THE JOURNAL OF TEACHING ENGLISH FOR SPECIFIC AND ACADEMIC PURPOSES
Vol. 13, Nº 2, 2025, pp. 315–330

UDC: (811+37.011.2) : 378.22 https://doi.org/10.22190/JTESAP250212024G
Review research paper

STUDY ON SYNTACTIC AND LEXICAL DEVELOPMENT THROUGH MASTER'S THESES

Julio Cesar Gomez

Universidad Externado de Colombia, Bogota, Colombia

ORCID iD: Julio Cesar Gomez https://orcid.org/0000-0002-0626-0669

Abstract. *This article explores the extent to which syntactic complexity and lexical diversity develop during the process of writing a master's thesis in an English language teaching program. A total of 21 students belonging to this program were asked to share drafts of their theses. Using the Wilcoxon signed-ranked test analysis in measurements of syntactic complexity for the whole group, the results revealed that there was a statistically significant difference. The two measurements were the mean number of dependent clauses per T-unit and sentence syntax similarity. In the intergroup analysis, the B1 level subgroup had significant gains in the mean number of modifiers per noun phrase. Regarding lexical diversity, the measurement lexical diversity, type-token ratio and content word lemmas showed statistically significant difference for the whole group and for the B1 level subgroup. The findings further indicate that the development of writing in syntactic terms occurs in measurements that involve more complex clauses and noun phrases. Regarding lexical diversity, the range of content words increased compared to other measurements. These findings reveal the limited extent to which features of syntax and lexicon may change as a result of writing a complex academic text such as master's thesis.*

Key words: *writing development, ELT master's thesis, teacher education, syntactic complexity, lexical diversity.*

1. INTRODUCTION

Writing development is usually framed within the context of a genre that could be an academic essay, a research article, or a dissertation. This is also the case in the master's thesis genre where research is reported. Second language (L2) learners face various challenges when writing theses. Paltridge and Starfield (2007) found that for L2 doctoral students, which may as well apply to master's L2 students, face psycho-affective/emotional challenges that include lack of confidence to write in English; behavioral issues related to the absence of writing routines and understanding writing as a process; rhetorical issues that involve the understanding of the characteristics of writing in the thesis genre and features such as metadiscourse; moderating claims and developing a voice; and social issues related to feelings of isolation during the writing task. In addition to all these, L2 students may not always possess the level of English proficiency and writing skills

Submitted February 12th, 2025, accepted for publication April 7th, 2025
Corresponding author: Julio Cesar Gomez, Universidad Externado de Colombia, Bogotá, Colombia
E-mail: julio.gomez@uexternado.edu.co

required to undertake complex writing tasks (Cadman, 1997). Therefore, a master's thesis becomes a type of assignment that may contribute to writing development, especially when it involves continuous feedback. Referring to Gomez' (2014) study on the perceptions of students and graduates from seven English Language Teaching (ELT) programs in Colombia about their writing development through master's theses, the current study examines three areas that participants in the aforementioned study improved significantly. It also explores how features of students' writing, such as syntactic complexity, lexical diversity, and cohesion, develop through the process of writing their theses in an ELT master's program. The study further compares the initial and final drafts of the components of 21 L2 students' theses on measurements based on existing literature for each of these features.

2. Literature Review

2.1. Writing Development as L2 Language Development

Since writing development is closely linked to language development, it is important to address syntactic complexity as a feature of L2 writing that can and should be measured to account for this growth. Lu (2011) defines syntactic complexity as the "the range of syntactic structures that are produced and the degree of sophistication of such structures" (p. 36). In other words, how the writer's syntactic repertoire expands is a reflection of L2 writing complexity (Ortega, 2003). Therefore, it is necessary to define what aspects of this syntactic complexity are to be considered in the current study. Various studies on the exploration of L2 syntactic complexity in relation to language proficiency, particularly two comprehensive reviews of studies (Wolfe-Quintero et al., 1998; Ortega, 2003), have shed some light on the lack of consistency and also the variety of factors each of these studies have considered. Based on the findings of these reviews, Lu (2011) identified 14 syntactic complexity measures of identifying language development. The measures were grouped into five categories: length of production, sentence complexity, subordination, coordination, and particular structures. Crossley and McNamara (2014) also explored the syntactic complexity of L2 writing while focusing on essays written by L2 learners and the analysis of 11 indices that measured "syntactic variety, syntactic transformations (e.g., negations and questions), syntactic embeddings, incidence of phrase types, and phrase length" (p. 70) using the computational tool Coh-Metrix. The aim of this study was to identify the syntactic features that were predictive of writing quality in human judgments compared to the analysis of the changes in syntactic complexity with computational indices. The findings of the computational analysis point out to significant changes between the first and third (last) essays in the number of modifiers per noun phrase, number of words before the main verb, and reduced syntactic similarity. Lorenzo and Gutierrez (2014) traced the development of complex syntax through 244 historical narratives from 9th to 12th graders. The study of Cognitive Academic Language Proficiency (CALP) structures development focused on verb tenses and sentence-level adverbial subordination and nominalization. The analysis of complexity was carried out using the Synlex software. The findings underscore constant change in complex nominals and sentence embeddedness, as well as a larger use of verbal tenses, all of which are important features of written L2 CALP. Yin et al. (2021) also explored the syntactic complexity in L2 writing, with their main focus being differences between emerging and expert Chinese research article writers. A sample of 60 research articles were analyzed as a whole, and each part-genre focused on

the measures of length of production unit, amount of subordination, amount of coordination, phrasal sophistication, and sentence complexity. The findings highlighted differences between expert and emerging writers in terms of the latter getting lower scores in overall sentence complexity, subordination, and verb phrases, but higher scores in coordination measures in some genre parts. Gaps in syntactic complexity between the two groups revealed differences in writing expertise related to choices and consistency in the use of complex syntactic structures in the different sections of the research articles.

