

Research paper

Syntactic Complexity in the Discussion Sections of Master's and Doctoral Theses: A Cross-Linguistic Study in Applied Linguistics

Zahra Shahbazi Sham Asbi^{*1}, Maryam Farnia²

^{*1}MA in English Language Teaching, Department of English Language and Literature, Payam Noor University, Tehran, Iran (Zahra.shahbazi1992@gmail.com)

²Assistant Professor of Applied Linguistics, Department of English Language and Literature, Payame Noor University, Tehran, Iran (mfarnia@pnu.ac.ir)

Shahbazi Sham Asbi, Z., & Farnia, M. (2024). Syntactic complexity in the discussion sections of master's and doctoral theses: A cross-linguistic study in applied linguistics, *Journal of new advances in English Language Teaching and Applied Linguistics*, 6(1), 124-148.

Doi: 10.22034/Jeltal.2024.6.1.6

*corresponding author's email: Zahra.shahbazi1992@gmail.com

Abstract

The current study investigates syntactic complexity in the discussion sections of master's and doctoral dissertations in applied linguistics. To this end, 30 local Iranian MA theses and 30 local Iranian PhD dissertations in applied linguistics were randomly selected from those published in IranDoc (Iran Document Center). Additionally, an international corpus comprising 30 doctoral dissertations in applied linguistics was randomly selected from research databases such as ProQuest. The data were then analyzed using syntactic complexity variables through both descriptive (frequency) and inferential (chi-square) analyses. Furthermore, eight indices of syntactic complexity were considered as the basis for this comparative study. The findings indicate a significant effect of education on the use of these eight syntactic complexity indices. At the conclusion of this paper, the significance of syntactic complexity in terms of structural writing is discussed, along with the implications of these findings for English for academic purposes writing research and pedagogy. Finally, the results and statistics support the conclusion of meaningful differences and considerable frequency in the use of certain syntactic structures over others. Given its practical implications, further research is necessary to better understand the impact of syntactic complexity on writing quality and development.

Keywords: academic writing theses, syntactic complexity, discussion section, academic writing.

Received: 2024-01-12 **Revised** 2024-04-20 **Accepted** 2024-04-25

Introduction

In recent decades, research on syntactic complexity has prominently featured in the literature on second language (L2) writing, with a significant increase in the number of studies conducted

over the past ten years. This field has a long-standing tradition of utilizing various measures of syntactic complexity as potential criteria for assessing language proficiency and development (Bi & Jang, 2020; Casal & Lee, 2019). While clausal coordination and subordination exhibit minimal variation in writing, it has been demonstrated that high-rated writings tend to display greater complexity. Syntactic complexity measures, such as the length or number of words per clause, t-unit, or sentence, the number of clauses per t-unit or sentence, and the types of clauses, have been extensively employed to establish benchmarks for language proficiency and improvements in speaking and writing (Bulté & Housen, 2014; Crossley & McNamara, 2014; Lu, 2011).

Additionally, syntactic elaboration has been identified as the most significant factor contributing to variations in writing scores. Furthermore, the combination of both syntactic expansion and diversity has been justified as accounting for the highest variance in research on writing scores. Writing quality is also influenced by short-term courses on syntactic and lexical complexity in ESL writings (Bulté & Housen, 2015). Prior research has noted differences in syntactic complexity regarding the length of the production unit, the extent of subordination, the degree of coordination, and the level of phrasal complexity between the writings of native speakers (NS) and non-native speakers (NNS) (Ai & Lu, 2015; Foster & Tavakoli, 2009; Mancilla et al., 2015).

While a notable concentration of analyses on syntactic complexity has focused on profiling linguistic features, rhetorical constructions have emphasized the conventions associated with specific structures of syntactic complexity. The application of complex syntactic structures varies according to the density of complexity in academic theses. In addition to a comprehensive understanding of the rhetorical demands of writing conventions, there is an increasing recognition

of how writers adapt their use of complex structures based on their rhetorical objectives. The insights gained from such analyses are valuable for assessing and teaching academic disciplinary genre writing. As understanding of the options available within academic disciplines evolves, it can illuminate how these options differ according to the various purposes of authors in academic theses. Each high-rated paper exhibits the highest densities of complexity, while low-rated papers demonstrate the lowest densities, indicating that features of phrasal complexity are significant elements in this context (Casal & Lee, 2019).

