left(1 + 2\,\frac{N}{D}\right)^{\frac{1}{2}} - 1\right]$$

An increasing number of studies have emphasized the importance of lexical richness and variation in L2 writing (Chen, 2020; González, 2017; Lu, 2012; Olinghouse & Leaird, 2009), as they are directly related to the writer's ability to make an effective communication (Lu, 2012). Researchers (e.g., Bestgen, 2017; González, 2017; Ha, 2019; Kim, 2014; Vögelin, Jansen, Keller, Machts, & Möller, 2019) note that LD predicts the quality of L2 writing; studies comparing L1 and L2 writing have shown a strong correlation between LD and the quality of writing. For instance, Engber (1995) found significant correlations between LD and quality of compositions implying that compositions with higher LD tend to receive better scores. Another study by Yu (2009) provided further support for the relationship between LD and the quality of spoken and written language produced by L2 learners. More recently, González (2017) and Vögelin et al. (2019) reported that L2 texts demonstrating greater lexical diversity were assessed more positively.

The relationship between LD and the quality of writing is not always straightforward. Research indicates that text genre impacts the relationship between LD and quality of writing (Olinghouse & Wilson, 2013). Analyzing narrative, persuasive, and informative genres, Olinghouse and Wilson (2013) found that genre influenced learners' vocabulary usage, and among the three genres, LD could only predict the quality of narrative texts. Hence, they concluded that genre is a decisive factor influencing the contribution of vocabulary to text quality.

Using different indices of lexical knowledge (e.g., lexical diversity) to assess the range of vocabulary, L2 writing studies have recognized the differences between texts produced by L1 and L2 writers. For example, Silva's (1993) overview of 72 studies comparing L1 and L2 writing showed that L2 writers exhibited less LD. More recent studies exploring the difference between L1 and L2 writing have supported the finding that texts composed by L1 writers have more LD (Eckstein & Ferris, 2017; Nasseri & Thompson, 2021). In previous studies, comparisons were made between different L1 and L2 writers (Crossley & McNamara, 2009; Eckstein & Ferris, 2017; Linnarud, 1986; Nasseri & Thompson, 2021), and no one to the best of our knowledge has compared L1 and L2 texts produced by the same writers.

2. Research questions

In the present study, we addressed the following questions:

1. Is there any statistically significant difference between the LD of L1 and L2 argumentative and narrative writing of L2 learners?

2. Do the SC and LD of L2 argumentative and narrative writings of L2 learners contribute to writing quality?

3. Method

3.1. Participants

The study utilized ex-post facto design to address the research questions. A total of 46 pre-intermediate and intermediate Iranian English language learners (17 males and 29 females) participated in the study. All participants were native speakers of Persian with little chance to use English language in natural communication contexts. At the time of study, they were studying English at a language Institute on Kish Island in Iran. The details of participants are presented in Table 1.

Table 1. *Distribution of Participants*

Level of Proficiency		Mean Age	Gender		Total
Pre-Intermediate	Intermediate		Male	Female	
23	23	29	17	29	46

3.2. Procedures

As participants were from four different intact classes, we needed to check their homogeneity in terms of language proficiency. Prior to the study, the Oxford Placement Test (OPT) (Allan, 2004) was distributed among participants. Based on the results, only those learners whose scores fell within the range of one standard deviation around the mean ($\bar{x}$ = *55.84, SD = 9.71*) with scores between 65.55 and 46.13 were recruited for the study.

Each participant wrote four essays—one argumentative and one narrative in L2, and one argumentative and one narrative in L1. Participants were asked to write an L2 argumentative essay on 'Do you think the reason for peoples' success is because of hard work or luck?' and then an L1 argumentative essay on 'How do movies or television influence people's behavior?' They were also required to write an L2 narrative essay on 'Describe one of your memorable journeys.' and an L1 narrative essay on 'Describe one thing you cannot live without.' Essays were written at one session interval.

As essays with greater length result in repetition of lexical items and increase the possibility of skewing the data (Crossley & McNamara, 2009; Engber, 1995), participants were instructed to write 250 words for each essay in 35 minutes. All L2 essays were scored by three experienced IELTS examiners based on IELTS scoring criteria: fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation.

3.3. Measures and instruments

D for L1 and L2 essays was calculated using free software called *vocd* in Computerized Language Analysis (CLAN) programs (Macwhinney, 2000) which is available online. *Vocd* does not use any special dictionary, and just (types) and (tokens) matter. In order for *vocd* to recognize Persian essays, they were Romanized, and then entered into CHAT Format to be

analyzed by *vocd*. As each Persian word equals an English word, and similar to English, Persian words are separated by space, Romanization was feasible.