Another feature of L2 language development is lexicon development. Lexicon development can also be considered as an indicator of language and writing development. The extent to which L2 writers possess an extensive and at times specialized vocabulary can be an indicator of their L2 proficiency. Kalantari and Gholami (2017) explored the development of lexical density, diversity, and sophistication as reflected in 50 essays written by five Iranian EFL learners over a period of six months in a TOEFL writing class from a Dynamic Systems Theory perspective. Their findings point out significant changes in two measurements of lexical sophistication: Academic Word Length and Beyond-2000. The remaining measurements of lexical diversity—Uber index, Squared Verb Variation, MLTD and Vocabulary Diversity (Vocd-D), and Lexical Density (LD-LCA and LD-VP)—seemed to have reached a plateau in their development. The improvement seen in lexical sophistication in this study can be attributed to the effect of testing preparation in terms of the types of tasks and feedback given.

In one more study exploring both syntactic complexity and lexical diversity, Mazgutova and Kormos (2015) examined the development of lexical diversity and grammatical complexity features of the academic writing of two groups of L2 learners—one at the undergraduate level and the other at graduate level—taking an EAP course in a British university. The findings reveal that both groups improved in terms of lexical diversity and participants in the undergraduate level enhanced their genre-specific grammatical structures. For lexical diversity, the measurement squared verb variation had significant changes for both groups. Nevertheless, it was the lower proficiency group that had significant changes in all the lexical diversity measures explored. Regarding syntactic complexity, both groups had significant changes in syntactic structure similarity, which reflects the use of a small variety of syntactic structures. However, the lower proficiency group also experienced changes in complex nominals and modifiers per noun phrase. These findings show the impact that intensive and immersive academic English programs have on learners' writing development.

2.2. Framework for the analysis of syntactic and lexical development

The current study uses the framework of analysis of syntactic and lexical development proposed by Mazgutova and Kormos (2015) as reference with some variations in terms of the measurements studied. Their framework took a developmental perspective and adapted Bulté and Housen's (2012) conceptualization of grammatical diversity as syntactic complexity and Jarvis' (2013) work on lexical diversity "as consisting of rarity, volume, variability, evenness, disparity and dispersion" (Mazgutova & Kormos, 2015, p. 5). Considering the variety of facets of grammatical and lexical complexity, the findings of the previous studies and the idea that scholarly writing needs to reflect a degree of language sophistication, the specific measurements that were explored to ascertain the syntactic complexity in the drafts of the various chapters of the theses were: Mean length of T-unit (MLT); mean number of dependent clauses per T-unit (DC/T); complex nominals (CN/C); Mean Number of Modifiers

per noun phrase (SYNNP); Sentence syntax similarity, all combinations, across paragraphs, mean (SYNSTRUTt), and Left Embeddedness, words before main verb, mean (SYNLE). Regarding lexical complexity, the parameters analyzed were Squared Verb Variation (SVV1); Lexical diversity, MTLD, all words (MLTD); Lexical diversity, type-token ratio, and content word lemmas (LDTTRc).

2.3. Computational Software Packages

One of the computational tools most recently developed to identify syntactic complexity as well as other relevant features such as lexical diversity is Coh-Metrix (Graesser et al., 2004; McNamara et al., 2014). This tool has been described extensively in studies (e.g., McCarthy et al., 2006) and implemented to assess the writing quality of argumentative essays written by both ESL and EFL students in state universities (Kyle, 2011), discriminate good from bad writing among L2 college students in China (Hongwei & Liqin, 2013), identify syntactic and lexical features of two advanced L2 writers' essays in Hungary (Wind, 2013), ascertain the syntactic and lexical characteristics of two groups of L2 writers in a British university through their argumentative essays in an English for academic purposes course (Mazgutova & Kormos, 2015), and assess the writing development of Japanese students in two EFL classrooms through repeated writing tasks (Baba & Nitta, 2010).

Synlex (Lu, 2012) is also another computational program that has been developed to identify features of syntactic complexity and lexical complexity. In particular, this program has been used to identify features of lexical complexity in several studies of L2 writing development (Desouky & Desouky, 2018; Kalantari & Gholami, 2017; Lorenzo & Rodriguez, 2014; Mazgutova & Kormos, 2015; Vasylets, et al., 2019; Wind, 2013).

These two computational tools have been widely used in longitudinal studies that have explored L2 writing development (Crossley, 2013). It is important to note that the most common sources of information to conduct these analyses are essays and writing tasks completed by L2 writers. Therefore, it is important to examine how this analysis plays out when the main source is drafts of the different components of theses that the participants in the current study wrote during their two-year program.

2.4. Research on L2 Writing Development in Graduate Level Education

Studies examining L2 writing development at the graduate level have focused on essay writing as the most common type of writing assignment. Zhao (2017) examined differences in the use of conjunctions and logical grammatical metaphors between native speakers and nonnative graduates and scholars on the basis of essays they had written. Mazgutova and Kormos (2015) explored the development of lexical and grammatical features of the academic writing of two groups of L2 learners—one at the undergraduate level and the other at graduate level—taking an EAP course in a British university.