The present study is going to provide the body of scholarship which works over the syntactic features in academic writing across theses contexts by exploring syntactic complexity of eight measures in Iranian local MA and PhD theses in applied linguistics and international theses in applied linguistics. The following questions will guide the study.

1 - Are there any differences in the syntactic complexity, as evaluated from eight established measures, of academic writings part-genres focus on theses discussion section in Iranian local MA and PhD theses in applied linguistics?

2. Are the use of syntactic complexity different in these two corpora from international theses?

Review of the Literature

Syntactic Complexity

Applied linguistics extensively utilizes syntactic complexity as a means to measure development in English as a Foreign Language (EFL) learning and teaching (Jin & Lu, 2020; Lambert & Nakamura, 2018; Mao & Jiang, 2017; Wu et al., 2020). Furthermore, studies focused on English

for Academic Purposes (EAP) have leveraged the use of syntactic complexity over the years (Devitt, 2014; Eka Sanhadi Rahayu et al., 2021; Kuiken et al., 2019). Collectively, these studies underscore the significance of syntactic complexity. According to Mazgutova and Kormos (2015), who investigated English for Specific Purposes (ESP) writing, both lexical and syntactic knowledge in L2 learners demonstrate that syntactic complexity plays a crucial role in writing development.

In the context of Second Language Acquisition (SLA), different types of syntactic complexity warrant critical discussion; however, in some studies, syntactic complexity has been treated as a secondary or dependent variable in relation to oral and written performance (Kuiken et al., 2019). The variation of syntactic complexity, as examined by Kuiken et al. (2019), outlines contrasting approaches to understanding how factors such as levels of complexity, linguistic proficiency, task type, and genre influence SLA. Their study suggests that learner motivation or personality may also contribute to variations in syntactic complexity. The construct of syntactic complexity and its measurement in L2 writing development are two additional norms that have been effectively considered (Larsson & Kaatari, 2020; Ortega, 2015).

Syntactic Complexity in Academic Writing

Writing development is a significant area of investigation in academic writing, particularly to explore the connection between knowledge development and writing features. A notable study on second language acquisition and syntactic complexity in L2 learners was conducted by Crossley and McNamara (2014). This computational study examined the impact of developmental changes on individual judgments of writing quality over the course of a semester. The findings revealed

that L2 writers predominantly used more nouns and demonstrated increased phrasal complexity; however, human raters tended to evaluate clausal complexity, indicating that syntactic complexity is often the most time-consuming aspect in writing classes. During this study, learners were able to produce longer noun phrases before the main verb and used phrasal verbs to avoid repetitive syntactic structures.

Hyland (2002) also investigated the syntactical effects on academic writing, providing reliable insights for writing academic papers. His study focused on the use of restricted first-person pronouns and demonstrated that not all structures and disciplines adhere to the coded impersonality often expected in academic writing. Among the various studies on writing quality in relation to syntactic complexity, one of the most practical is Jagaiah's (2017) investigation, which considered four potential variables in sentences: pattern, length, connectors, and sophistication. This study aimed to examine the relationship between these variables and writing quality for both risk-taking and non-risk-taking students. Ultimately, the findings indicated a moderate to fair positive relationship among all four variables and writing quality, highlighting that this relationship is influenced by the type of student.

Disciplinary Variation in Academic Writings

Cross-disciplinary in academic writing is one of the most useful studies in case of writing developments. Hu and Liu (2018) represented a cross-disciplinary of rhetorical structure study with the focus on Three Minute Thesis (3MT), and a corpus of 142 presentations by Doctoral students. The findings were categorized in hard and soft discipline. A hard discipline is inclined to employ quantitative data, tended to be predictive and experimental; and a soft discipline is inclined

to apply qualitative data, generally, no experiment is performed, and no prediction concerning the future are made based on the research performed. Based on this studies' finding most important differences were in three moves which are: framework, result and methods. the theoretical framework was chosen by soft-discipline writers more than hard-discipline writers and in contrast hard-discipline writers prefer to have methods. There was one other category on disciplinary study by Hu and Liu (2018) and it was based on pure and applied research which in this case the focus is on the purpose and relationship of a study. An applied relation means to use of the study's result for some specific purpose, under this study more results move in pure-discipline was presented than applied-discipline.