Essays in English were also typed and converted to CHAT format for analysis by *vocd*. In order to avoid entering the same word with different misspellings which can potentially increase D, we followed Yu's (2009) suggestion. A spelling mistake was corrected if it was spelt correctly elsewhere in the same essay. For misspelt words which were repeated once, no correction was made. If a word was misspelt differently in the same text, they were replaced with the correct spelling.

The syntactic complexity analyzer (SCA) developed by Lu (2010) was used to measure SC of our L2 data. Five measures of the 14 different indices analyzed by SCA were adopted in this study: three length-related measures (i.e., mean length of clause, mean length of sentence, and mean length of T-unit), one subordination-related (i.e., clauses per T-unit), and one nominal-related measure (complex nominals per clause). Table 2 presents indices adopted by this study.

Table 2. *Description for Complexity Measures Deployed in this Study*

Sub-construct	Definition	Code
Length of production unit	Mean length of clause	MLC
	Mean length of sentence	MLS
	Mean length of T-unit	MLT
Subordination	Clauses per T-unit	C/T
Phrasal measure	Complex nominals per clause	CN/C
Lexical Diversity	vocd-D	D

4. Results

A paired-samples *t*-test was used to examine the difference between LD in L1 and L2 argumentative essays. Before conducting *t*-test, the normality assumption was checked. As Table 3 displays, Skewness and Kurtosis ratios fell within the range +/- 2 indicating that the normality assumption for running *t*-test was not violated.

Table 3. *Descriptive Statistics for L1 and L2 Argumentative Writings (N= 46)*

D scores	Min	Max	Mean	Std. Deviation	Skewness	Std. Error	Kurtosis	Std. Error
L1 argumentative	59.43	108.67	83.84	11.03	0.16	0.35	0.37	0.69
L2 argumentative	46.50	125.05	74.12	15.72	1.20	0.35	1.86	0.69

As shown in Table 3, D in L1 argumentative essays ($\bar{x}$ = 83.84, SD = 11.03) was higher than D in L2 argumentative essays ($\bar{x}$ = 74.12, SD = 15.72). The paired-samples *t*-tests indicated that the difference in D in L1 and L2 argumentative essays was statistically significant, t (45) = 3.74, p = .00, with the Cohen's effect size value being .50, which can be considered as a medium-size effect (Sawilowsky, 2009) (see Table 4 for mean difference).

Table 4. *Paired-Samples t-test Examining the Difference between D in L1 and L2 Argumentative Essays*

		Paired Differences								Cohen's d
				Std. Error	95% CI				*p*	
		Mean	Std. Deviation	Mean	Lower	Upper	t	df	value	
Writing	L1 argumentative D L2 argumentative D	9.72	17.60	2.59	4.49	14.94	3.74	45	.00	.50

Another paired-samples *t*-test was utilized to examine the difference between LD in L1 and L2 narrative writings. Prior to conducting *t*-test, the normality assumption was checked. As Table 5 shows, Skewness and Kurtosis ratios fell within the range +/- 2 indicating that normality assumption for running *t*-test was tenable.

Table 5. *Descriptive Statistics for Two Narrative Writings (N= 46)*

	Min	Max	Mean	Std. Deviation	Skewness		Kurtosis	
D scores						Std. Error		Std. Error
L1 narrative	46.89	119.95	82.15	15.95	0.21	0.35	0.30	0.69
L2 narrative	45.14	103.13	70.51	14.89	0.44	0.35	-0.51	0.69

As presented in Table 5, D in L1 narrative essays ($\bar{x}$ = 82.15, *SD* = 15.95) was higher than D in L2 narrative essays ($\bar{x}$ = 70.51, *SD* = 14.89). The paired-samples *t*-tests showed that the difference between D in L1 and L2 narrative essays was statistically significant, *t* (45) = 4.30, *p* = .00, with the Cohen's effect size value of a medium-size effect (0.53) (Sawilowsky, 2009) (see Table 6 for mean difference).

Table 6. *Paired-Samples t-test Examining the Difference between D in L1 and L2 Narrative Essays*

		Paired Differences								Cohen's d
				Std. Error	95% CI				*p*	
		Mean	Std. Deviation	Mean	Lower	Upper	T	df	value	
Writing	L1 narrative D L2 narrative D	11.64	18.35	2.70	6.19	17.09	4.30	45	.00	.53

To address the second research question, a multiple linear regression was utilized to ascertain whether SC of L2 argumentative writing contributes to writing quality. To do so, MLC, MLS, MLT, C/T, and CN/C, as indicators of SC, were entered into the regression model to determine whether they can predict writing quality in L2 argumentative essays ($\bar{x}$ = 5.51, *SD* = 0.72). Table 7 presents the descriptive statistics of the five SC indices of our data (i.e., MLC, MLS, MLT, C/T, CN/C) and the criterion variable (i.e., L2 argumentative IELTS scores) in the regression model.