The studies presented so far have explored certain features of writing usually within the time frame of a course or a semester. There are very few longitudinal studies that have explored L2 writing development. One study analyzed the written narratives and oral presentations done by five Chinese English learners, which were repeated tasks carried out four times over a period of six months (Larsen-Freeman, 2006). This study focused on fluency, grammatical complexity, accuracy, and vocabulary complexity as the four main indices of determining L2 writing development. The findings reveal that the participants improved in terms of accuracy and fluency. Their lexicon and grammar also

became more complex. However, each participant followed a different path for this development. Notably, this study did not use computational software to conduct the analysis or any of the measurements determined for the analysis of features of syntactic complexity or lexical diversity in the current study. Baba and Nitta (2010) also used repeated tasks involving written compositions of 46 L2 college students over a period of one year in Japan to determine L2 writing development. They used three categories of measures to study quality in the development of L2 writing, which were fluency, lexical complexity, and grammatical complexity. The measurement defined for the analysis of fluency was Latent Semantic Analysis, for lexical complexity was word frequency values from CELEX and MLTD, and for grammatical complexity was average sentence length and Sentence similarity, all sentences across paragraphs (STRUT). The results revealed that grammatical complexity and lexical diversity improved over the year period, but not writing fluency. The individual paths of development were also quite varied for the participants in this study.

Studies related to the master's thesis writing have explored aspects of the writing process of L2 writers, and they include studies that have examined difficulties in understanding demands of their audience and genre in lexical terms (Shaw,1991); difficulties with features of language such as coherence, organization, and grammatical correctness (Powers, 1994); and those that have focused on challenges in writing certain sections, such as the discussion section (Bitchener & Basturkmen, 2006), the evaluation of sources in the literature review (Xie, 2016), and the strategies to write their theses (Khozaei Ravari & Tan, 2019).

The only study that has focused on the writing development of home L2 students as they wrote their master's theses is Gomez' (2014) study, which explored perceptions of students, graduates, and supervisors from seven ELT master's programs in Colombia in relation to writing development through theses. Data were gathered through surveys and in-depth interviews with some stakeholders and document analysis was conducted on some samples of theses. The findings pointed out to writing development in terms of perceived improvement in features, such as vocabulary, grammar, syntax, and text organization based on students' initial writing level.

Building on the recurrent findings that highlight the uniqueness of writing development for each individual, the current study goes beyond the self-reported data of the participants' perceived writing development and the analysis of the two drafts from three students in Gomez' study (2014) to consider drafts from 21 master's students in Colombia and determine the features of syntactic complexity and lexical diversity that are in each component of their final theses. To achieve this, the study uses computational software programs. It is important to highlight that the experience of these students as home L2 novice writers is worth exploring since it reflects the challenges and opportunities of many other ELT master's students in EFL contexts.

3. Research Question

The current study seeks to answer the following research question: What features of syntactic complexity and lexical diversity do students develop as they write their ELT master's theses?

4. METHODS

4.1. Research Methodology

A case study methodology was chosen for the current study since it allows the integration of quantitative data "which facilitates reaching a holistic understanding of the phenomenon being studied" (Baxter Jack, 2008, p. 554). Data from several participants are merged during the analysis to build what would be a more complete picture of the phenomenon. Braine (2002) pointed out the value of this type of study in exploring linguistic growth since "case studies are also descriptive, dynamic, and rely upon naturally occurring data" (p. 66).

4.2. Participants

The 21 participants in this study were in-service English teachers taking a master's program in Education with an emphasis in English didactics from the Universidad Externado de Colombia. These students belonged to four different cohorts that were part of the program from 2014 to 2018. Each cohort took two years to complete their master's program and write their theses. In this master's program, there was no English language level or research proposal requirement upon entry. Students in this program took variety of courses during the four semesters and gradually wrote the sections of the theses submitting several drafts each semester in the research class. All participants were native speakers of Spanish and started learning English as a foreign language in high school and then continued learning in their undergraduate programs and some in English language centers. On average, they had been learning English for 8 to 10 years. They taught English as a foreign language in private and public schools at the elementary and secondary levels, and a few at the college level. The 21 participants were classified by three external reviewers using their first drafts to determine their English writing proficiency while considering the descriptors of the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) (Council of Europe, 2001). This framework has been adopted and adapted by Ministry of Education of Colombia and became standards for foreign language teaching and learning in the country (Ayala & Alvarez, 2005). This placement determined three main subgroups (Table 1)—Independent user (B2), Independent user (B1), and Basic user (A2)—which were later considered for the second stage of the analysis.

Table 1 Participants Profiles

Gender	Female	18
	Male	3
Teaching Context	Public School	4
	Private School	12
	Private University	3
	Public University	1
	Language Institute	1
English Writing Proficiency	Level A2	4
	Level B1	9
	Level B2	8
Cohort	Cohort 1	7
	Cohort 2	2
	Cohort 3	7
	Cohort 4	5

4.3. Instruments

Students' drafts of the different components of their theses were gathered through the four semesters of their program. However, the researcher focused on the first and last drafts of each major component of the theses to conduct the analysis of the different measurements of syntactic complexity and lexical diversity (Table 2) using the two main computational software programs (i.e., Coh-Metrix and Synlex). The first draft was compiled with the first draft of every component each student wrote, while the last draft was the final submission prior to final feedback and editing. In each draft, headings and direct quotes from other sources included in the text were omitted to allow the software programs to focus on the students' own production.