Success in higher educations is always an important goal for EFT student, and it depends on student comprehension ability and writing skills of their chosen courses; obviously this level of being a skillful person needs literacy practice and academic discipline, therefore from this view point North (2005) studied on disciplinary deference like pure science and humanities and pure social science, considering the relation of nature of knowledge also nature of disciplinary culture. Based on the finding under-graduated tent to use requirements essay writing, and also it is proven that student tendency on showing constructed and contested, oriented them more interpersonal. The findings also included the variety of learner's knowledge in soft and hard discipline.

Doctoral students are always a very good target to study disciplinary based on their writing skills. What has been investigated by Wang and Zeng (2021) is on self-preference in disciplinary but in narrowed sampling of hedges and boosters Doctoral writing and their self-mentioned used in order to manifest the preferred use of authorial presence and functions. Disciplinary use is a very important subject in academic writings, in Kaidan et al., (2021), nominalization was explored as

an important controversial topic. The findings indicated in applied linguistic writings the nominalization has been more used than counterparts; however, not many significant differences among them announced.

Previous Studies

The amount of empirical studies representing the inconclusive results are often contradictory in SLA research, in relation to L2 proficiency level, task modality, task type and aim of the task (Bulte & Housen, 2015; Ortega & Norris, 2009), also in some other empirical studies the source of language is involved (Kuiken et al, 2019). The way of the most complexity, language competence or improvement has been also investigated (Bulte & Housen, 2015; Crossley & Mcnamara, 2014; jin and Lu, 2020, Mazgutova & Kormos, 2015) one more popular measurement is contemplated to be the average of subordination clause by calculating them (Casal et al.2021). Most of the researches in any investigation obtain their data from written text. (Arya et al., 2011; Johnson & William, 2017; Kormos & Mazgutova, 2015)

Researches on linguistic aspects are using of English as a lingua franca (ELF) from a syntactic complexity perspective, to realize in what way ELF users state meaning connections in any academic writing (Wu et al, 2020). Moreover, a national Corpus like academic text, common science, news, and fiction was the target of some other studies (Kaatari & Larsson, 2020).

The study of Wu et al (2020) extracted the samples of the study from SciELF corpus and Corpus of Contemporary American English (COCA) SciELF, in order to reach the objective of study which is a completion of an earlier research on linguistic features of English as a lingua franca

(ELF). Mazgutova and Kormos (2015) in order to find out the process of changing in the lexical and syntactic features of L2 learners' in academic text, designed a course of a one-month-long intensive English for Academic Purposes (EAP) and investigated the process and programmed at a British university during summers of 2012 and 2013, to get the main aims, which the first is to improve learner's employment of English in an academic setting and second promote a consciousness of the acquiring abilities and tactics. Another process-based study has been done by Mao and Jiang (2017) who investigated the influences of the agreement involved in the continuing duty with a focus on syntactic complexity in L2 written production in to two groups with a focused on 7 indexes. The result represents some significant differences and proves that dissimilarities occur in six of the syntactic complexity measures.

Several quantitative researches like (1a) studies between L1 and L2 suggest a meaningful dissimilarity in syntactic complexity between L1 and L2 texts. Individuals who examined syntactic complexity applied different syntactic complexity measures as a potential criterion of language competence or development (Bulte & Housen, 2015; Crossley & Mcnamara, 2014; Kormos & Mazgutova, 2015). Mostly, outcomes reveal that the complexity measures chosen can denote substantial variations in writing proficiency while contrasting diverse proficiency levels.