Table 7. *Descriptive Statistics for SC Indices and L2 Argumentative IELTS Scores (N = 46)*

	Mean	Std. D	Skewness		Kurtosis	
	Value	Value	Value	Std. Error	Value	Std. Error
L2 Argumentative IELTS Scores	5.51	.72	-.42	.35	-.35	.69
MLC	7.73	1.53	1.56	.35	1.18	.69
MLS	19.40	3.17	.10	.35	.59	.69
MLT	14.65	2.22	-.31	.35	.73	.69
C/T	1.93	.33	-.09	.35	-.32	.69
CN/C	.87	.29	1.10	.35	1.80	.69

Note: MLC $=$ Mean length of clause; MLS = Mean length of sentence; MLT= Mean length of T-unit; C/T= Clauses per T-unit; CN/C = Complex nominals per clause.

The five syntactic measures (i.e., MLC, MLS, MLT, C/T, and CN/C) were entered into a stepwise regression. Table 8 and Table 9 show the significance of the regression model, F (5, 40) = 2.84, p = .02, with an R^2 of 0.26. On the basis of the adjusted R^2 value (i.e., R^2 = .17), it can be argued that the five measures as a whole predicted 17% of variability (a modest model in terms of variance exaplained, see Plonsky & Ghanbar, 2018) in L2 argumentative writing.

Table 8. *Regression Equation Test of Significance*

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	6.06	5	1.21	2.84	0.02
Residual	17.18	40	0.42		
Total	23.24	45	-	-	-

The assumptions concerning normality, linearity and homoscedasticity of residuals were checked for regression analysis, and the outcomes did not violate the assumptions (see Plonsky & Ghanbar, 2018). As seen in Table 9, the Durbin-Watson value of 1.8 (between 1.5 and 2.5) represents no autocorrelation. The condition index lower than 15 indicates no collinearity (Tabachnick & Fidell, 2013). No cases of multicollinearity were observed when VIF was used to detect multicollinearity. Since skewness and kurtosis were between -2 and +2 (see Table 7), the data can be considered normal.

Table 9. *R, R^2, adjusted R^2, and Test of Independence of Residuals of Simple Regression Analysis*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
Estimates	0.51	.26	.17	0.65	1.8

As can be seen in Table 10, MLT (B = .49, S.E = .19 , $\hat{a}$ = 1.51, t = 2.53, p =.02) and C/T (B = -3.48, S.E = 1.40 , $\hat{a}$ = -1.59, t = -2.49, p = .02) were significant predictors of L2 argumentative essays. It can be argued that MLT and C/T predicted 17% of variance in L2 argumentative essays. Hence, it can be claimed that among all the indicators of SC (i.e., MLC, MLS, MLT, C/T, CN/C), MLT and C/T can predict the quality of L2 argumentative writing (see Table 10 for other non-significant predictors).

Table 10. *Regression Coefficients of Regression Analysis Pertaining to L2 Argumentative IELTS Scores*

Predictor	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
MLC	-.47	.31	-1.00	-1.50	.14	.42	6.08
MLS	-.07	.04	-.30	-1.59	.12	.50	1.97
MLT	**.49**	**.19**	**1.51**	**2.53**	**.02**	**.52**	**9.33**
C/T	**-3.48**	**1.40**	**-1.59**	**-2.49**	**.02**	**.45**	**2.16**
CNC	-.54	.70	-.22	-.78	.44	.33	4.31
MLC	-.47	.31	-1.00	-1.50	.14	.42	4.08

Note: MLC = Mean length of clause; MLS = Mean length of sentence; MLT= Mean length of T-unit; C/T= Clauses per T-unit; CN/C = Complex nominals per clause.

To find out whether SC of narrative writing of L2 learners contributes to writing quality, another multiple linear regression was conducted and MLC, MLS, MLT, C/T, and CN/C, as indicators of SC, were entered to a regression model to determine whether they can predict L2 narrative quality ($\bar{x}$ = 5.30, *SD* = 0.73). Table 11 presents the descriptive statistics of the five SC indices of our data (i.e., MLC, MLS, MLT, C/T, CN/C) and the criterion variable (i.e., L2 narrative IELTS scores) in the regression model.

Table 11. *Descriptive Statistics for SC Indices and L2 Narrative IELTS Scores (N = 46)*

	Mean	Std. D	Skewness		Kurtosis	
	Value	Value	Value	Std. Error	Value	Std. Error
L2 Narrative IELTS Scores	5.30	.73	-.06	.35	-.59	.69
MLC	7.73	1.53	1.56	.35	1.18	.69
MLS	19.40	3.17	.10	.35	.59	.69
MLT	14.65	2.22	-.31	.35	.73	.69
C/T	1.93	.33	-.09	.35	-.32	.69
CN/C	.87	.29	1.10	.35	1.80	.69

Note: MLC = Mean length of clause; MLS = Mean length of sentence; MLT= Mean length of T-unit; C/T= Clauses per T-unit; CN/C = Complex nominals per clause.