Table 2 Measurements of Syntactic Complexity and Lexical Diversity

	Measurements					
Syntactic Complexity	Mean Length of T-unit (MLT) *	Mean Number of Dependent Clauses per T-unit (DC/T) *	Complex Nominals (CN/C) *	Mean Number of Modifiers per noun phrase (SYNNP) **	Sentence Syntax Similarity, all combinations, across paragraphs, mean (SYNSTRUTt)**	Left Embeddedness, words before main verb, mean (SYNLE) **
Lexical Diversity	Squared Verb Variation (SVV1) *	Lexical Diversity, all words (MLTD)**	Lexical Diversity, type-token ratio, content word lemmas (LDTTRc).**			

*Synlex **Coh-Metrix

4.4. Data Analysis Procedure

Two types of analyses were carried out with the students' first and last drafts of each component of the theses (Table 3) to determine changes. Table 4 presents the composition of the data set, which includes the total number of words in the first and last drafts for each participant, as well as the mean and standard deviation. A computational analysis of specific measurements of syntactic complexity and lexical diversity was conducted using the computerized software programs Coh-Metrix (Graesser et al., 2004) and Synlex (Lu, 2010). The results for every measurement were recorded in an Excel spreadsheet and then transferred to SSPS statistical analysis software. The statistical analysis aimed to determine whether there was development in any of the measurements of syntactic complexity and lexical diversity in the students' writing. Wilk test and determining the normality of the data, the Wilcoxon signed-ranked test was used instead of the *t*-The non-parametric test, Wilcoxon signed-ranked test, was used to identify differences in the Coh-Metrix and Synlex indices for each of the defined measurements for the whole group of participants (N=21) and the three subgroups between the first and last drafts of the selected sections of their theses. It is important to note that despite conducting the Shapiro test because of the number of subjects in the sample, particularly for the inter group analysis, the results of the t-test were very similar to the results of the Wilcoxon signed-ranked test presented here.

Table 3 Drafts Gathered for Every Section of the Thesis for Each Participant

Section of Thesis	Draft	Draft	Draft	Draft	Draft	Draft
Introduction	D1*	D2	D3	D7	D13	D16**
Setting & Rationale	D2*	D3	D7	D13	D16**	
Statement of the Problem	D2*	D3	D7	D13	D16**	
Construct 1	D4*	D5	D7	D13	D16**	
Construct 2	D5*	D7	D13	D16**		
Construct 3	D5*	D7	D13	D16**		
Research Design	D10*	D11	D12	D13	D16**	
Data Analysis Procedure	D14*	D15	D16**			
Conclusion	D14*	D15	D16**			

*First draft analyzed ** Last draft analyzed

Table 4 Mean Number of Words in Initial and Final Drafts for Each Participant

Participant	Total words draft 1	Total words draft 2
1	9.636	8.073
2	8.073	7.520
3	6.237	6.523
4	7.649	7.479
5	6.910	8.445
6	6.382	8.158
7	5.424	7.174
8	10.039	12.056
9	10.360	11.541
10	8.083	10.630
11	9.615	9.441
12	6.602	9.781
13	9.165	9.845
14	9.982	9.986
15	8.192	9.990
16	6.561	7.454
17	8.994	9.670
18	7.166	9.994
19	9.642	9.999
20	8.024	8.914
21	8.567	9.024
Mean	8.157	9.128
SD	1453.04	1453.61

5. RESULTS

5.1. Syntactic Complexity

The results of a Wilcoxon signed-ranked test for the whole group of participants (N=21) show that there were two measurements of syntactic complexity that showed significant difference between the first and last drafts of the theses. The first measurement is *mean number of dependent clauses per T-unit* (DC/T), which showed a significant difference in the indices for the last drafts (*Mdn* = 0.31) were statistically significantly higher than

the indices for the first drafts (*Mdn* = 0.29), $T = 172$, $z = -1.99$, $p < 0.046$, with a large effect size, $r = 0.30$. The second measurement was *Sentence syntax similarity, all combinations, across paragraphs, mean* (SSS) and the difference was significant between the last drafts (*Mdn* =0.073) and the first drafts (*Mdn* = 0.069), $T = 125$, $z = -2.954$, $p < 0.003$, with a large effect size, $r = 0.45$. There were no significant changes in the remaining measurements of syntactic complexity considered, i.e., MLT, complex nominals, mean number of modifiers, and left embeddedness. Table 5 shows the descriptive statistics on how the mean number of dependent clauses per T-unit and the sentence syntax similarity mean increased from the first drafts to the last drafts.

Table 5 Descriptive Statistics for All Measurements of Syntactic Complexity

Measurements	Drafts 1		Drafts 2	
	Mean	*SD*	*Mean*	*SD*
Mean Length of T-unit (MLT)	70.2457	7.64091	73.8133	7.76525
Mean Number of Dependent Clauses per T-unit (DC/T)	0.2886	0.05313	0.3138	0.03681
Complex Nominals (CN/C)	1.6495	0.07180	1.6757	0.08908
Mean Number of Modifiers per Noun Phrase (SYNNP)	0.9062	0.04533	0.9062	0.04353
Sentence Syntax Similarity, all Combinations, Across Paragraphs, Mean (SYNSTRUTt)	0.0691	0.00574	0.0728	0.00661
Left Embeddedness, Words Before Main Verb, Mean (SYNLE)	4.4476	0.54566	4.3510	0.51295

Table 6 shows the effect sizes for the measurements that showed significant change based on the results of the Wilcoxon signed-ranked test for all the measurements of syntactic complexity for the whole group.

Table 6 Effect Sizes for All Measurements of Syntactic Complexity

Measurements	*N*	*Z*	*p*	*r*
Mean Length of T-unit (MLT)	21	−1.790	0.073	0.27
Mean Number of Dependent Clauses per T-unit (DC/T)	21	−1.991	0.046	0.30
Complex Nominals (CN/C)	21	−.915	0.360	0.14
Mean Number of Modifiers per Noun Phrase (SYNNP)	21	−.035	0.972	0.005
Sentence Syntax Similarity, all Combinations, Across Paragraphs, Mean (SYNSTRUTt)	21	−2.954	0.003	0.45
Left Embeddedness, Words Before Main Verb, Mean (SYNLE)	21	−.710	0.478	0.11

Out of the three subgroups of participants, one subgroup showed significant changes in particular measurements of syntactic complexity. The subgroup of participants (N=9) in the B1 level of writing proficiency (CEFR) showed a significant difference in the measurement *Mean Number of Modifiers per noun phrase (SYNNP)* in the indices for the first drafts (*Mdn* = 0.88) and for the last drafts (*Mdn* = 0.89), $T = 39$, $z = -1.981$, and $p < 0.048$, with a large effect size, $r = 0.47$. Table 7 shows the effect size for all the three groups of language proficiency in the measurements of syntactic complexity.

Table 7 Effect Sizes for Each Group for All Measurements of Syntactic Complexity

Measurements	B2				B1				A2			
	N	*Z*	*p*	*r*	*N*	*Z*	*p*	*r*	*N*	*Z*	*p*	*r*
Mean Length of T-unit (MLT)	8	-1.540	0.123	0.38	9	-.415	0.678	0.09	4	-1.461	0.144	0.51
Mean Number of Dependent Clauses per T-unit (DC/T)	8	-1.689	0.091	0.42	9	-.299	0.765	0.07	4	-1.841	0.066	0.65
Complex Nominals (CN/C)	8	-1.103	0.270	0.27	9	-.296	0.767	0.06	4	.000	1.000	0
Mean Number of Modifiers per noun phrase (SYNNP)	8	-1.294	0.196	0.32	9	-1.981	0.048	0.47	4	-1.105	0.269	0.39
Sentence Syntax Similarity, all combinations, across paragraphs, mean (SYNSTRUTt)	8	-1.682	0.093	0.42	9	-1.778	0.075	0.41	4	-1.604	0.109	0.56
Left Embeddedness, words before main verb, mean (SYNLE)	8	-.593	0.553	0.14	9	-.889	0.374	0.20	4	-.365	0.715	0.12

5.2. Lexical Diversity

The only measurement of lexical diversity that showed significant change for the whole group was the *Lexical diversity, type-token ratio, content word lemmas* (LDTTRc). This difference was significant between the first drafts (*Mdn* = 0.334) and the last drafts (*Mdn* = 0.329), $n = 21$, $T = 24$, $z = -3.181$, and $p < 0.001$, with a large effect size, $r = 0.49$. The measurements of squared verb variation and lexical diversity, measure of textual lexical diversity, did not show significant differences for all words. The descriptive statistics in Table 8 show how the means of lexical diversity, type-token ratio, content word lemma decreased from the first drafts to the last drafts.

Table 8 Descriptive Statistics for All Measurements of Lexical Diversity

Measurements	Drafts 1		Drafts 2	
	Mean	*SD*	*Mean*	*SD*
Squared Verb Variation (SVV1)	63.4324	12.85818	64.2333	12.13039
Lexical Diversity, all words (MLTD)	76.8470	8.42018	77.9957	8.86303
Lexical Diversity, type-token ratio, content word lemmas (LDTTRc)	0.3410	0.02951	0.3236	0.02753

Table 9 shows the effect sizes for the measurements that showed significant change in the results of Wilcoxon signed-ranked test for all the measurements of lexical diversity for the whole group.

Table 9 Effect Sizes for All Measurements of Lexical Diversity

Measurements	*Z*	*p*	*r*
Squared Verb Variation (SVV1)	-.261	0.794	0.04
Lexical Diversity, all words (MLTD)	-.921	0.357	0.14
Lexical Diversity, type-token ratio, content word lemmas (LDTTRc)	-3.181	0.001	0.49

Similarly, the only subgroup of participants that showed a significant difference in a measurement of lexical diversity was the B1 group. The measurement was *Lexical diversity, type-token ratio, content word lemmas* (LDTTRc), and this subgroup showed a significant difference in the indices for the first drafts (*Mdn* = 0.35, $n = 9$) and for the last drafts (*Mdn* = 0.329, $n = 9$), $T = 4$, $z = -2.192$, and $p < 0.028$, with a large effect size, $r = 0.52$. The effect

size for every measurement of lexical diversity for each group of students is presented in Table 10.

Table 10 Effect Sizes for Each Group for All Measurements of Lexical Diversity

Measurements	B2			B1			A2		
	Z	*p*	*r*	*Z*	*p*	*r*	*Z*	*p*	*r*
Squared Verb Variation (SVV1)	.000	1.000	0	-.889	0.374	0.20	-1.826	0.068	0.64
Lexical Diversity, all words (MLTD)	-1.400	0.161	0.35	-.652	0.515	0.15	-1.095	0.273	0.38
Lexical Diversity, type-token ratio, content word lemmas (LDTTRc)	-1.540	0.123	0.38	-2.192	0.028	0.52	-1.826	0.068	0.64

6. Discussion

Regarding syntactic complexity, two measurements had significant changes for the whole group. The number of dependent clauses per T-unit that these writers used in the last drafts of their theses increased compared to the first drafts. The significant change in this measurement corresponds with the findings in the study conducted by Lorenzo and Gutierrez (2014), which also demonstrated higher scores in sentence embeddedness comparing 9th graders' to 12th graders' historical narratives writing. Although this change was not significant in that study, it points out a tendency toward the production of more complex sentence structures through time. Yin et al. (2021) found that subordination measures in general, including the CD/T measure explored in this study, had lower values for emerging IP writers publishing in an entry level journal compared to expert writers publishing in a top journal. The difference between these results and the results of the current study may point to a different stage in the writing development of L2 graduate students in the current study, their purpose for writing, and even the editing process that Yin et al.'s participants went through for publication.