The five syntactic measures (i.e., MLC, MLS, MLT, C/T, and CN/C) were used in a stepwise regression, $F (5, 40) = 8.12$, $p = .00$, $R^2 = .50$ (see Tables 12 and 13). The measures account for 45% of variance in L2 narrative essays. Before running the linear regression analysis, the assumptions of normality, linearity and homoscedasticity of residuals were checked, and no violations of the assumptions were detected. As it is shown in Table 13, the data met the assumption of independent errors (Durbin-Watson value = 1.3). The condition index lower than 15 indicates no collinearity (Tabachnick & Fidell, 2013). When VIF was deployed to check for multicollinearity, no multicollinearity problems were detected. Likewise, the normality assumption was checked through the values of skewness and kurtosis which were between -2 and +2 (see Table 11).

Table 12. *Regression Equation Test of Significance*

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	12.21	5	2.43	8.12	0.00
Residual	12.02	40	0.30		
Total	24.23	45			

Table 13. *R*, R^2, adjusted R^2, and Test of Independence of Residuals of Simple Regression Analysis

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
Estimates	0.71	.50	.45	0.54	1.3

According to Table 14, CN/C (B = 3.13, S.E = .74 , $\hat{a}$ = 7.60, *t* = 4.22, *p* =.00) was the only significant predictor of L2 narrative essays. Therefore, it can be claimed that among all the indicators of SC (i.e., MLC, MLS, MLT, C/T, CN/C), only CN/C can predict the quality of L2 narrative essays (see Table 14 for other non-significant predictors).

Table 14. *Regression Coefficients of Regression Analysis Pertaining to L2 Narrative IELTS Scores*

Predictor	Unstandardized Coefficients		Standardized Coefficients		t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta				Tolerance	VIF
MLC	.22	.73	.25	.31	.76	.49	2.98	
MLS	-.03	.05	-.09	-.50	.62	.37	2.70	
MLT	-.37	.56	-.93	-.67	.51	.60	5.77	
C/T	4.47	3.86	.93	1.16	.25	.62	3.44	
CN/C	**3.14**	**.74**	**.76**	**4.22**	**.00**	**.38**	**2.61**	

Note: MLC = Mean length of clause; MLS = Mean length of sentence; MLT= Mean length of T-unit; C/T= Clauses per T-unit; CN/C = Complex nominals per clause.

We conducted a simple linear regression analysis to examine the contribution of LD to writing quality in L2 argumentative writing. In line with the recommendations of Plonsky and Ghanbar (2018), before running this regression analysis, its underlying statistical assumptions, namely normality, linearity, homoscedasticity, and the independence of error terms were checked and no violations were observed. All the skewness and kurtosis estimates were between -2 and +2, and Durbin-Watson test of autocorrelation of residuals (with all the values being in the range of 1.5 to 2.5) was examined. No outliers were found in our data. Table 15 presents the descriptive statistics of the criterion and predictor variable.

Table 15. *Descriptive Statistics for LD Index and L2 Argumentative IELTS Scores (N = 46)*

	Mean	Std. D	Skewness		Kurtosis	
	Value	Value	Value	Std. Error	Value	Std. Error
L2 Argumentative D	74.12	15.72	1.17	.35	1.48	.68
L2 Argumentative IELTS Scores	5.51	0.71	-0.41	.35	0.34	.68

The model can significantly account for the variance in L2 argumentative D, *F* (1, 44) = 67.45, *p* = .00, R^2 = .60. The adjusted R^2 value of .59 revealed that 59% of variability in L2

argumentative essay was predicted by L2 Argumentative D (Plonsky & Ghanbar, 2018) (B = .03, S.E = .004, â = .78, *t* = 8.21, *p* = .00) (see Tables 16 and 17).

Table 16. *Regression Equation Test of Significance*

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	14.06	1	14.06	67.45	0.00
Residual	9.17	44	0.20		
Total	23.24	45			

Table 17. *R, R^2, adjusted R^2, and Test of Independence of Residuals of Simple Regression Analysis*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
Estimates	0.77	.60	0.59	1.78	1.58

Another simple linear regression analysis was conducted to examine whether LD of narrative writing of L2 learners contributes to writing quality. Before running the regression analysis, the assumptions of normality, linearity, homoscedasticity, and the independence of error terms were checked and no violations were observed. Moreover, no outlier was detected after data screening. Table 18 presents descriptive statistics of the criterion and predictor variable.