Similarly, there was a change in sentence syntax toward more similar structures comparing the two drafts of each of the components of the theses from the participants. These findings correspond with Azadnia, et al.'s (2019) study that identified this measurement as one that reflects a distinctive characteristic of the academic writing of L2 writers in a study that compared their writing to that of L1 writers in their dissertations. This measurement deals with the degree of consistency in the syntactic structures used in the text. Similarly, Mazgutova and Kormos (2015) found that syntactic structures similarities increased in their study instead of decreasing in both lower and higher proficiency groups. This trend somehow reflects a feature of professional writers in certain cultures and genres that would prefer to use a smaller range of grammatical structures in their writing (Baba & Nitta, 2010) or a developmental stage in their L2 writing development. Conversely, this finding contradicts Crossley and McNamara's (2014) study, which found that syntactic similarity reduced in essays produced over the development of a semester writing course. Similarly, Baba and Nitta (2010) found that their participants used a wider range of grammatical structures by the end of their yearlong language course. This disparity with the thesis writing process explored in the current study could be as a result of the intensity and probably focused practice in both the writing course and the English language course compared to the mostly autonomous

process of adjustment and confidence building through the extensive work with various drafts of the different components of thesis documents.

Interestingly, when exploring how each measurement of syntactic complexity behaved according to each subgroup of students classified by language proficiency, a different measurement of syntactic complexity showed significant change. This measurement was related to the number of modifiers per noun phrase, and the subgroup that had these results was the B1group. Although this subgroup represents a sample of the total group of participants, it still reflects a tendency highlighted in Azadnia et al.'s (2019) study in which texts written by L2 writers show a high number of modifiers in their texts. Similarly, Crossley and McNamara's (2014) study found that specifically intermediate L2 learners in their study produced longer noun phrases. However, the findings of the study conducted by Mazgutova and Kormos (2015) found that their lower proficiency group (B2 level in the CEFR scale) had significant gains in the number of modifiers per noun phrase, which was not the case for our B2 participants who had the most advanced language proficiency level in the sample. This discrepancy is underscored by the fact that there was a slight difference between these participants' overall scores and the higher proficiency group's (C1 level in the CEFR scale) scores in Mazgutova and Kormos's (2015) study. The advanced group in the current study did not have significant changes in any of the measurements of syntactic complexity contrary to the study conducted by Mazgutova and Kormos, which found that the syntactic structure similarity changed significantly for this group. Issues related to the English proficiency level of the participants and the intensity of the one-month English for academic purposes course could account for the differences in their results. Although the B1 group in the current study was somehow intermediate between A2 and B2 participants in the sample, it seems to be a group that is developing toward a more proficient level. However, the current study also contradicts other studies that found that more proficient writers tend to use more modifiers (e.g., Guo et al., 2013; McNamara et al., 2010; Parkinson & Musgrave, 2014; Azadnia et al. (2019). In this study, B2 participants did not record a significant improvement in this measurement maybe due to the fact that they were already using a substantial number of modifiers. However, this would not be explicitly explored using the analysis conducted.

In terms of lexical diversity, the only measurement that showed a significant change was related to type-token ratio in content word lemmas for the whole group. All the other measurements did not show any significant change. This was also the case in Kalantari and Gholami's (2017) study, which did not have significant changes in the measurements of lexical diversity such as SVV and MLTD that were also considered in the current study. On the contrary, the MLTD measurement changed significantly for L2 learners in Baba and Nitta's (2010) study. The analysis of the subgroups in these measurements also showed that the B1 group had significant change, and the measurement in the study was lexical diversity, type-token ratio, content word lemmas as well. The change in this measurement can be explained by the fact that the participants may have already mastered a great deal of the grammar structures and could experiment with a variety of new vocabulary in their drafts.

Some limitations of the current study include the extent to which all the measurements of syntax complexity and lexical diversity chosen fully reflected all the features that could have been affected through the thesis writing process cannot be ascertained. Another limitation of the study is the lack of certification in language proficiency of the participants upon arrival at the program, which made it necessary to appoint three external reviewers to facilitate the placement of the participants based on their writing samples. Another limitation is the small

number of participants in the proficiency groups that may not deem these statistical results meaningful to determine no development instead of insufficient power of the statistical test. Despite representing four different cohorts of participants in this program over a period of five years, this number is still not representative of the amount of ELT programs in this context, making these findings transferable only to ELT programs with similar characteristics. The types of studies mainly reported in the theses written by the 21 participants were action research studies that provided certain consistency in the information. This becomes a limitation since these theses did not include other types of studies also conducted in ELT master's programs. Similarly, the fact that these participants' tutor later became the researcher guarantees that they had been exposed to similar writing practices (i.e., drafting, language, and content feedback), which may have limited their exposure to other strategies or ways of writing their theses. One last limitation also related to the tutor is related to the amount of linguistic input that may have made its way to the last drafts, in particular to the drafts of lower proficiency students. However, it is important to highlight that the writing process does not occur in a vacuum, and it is not possible to control for the influence of the feedback, peer feedback and other sources that students may have used along the process.

Future studies should explore how other features of writing, such as coherence and cohesion, or even other features of lexical development such as density or sophistication evolve through the process. The studies can also investigate the gradual changes of every component of the theses through multiple drafts to see where changes start to happen and even complement the research process with qualitative accounts to identify factors that contributed to these changes.