Table 18. *Descriptive Statistics for LD Index and L2 Narrative IELTS Scores (N = 46)*

	Mean	Std. D	Skewness		Kurtosis	
	Value	Value	Value	Std. Error	Value	Std. Error
L2 Narrative D	70.51	14.89	0.44	.35	-0.59	.68
L2 Narrative IELTS Scores	5.30	0.73	-0.06	.35	-0.59	.68

Table 19 and Table 20 present the significance of the regression model, *F* (1, 44) = 83.32, *P* = .00, with an R^2 of .65. The adjusted R^2 value of .64 indicated that 64% of variability in L2 narrative essays was significantly predicted by L2 narrative D (Plonsky & Ghanbar, 2018) (B = .04, S.E = .004, â = .80, *t* = 9.12, *p* = .00).

Table 19. *Regression Equation Test of Significance*

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	86.15	1	15.86	83.32	0.00
Residual	8.37	44	0.19		
Total	24.23	45			

Table 20. *R, R^2, adjusted R^2, and Test of Independence of Residuals of Simple Regression Analysis*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
Estimates	0.80	0.65	0.64	0.43	1.44

5. Discussion

This study examined the difference between the LD of argumentative and narrative writing of L2 learners in their L1 and L2. The difference between L1 (i.e., Persian) and L2 (i.e., English)

narrative and argumentative writings composed by the same writers has scarcely been investigated in terms of LD. There are studies, however, which compared L1 and L2 writings composed by different L1 and L2 writers. These studies show that compositions produced by L1 writers contain more LD than those produced by L2 writers. For example, Eckstein and Ferris (2017) analyzed L1 and L2 writing samples for language error counts, lexical complexity and SC and found a significant difference in D suggesting L2 writers had less LD than their L1 peers. Similar findings were reported by Nasseri and Thompson (2021) who investigated lexical density and diversity in L1 and L2 (ESL and EFL) academic writing and indicated that texts produced by EFL group contained the least LD. This study used paired-samples *t*-tests to demonstrate that there was a statistically significant difference for LD between L1 and L2 narrative and argumentative essays. In both argumentative and narrative writings, the observed LD was higher in L1 than in L2. This finding highlights the need for L2 writers to develop greater lexical variation and for writing instructors to consider the potential limitation in L2 writers' vocabulary.

Regarding the relationship between SC and writing quality in argumentative and narrative genres, our results suggest that SC contributes to writing quality; however, measures of SC (MLC, MLS, MLT, C/T, CN/C) predicting the quality of argumentative and narrative writings are distinct. This finding seems to be partially in tune with previous observations on the relationship between SC and writing quality (Beers & Nagy, 2009; Bi & Jiang, 2020; Bulté & Housen, 2014; Casal & Lee, 2019; Li, 2015; Yang et al., 2015). In line with Beers and Nagy (2009), Lu (2017), and Qin and Uccelli (2016), we showed that SC has a relationship with writing quality, but this relationship depends on genre. Similar to Beers and Nagy (2009) and Qin and Uccelli (2016), we found consistent evidence that measures of SC as predictors of writing quality are distinctive for each genre.

With regard to SC measures as predictors of the quality of writing, we reported that for argumentative writing quality, MLT and C/T could contribute to L2 argumentative writing. This is in agreement with Jagaiah, Olinghouse, and Kearns (2020) whose review of 36 studies concerning SC in writing research showed that MLT and C/T surfaced more in the argumentative genre. Concerning global complexity, the scores for argumentative essays were predicted by MLT which is in line with Bulté and Housen's (2014), Casal and Lee's (2019), Yang et al.'s (2015), and Li's (2015) results showing MLT as a writing quality predictor. Regarding subordination (C/T) and its role in predicting writing quality, our findings showed that subordination could be employed to predict the quality of writing composed by learners at intermediate or lower levels. This finding is consistent with Bi and Jiang (2020), Khushik and Huhta (2019), and Lahuerta Martínez (2018). In contrast, studies investigating essays composed by more proficient writers did not find subordination-related measures as a distinguishing factor in writing quality (Casal & Lee, 2019; Kyle & Crossley, 2018). As Bi and Jiang (2020) argue, subordination has its own importance, but subordination-related indices cannot indicate high proficient learners' writing quality. It seems that language proficiency influences the choice of syntactic structures since "at different moments in the developmental process (at different proficiency levels) the very make-up of the learners L2 interlanguage system is different" (Verspoor, Lowie, Chan, & Vahtrick, 2017, 20-21). In other words, learners at different levels of proficiency tend to deploy certain syntactic structures to produce

complex language (Ortega, 2015), and for low to intermediate learners, an essential grammatical structure to create this complexity is subordination (Bi & Jiang, 2020). The quality of narrative essays, however, was predicted by CN/C which is in line with previous studies investigating the relationship between phrasal measures and quality or development of writing (Bi & Jiang, 2020; Casal & Lee, 2019; Li, 2015; Lu, 2011). These studies showed that CN/C could be used to measure the quality of writing, regardless of learners' proficiency level. Due to the fact that the argumentative genre demands that writers hold a viewpoint and produce complicated relationships among ideas to support it, writers attempt to make these relationships by including complex sentence structures. To express more complex relationships among ideas, argue and provide evidence, writers use more main and subordinate clauses to put ideas together in a single T-unit, which results in more C/T and longer T-units. The narrative genre, however, does not demand such syntactically complex sentences (Beers & Nagy, 2011; Jagaiah et al., 2020).

Concerning LD and its contribution to writing quality, our findings are consistent with previous studies by González (2017) and Yu (2009) who reported positive correlations between LD and writing quality. In the present study, the simple linear regressions indicated that LD could explain 64 percent of the variance in the narrative and 59 percent in the argumentative writing, suggesting LD could contribute to writing quality, regardless of the genre.

Conclusion

Despite the wide range of studies on LD, no study has compared LD in L1 and L2 writings of the same writers in argumentative and narrative genres; therefore, in the present study we addressed this issue. Results indicated that participants use more varied words in their L1 argumentative and narrative writings than their L2. Findings of the present study evidently indicate a significant relationship exists between SC and L2 writing quality. This study has demonstrated that specific SC indices correlate with L2 writing quality in argumentative and narrative genres suggesting that genre has an impact on learners' syntactic choices. In terms of LD, results revealed that LD is positively related to L2 writing quality in argumentative and narrative genres. Therefore, it can be argued that the relationship between LD and overall quality rating of L2 writing is not genre-dependent. In general, the findings of the present study suggest that writing instructors should instruct L2 writers to use a variety of words to express their opinions and include more syntactically complex structures in their compositions.

A few limitations exist in this study. First, similar to Casal and Lee (2019) and González (2017) we did not control the topic; since we used a within-subjects design, the same group of participants wrote on different topics in their L1 and L2 to avoid repetition of ideas or simply translation from either L1 to L2 or vice versa. Future studies could control this variable and include participants from matched groups to control the topic and avoid order effects while they write either in their L1 (e.g., Persian) or L2 (i.e., English). In addition, this study was limited to examination of the contribution of SC and LD to writing quality at pre-intermediate and intermediate levels of proficiency. Therefore, future research can examine this relationship at other levels of proficiency so as to ascertain whether the similar results will be obtained.

References

Allan, D. (2004). *Oxford Placement Test1.* Oxford University Press.

Beers, S. F., & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? . *Reading and Writing,* 22(2), 185-200. http://dx.doi.org/10.1007/s11145-007-9107-5

Beers, S. F., & Nagy, W. E. (2011). Writing development in four genres from grades three to seven: syntactic complexity and genre differentiation. *Reading and Writing*, 24(2), 183-202. http://dx.doi.org/10.1007/s11145-010-9264-9

Berman, R. A. (2008). The psycholinguistics of developing text construction. *Journal of Child Language,* 35(4), 735-771. http://dx.doi.org/10.1017/S0305000908008787

Berman, R. A., & Nir-Sagiv, B. (2007). Comparing Narrative and Expository Text Construction Across Adolescence: A Developmental Paradox. *Discourse Processes,* 43(2), 79-120. http://dx.doi.org/10.1080/01638530709336894

Bestgen, Y. (2017). Beyond single-word measures: L2 writing assessment, lexical richness and formulaic competence. *System,* 69, 65-78. http://dx.doi.org/10.1016/j.system.2017.08.004

Bi, P., & Jiang, J. (2020). Syntactic complexity in assessing young adolescent EFL learners' writings: Synactic elaboration and diversity. *System,* 91(1022), 1-10. http://dx.doi.org/10.1016/j.system.2020.102248

Biber, D., Gray, B., & Poonpon, K. (2013). Pay attention to the phrasal structures: Going beyond T-units—A response to WeiWei Yang. *TESOL Quarterly,* 47(1), 192-201. http://dx.doi.org/10.1002/tesq.84

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing,* 26, 42–65. http://dx.doi.org/10.1016/j.jslw.2014.09.005

Casal, J. E., & Lee, J. J. (2019). Syntactic complexity and writing quality in assessed first-year L2 writing. *Journal of Second Language Writing,* 44, 51-62.

Chen, Y.-c. (2020). Assessing the lexical richness of figurative expressions in Taiwanese EFL learners' writing. *Assessing Writing,* 43, 1-12. http://dx.doi.org/10.1016/j.asw.2019.06.007

Choi, Y. H. (2017). Roles of receptive and productive vocabulary knowledge in L2 writing through the Mediation of L2 reading ability. *English Teaching,* 72(1), 3-24. http://dx.doi.org/10.2139/ssrn.4593605

Chon, Y. V., & Lee, S. W. (2015). L2 collocational knowledge by congruency and writing Proficiency of EFL university learners. *English Language and Literature Teaching,* 21(1), 49-74. http://dx.doi.org/10.17250/khisli.33..201609.006

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition,* 11(4), 367-383.