7. Conclusion

The current study aimed at identifying the features of syntactic complexity and lexical diversity that develop through the process of writing a master's thesis in an ELT program in an EFL context. The findings highlight the features of writing that may tend to indicate particular changes in syntactic complexity and lexical diversity for L2 students writing their master's theses. From the L2 thesis writers' perceptions of improvement in grammar and vocabulary in Gomez' (2014) study, it is possible to identify that in the group of 21 master's students in the current study, changes in syntax are related to the level of complexity of the sentences through embedded clauses and the number of modifiers per noun phrases for the intermediate level subgroup. Another aspect of syntax that changed significantly was the use of similar structures, which may suggest less complexity in the syntax in itself and more use of structures that have been gradually mastered through the writing practice. Concerning changes in vocabulary, it was the increase in the use of new content words that stood out for the whole group. The findings further point out to the development of specific features of syntactic complexity that have been consistently seen as indicators of development, especially in longitudinal studies of L2 writers. Concerning the feature of lexical diversity identified, only one measurement was not found to be significantly changing in other studies of L2 writing development. It is worth highlighting that the B1 subgroup showed the most significant changes in syntactic complexity and lexical diversity measurements compared to the A2 and B2 subgroups, which seems to point to a stage of ongoing work on the mastery of the language in the former group or to a plateau stage in the latter.

The characteristics of a master's thesis as an academic undertaking that combines independent and autonomous work with support from supervisors and academic writing specialists make it hard to identify the effect of any given agent on specific syntactical and lexical features. However, the drafting processes involved in this ELT program may provide an opportunity to refine the writing to produce the language required to comply with the requirements of every component of a thesis. Based on the results obtained in this study, it is clear that only a few aspects of syntactic complexity and lexical diversity seem to have changed for this group of participants. Nevertheless, the features of syntactic complexity that changed for the whole group reflect a tendency in advanced L2 writers to produce more complex clauses while adhering to familiar structures in general. The changes in vocabulary focused on the wider range of content words they used in latter drafts.

The implications of this study for ELT education programs include the importance of exposing learners to the diverse models and extensive writing practice that engage L2 emerging academic writers in the exploration and implementation of variety of structures and vocabulary used in academic genres. These practices can contribute to better prepare students to face the challenges of academic writing at the graduate and postgraduate levels. The current study aimed to extend research on L2 writing development that involves the exploration of complex academic texts, such as a master's thesis, to go beyond the realm of essay and narrative writing in graduate writing or abstract writing in postgraduate writing. The identification of aspects of syntax and lexicon that significantly change—as well as those that do not—through to the thesis writing process provides insights on features that require further attention to move to more demanding writing genres such as research articles or doctoral dissertations.

REFERENCES

Ayala Zárate, J., & Álvarez V., J. A. (2005). A perspective of the implications of the common European framework implementation in the colombian socio-cultural context. *Colombian Applied Linguistics Journal*, (7), 7–26. https://doi.org/10.14483/22487085.162

Azadnia, M., Lotfi, A., & Biria, R. (2019). A study of syntactic complexity 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. doi: 10.30486/relp.2018.663453

Baba, K., & Nitta, R. (2010, May). Dynamic effects of task type practice on the Japanese EFL university student's writing: Text analysis with Coh-Metrix. In *Twenty-Third International FLAIRS Conference*.

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report*, *13*(4), 544–559. https://doi.org/10.46743/2160-3715/2008.1573

Bitchener, J., & Basturkmen, H. (2006). Perceptions of the difficulties of postgraduate L2 thesis students writing the discussion section. *Journal of English for Academic Purposes*, *5*(1), 4–18. https://doi.org/10.1016/j.jeap.2005.10.002

Braine, G. (2002). Academic literacy and the nonnative speaker graduate student. *Journal of English for academic purposes*, *1*(1), 59–68. https://doi.org/10.1016/s1475-1585(02)00006-1

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. House, F. Kuiken, & I. Vedder (Eds.). *Dimensions of L2 performance and proficiency: Complexity, accuracy, and fluency in SLA* (pp. 21-46). John Benjamins.

Cadman, K. (1997). Thesis writing for international students: A question of identity. *English for Specific Purposes, 16* (1), 3–14. https://doi.org/10.1016/s0889-4906(96)00029-4

Council of Europe. (2001). *The Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press.

Crossley, S. A. (2013). Advancing research in second language writing through computational tools and machine learning techniques: A research agenda. *Language Teaching*, *46*(2), 256–271. https://doi.org/10.1017/s0261444812000547

Crossley, S. A., & McNamara, D. S. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing*, *26*, 66–79. https://doi.org/10.1016/j.jslw.2014.09.006

Desouky, D., & Desouky, E. S. A. (2018). The effects of task complexity on FL learners' online writing behaviors and the correlated underlying cognitive processes. *CDELT Occasional Papers in the Development of English Education*, *64*(1), 33–68

Ferguson, T. (2009). The 'write'skills and more: A thesis writing group for doctoral students. *Journal of Geography in Higher Education*, *33*(2), 285–297. https://doi.org/10.1080/03098260902734968

Gomez, J. C. (2014). *The development of writing skills in master's level English as a foreign language teacher education programs: insight into the process and perceptions from stakeholders in Colombian universities* (Unpublished Doctoral dissertation) University of Alabama, Tuscaloosa, Alabama.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior research methods, instruments, & computers*, *36*(2), 193–202. https://doi.org/10.3758/bf03195564

Guo, L., Crossley, S. A., & McNamara, D. S. (2013). Predicting human judgments of essay quality in both integrated and independent second language writing samples: A comparison study. *Assessing Writing*, *18*(3), 218–238. https://doi.org/10.1016/j.asw.2013.05.002

Halliday, M. A., & Hasan, R. (1976). *Cohesion in English*. Longman Group Ltd. Hongwei, W., & Liqin, Y. A. (2013). A computational analysis of textual features and L2 writing proficiency. *International Journal of Academic Research in Progressive Education and Development, 2*(4), 170–185. https://doi.org/10.6007/ijarped/v2-i4/396

Jarvis, S. (2013). Defining and measuring lexical diversity. In S. Jarvis & M. Daller (Eds.), *Vocabulary knowledge: Human ratings and automated measures* (pp. 13-45). John Benjamins.