Crossley, S. A., & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing,* 18(2), 119-135. http://dx.doi.org/10.1016/j.jslw.2009.02.002

Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading,* 35(2), 115-135. http://dx.doi.org/10.1111/j.1467-9817.2010.01449.x

Crowhurst, M. (1980). Syntactic Complexity and Teachers' Quality Ratings of Narrations and Arguments. *National Council of Teachers of English,* 14, 223-231.

Daller, H., van Hout, R., & Treffers‐Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics,* 24(2), 197–222. http://dx.doi.org/10.1093/applin/24.2.197

Durán, P., Malvern, D., Richards, B., & Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics,* 25(2), 220-242. http://dx.doi.org/10.1093/applin/25.2.220

Eckstein, G., & Ferris, D. (2017). Comparing L1 and L2 texts and writers in first-year composition. *TESOL Quarterly,* 52(1), 137-162. http://dx.doi.org/10.1002/tesq.376

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics,* 30(4), 474-509. http://dx.doi.org/10.1093/applin/amp042

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford University Press.

Engber, C. A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing,* 4(2), 139-155.

González, M. C. (2017). The Contribution of Lexical Diversity to College-Level Writing. *TESOL Journal,* 8(4), 899-919. http://dx.doi.org/10.1002/tesj.342

Ha, H.S. (2019). Lexical Richness in EFL Undergraduate s' Academic Writing. *English Teaching,* 74(3), 3-28. http://dx.doi.org/10.15858/engtea.74.3.201909.3

Halliday, M., & Hasan, R. (1985). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Oxford University Press.

Hammerly, H. (1991). *Fluency and Accuracy : Toward Balance in Language Teaching and Learning*. Multilingual Matters.

Housen, A., & Kuiken, F. (2009). Complexity, Accuracy, and Fluency in Second Language Acquisition. *Applied Linguistics,* 30(4), 461-473. http://dx.doi.org/10.1093/applin/amp048

Housen, A., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 Performance and Proficiency: Complexity, accuracy and fluency in SLA*. John Benjamins.

Jagaiah, T., Olinghouse, N. G., & Kearns, D. M. (2020). Syntactic coomplexity measure: variation by genre, grade-level, students' writing ability, and writing quality. *Reading and Writing,* 33(3), 2577-2638. http://dx.doi.org/10.1007/s11145-020-10057-x

Khushik, G. A., & Huhta, A. (2019). nvestigating syntactic complexity in EFL learners' writing across common European framework of reference levels A1, A2, and B1. *Applied Linguistics,* 41(4), 506-532. http://dx.doi.org/10.1093/applin/amy064

Kim, J.-y. (2014). Predicting L2 Writing Proficiency Using Linguistic Complexity Measures: A Corpus-Based Study. *English Teaching,* 69(4), 27-51. http://dx.doi.org/10.15858/engtea.69.4.201412.27

Kim, M., & Crossley, S. A. (2018). Modeling second language writing quality: A structural equation investigation of lexical, syntactic, and cohesive features in source-based and independent writing. *Assessing Writing,* 37, 39-56. http://dx.doi.org/10.1016/j.asw.2018.03.002

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. http://dx.doi.org/10.1111/modl.12468

Lahuerta Martínez, A. C. (2018). Analysis of syntactic complexity in secondary education EFL writers at different proficiency levels. *Assessing Writing, 35,* 1-11. http://dx.doi.org/10.1016/j.asw.2017.11.002

Li, H. (2015). Relationship between measures of syntactic complexity and judgments of EFL writing quality. In Proceedings of 2015 Youth Academic Forum on Linguistics, Literature, Translation and Culture (pp. 216–222). American Scholars Press.

Lieven, E. (1978). 'Conversations between mothers and young children: individual differences and their possible implication for the study of child language learning'. *The Development of Communication*, 173-187.

Linnarud, Moira. (1986). *Lexis in composition*. CWK Gleerup.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics,* 15(4), 474-496. http://dx.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. http://dx.doi.org/10.5054/tq.2011.240859

Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal,* 96(2), 190-208. http://dx.doi.org/10.1 111/j. 1540-4781.201 1.01232.

Lu, X. (2017). Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. *Language Testing,* 34(4), 493-511. http://dx.doi.org/10.1177/0265532217710675

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Lawrence Erlbaum Associates.

Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development: Quantification and assessment*. Basingstoke, Palgrave MacMillan.

McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Metrix: Capturing linguistic features of cohesion. *Discourse Processes,* 47(2), 292-330. http://dx.doi.org/10.1080/01638530902959943

Nasseri, M., & Thompson, P. (2021). Lexical density and diversity in dissertation abstracts: Revisiting English L1 vs. L2 text differences. *Assessing Writing,* 47(2), 1-11. http://dx.doi.org/10.1016/j.asw.2020.100511

Nation, P., & Webb, S. (2011). *Researching and analysing vocabulary*. Heinle, Cengage Learning.

Nippold, M. A., Hesketh, L. J., & Duthie, J. K. (2005). Conversational versus expository discourse: A study of syntactic development in children, adolescents, and adults. *Journal of Speech, Language, and Hearing Research,* 48(5), 1048-1064. http://dx.doi.org/10.1044/1092-4388(2005/073)

Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. Applied Linguistics, 30(4), 555–578. http://dx.doi.org/10.1093/applin/amp044

Olinghouse, N. G., & Leaird, J. T. (2009). The relationship between measures of vocabulary and narrative writing quality in second- and fourth-grade students. *Reading and Writing,* 22(5), 545-565. http://dx.doi.org/10.1007/s11145-008-9124-z

Olinghouse, N. G., & Wilson, J. (2013). The Relationship between Vocabulary and Writing Quality in Three Genres. *Reading and Writing: An Interdisciplinary Journal,* 26(1), 45-65. http://dx.doi.org/10.1007/s11145-012-9392-5

Ortega. (2003). Syntactic Complexity Measures and their Relationship to L2 Proficiency: A Research Synthesis of College‐level L2 Writing. *Applied Linguistics,* 24(4), 492–518. http://dx.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 (Supple1), 82-94. http://dx.doi.org/10.1016/j.jslw.2015.06.008

Plonsky, L., & Ghanbar, H. (2018). Multiple Regression in L2 Research: A Methodological Synthesis and Guide to Interpreting R^2 Values. *The Modern Language Journal,* 102(4), 713-731. http://dx.doi.org/10.1111/modl.12509

Qin, W., & Uccelli, P. (2016). Same language, different functions: A cross-genre analysis of Chinese EFL learners' writing performance. *Journal of Second Language Writing,* 33(1), 3-17. http://dx.doi.org/10.1016/j.jslw.2016.06.001

Ravid, D., & Berman, R. A. (2010). Developing noun phrase complexity at school age: A text-embedded cross-linguistic analysis. *First Language,* 30(1), 3-26. http://dx.doi.org/10.1177/0142723709350531

Read, J. (2000). *Assessing vocabulary*: Oxford University Press.

Sawilowsky, S. S. (2009). New Effect Size Rules of Thumb. *Journal of Modern Applied Statistical Methods,* 8(2), 597-599. http://dx.doi.org/10.22237/jmasm/1257035100

Schultz, J. M. (1991). Mapping and Cognitive Development in the Teaching of Foreign Language Writing. *American Association of Teachers of French,* 64(6), 978-988.

Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly,* 27(4), 657-677. http://dx.doi.org/10.2307/3587400

Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford University Press.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Allyn and Bacon.

Taguchi, N., Crawford, W., & Wetzel, D. Z. (2013). What Linguistic Features Are Indicative of Writing Quality? A Case of Argumentative Essays in a College Composition Program. *TESOL Quarterly,* 47(2), 420-430. http://dx.doi.org/10.4324/9781003177579

Verspoor, M., Lowie, W., Chan, H. P., & Vahtrick, L. (2017). Linguistic complexity in second language development: variability and variation at advanced stages. *Recherches en didactique des langues et des cultures,* 14(1), 1-27. http://dx.doi.org/10.4000/rdlc.1450

Vögelin, C., Jansen, T., Keller, S. D., Machts, N., & Möller, J. (2019). The influence of lexical features on teacher judgements of ESL argumentative essays. *Assessing Writing,* 39(7), 50-63. http://dx.doi.org/10.1016/j.asw.2018.12.003

Wang, S., & Slater, T. (2016). Syntactic complexity of EFL Chinese students' writing. *English Language and Literature Studies,* 6(1), 81-86. http://dx.doi.org/10.5539/ells.v6n1p81

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

Wu, S., Dixon, L. Q., Sun, H., & Zhang, P. (2019). Breadth or depth: the role of vocabulary in Chinese English-Language beginning writers' development. *International Journal of Bilingual Education and Bilingualism,* 24(9). http://dx.doi.org/10.1080/13670050.2019.1572066

Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing,* 28 (S1), 53-67. http://dx.doi.org/10.1016/j.jslw.2015.02.002

Yu, G. (2009). Lexical Diversity in Writing and Speaking Task Performances. *Applied Linguistics,* 31(2), 236–259. http://dx.doi.org/10.1093/applin/amp024