Kalantari, R. & Gholami, J. (2017). Lexical Complexity Development from Dynamic Systems Theory Perspective: Lexical Density, Diversity, and Sophistication. *International Journal of Instruction, 10*(4), 1–18. https://doi.org/10.12973/iji.2017.1041a

Khozaei Ravari, Z., & Tan, K. E. (2019). A qualitative investigation of strategies and experiences of non-native students writing master's theses. *Journal of Applied Research in Higher Education, 11*(2), 310-324. https://doi.org/10.1108/jarhe-07-2018-0120

Kyle, K. (2011). *Objective measures of writing quality*. [Unpublished doctoral dissertation]. Colorado State University.

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

Larcombe, W., McCosker, A., & O'Loughlin, K. (2007). Supporting education PhD and DEd students to become confident academic writers: An evaluation of thesis writers' circles. *Journal of University Teaching and Learning Practice*, *4*(1), 61–71. https://doi.org/10.53761/1.4.1.6

Li, L. Y., & Vandermensbrugghe, J. (2011). Supporting the thesis writing process of international research students through an ongoing writing group. *Innovations in Education and Teaching International*, *48*(2), 195–205. https://doi.org/10.1080/14703297.2011.564014

Lorenzo, F., & Rodríguez, L. (2014). Onset and expansion of L2 cognitive academic language proficiency in bilingual settings: CALP in CLIL. *System*, *47*, 64–72. https://doi.org/10.1016/j.system.2014.09.016

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics, 15*(4), 474–496. https://doi.org/10.1075/ijcl.15.4.02lu

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *Tesol Quarterly*, *45*(1), 36–62. https://doi.org/10.5054/tq.2011.240859

Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, *96*(2), 190–208. https://doi.org/10.1111/j.1540-4781.2011.01232_1.x

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, *29*, 3–15. https://doi.org/10.1016/j.jslw.2015.06.004

McCarthy, P. M., Lewis, G. A., Dufty, D. F., & McNamara, D. S. (2006, May). Analyzing Writing Styles with Coh-Metrix. In *FLAIRS Conference* (pp. 764–769).

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic Features of Writing Quality. *Written Communication*, *27*(1), 57–86. https://doi.org/10.1177/0741088309351547

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

Odena, O., & Burgess, H. (2017). How doctoral students and graduates describe facilitating experiences and strategies for their thesis writing learning process: A qualitative approach. *Studies in Higher Education*, *42*(3), 572–590. https://doi.org/10.1080/03075079.2015.1063598

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24,* 492–518. https://doi.org/10.1093/applin/24.4.492

Paltridge, B., & Starfield, S. (2007). *Thesis and dissertation writing in a second language: A handbook for supervisors*. Routledge. https://doi.org/10.4324/9780203960813

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*, 48–59. https://doi.org/10.1016/j.jeap.2013.12.001

Powers, J. (1994). What faculty say about working with graduate ESL Writers. Paper presented at the Annual Meeting of Teaching English as a Second Language, Baltimore, MD, March, 1994.

Shaw, P. (1991). Science research students' composing processes. *English for Specific Purposes 10*, 189–206. https://doi.org/10.1016/0889-4906(91)90024-q

Törnqvist, J. (2015). *Using Coh-Metrix to investigate changes in student texts: Comparing student writing from 1999 and 2009* [Unpublished undergraduate thesis]. Stockholm University.

Vasylets, O., Gilabert, R., & Manchón, R. M. (2019). Differential contribution of oral and written modes to lexical, syntactic and propositional complexity in L2 performance in instructed contexts. *Instructed Second Language Acquisition*, *3*(2), 206–227. https://doi.org/10.1558/isla.38289

Wagener, B. (2018). The importance of affects, self-regulation and relationships in the writing of a master's thesis. *Teaching in Higher Education*, *23*(2), 227–242. https://doi.org/10.1080/13562517.2017.1379480

Wang, X. and Yang, L. (2012). Problems and strategies in learning to write a thesis proposal: a study of six MA students in a TEFL program. *Chinese Journal of Applied Linguistics, 35* (3), 324–341. https://doi.org/10.1515/cjal-2012-0024

Wind, A. M. (2013). Second language writing development from a Dynamic Systems Theory perspective. In *Proceedings of the 8th Lancaster University Postgraduate Conference in Linguistics & Language Teaching* (pp. 90–123).

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

Xie, J. (2016). Direct or indirect? Critical or uncritical? Evaluation in Chinese English-major MA thesis literature reviews. *Journal of English for Academic Purposes*, *23*, 1–15. https://doi.org/10.1016/j.jeap.2016.05.001

Yin, S., Gao, Y., & Lu, X. (2021). Syntactic complexity of research article part-genres: Differences between emerging and expert international publication writers. *System, 97*, 102427. https://doi.org/10.1016/j.system.2020.102427

Zhao, J. (2017). Native speaker advantage in academic writing? Conjunctive realizations in EAP writing by four groups of writers. *Ampersand, 4*, 47–57. https://doi.org/10.1016/j.amper.2017.07.001