Method

Corpus of the Study

This study was an attempt to examine the type, frequency, and different use of syntactic complexity, employed in the discussion section of final academic thesis in English teaching from

level of Iranian local MA and PhD theses and international PhD theses. The corpus of this study comprised of 195,675 Words extracted from 30 Iranian MA theses in applied linguistics total 50,822 words. A number of 30 theses was chosen from Iranian PhD theses from the top universities in Iran, total 55,030 words and the last corpus 30 number of PhD theses from top university of US total 89,823 words. Details of the corpus can be found in Table 1.

Table 1

Detail of Corpora

Corpora	Level	Size	Number of words
Local Corpus	Iranian MA in applied linguistics	30	50,822
	Iranian PhD in applied linguistics	30	55,030
International Corpus	PhD in applied linguistics	30	89,823
Total		90	195,675

Instruments

In this study eight syntactic complexity which were measured and computed by employing a modified form of the Syntactic Complexity Analyzer (SCA Casal et al, 2021.). This analyzer is an easily accessible automatic syntactic complexity analyzer. The adapted analyzed information shows eight element sub-constructs explained in detail in the data analysis part of this paper. The complexity index in which it is computed, with the terminal nodes, represents “nonoverlapping discrete sub-constructs” and nonterminal nodes represents “composites of discrete sub-constructs” (Casal et al, 2021). These measures and main computation processes are displayed in Table 2.

Table 2

Syntactic Complexity Measures, Labels, And Calculations. (Casal et al. 2021)

Sub-construct	Measure name	Label	Calculation
Overall sentence complexity	Mean length of sentence	MLS	Word count/sentence count
Clausal coordination	T-units per sentence	TU/S	T-unit count/sentence count
Overall T-unit complexity	Mean length of T-unit	MLTU	Word count/T-unit count
Clausal subordination	Dependent clauses per T-unit	DC/TU	Dependent clause count/T-unit count
Elaboration at clause level	Mean length of clause	MLC	Word count/clause count
Phrasal coordination	Coordinate phrases per clause	CP/C	Coordinate phrase count/clause count
Noun phrase complexity Complex	NPs per clause	CN/C	Complex NP count/clause count
Nonfinite elements/ subordination	Nonfinite elements per clause	NFE/C	Nonfinite element count/clause count

Result

Syntactic Complexity in Iranian MA and PhD Corpus

As stated before, the first research question intended to compare the syntactic complexity of discussion sections of Iranian MA and PhD theses in terms of their syntactic complexity. In so doing, two corpora including local Iranian MA and PhD theses were chosen randomly and they were analyzed using Syntactic Complexity Analyzer which scores words, sentences, clauses, dependent clauses, T-units, complex T-units, coordinated phrases, complex nominal and verb

phrases and it was mentioned and explained in chapter three. The descriptive statistics are exhibited in Table 3.

Table 3

Complexity Measures in MA and PhD Theses

Complexity measure	MA theses		PhD theses		Total
	Frequenc y	Percentag e	Frequenc y	Percentag e	
Mean length of sentence (MLS)	1544.87	36.58	2678.02	63.42	4222.90
Mean length of T-unit (MLTU)	1435.60	36.66	2479.57	63.34	3915.17
Mean length of clause (MLC)	771.93	36.93	1317.99	63.07	2089.92
Clause per sentence(C/S)	26.16	39.96	39.30	60.04	65.46
Dependent clauses per T-unit (DC/TU)	49.612	37.48	74.7375	62.52	124.34
T-units per sentence (TU/S)	64.76	39.92	97.462	67.08	162.22
Coordinate phrases per clause (CP/C)	23.16	38.10	37.62	61.9	60.78
NPs per clause (CN/C)	100.53	49.20	193.76	50.8	294.76

Table 3 shows the descriptive statistics of various complexity computes of the discussion sections in the two local MA and PhD corpus. Generally, the complexity measures were employed more frequently in PhD corpus in comparison with MA these. As it is informed in the table the MLS, MLTU and CN/C are three the most frequent use of a complexity in Iranian MA theses respectively, which are exactly the most frequency in Iranian PhD theses. Furthermore, all understudy items in PhD has been grown more than 50% except of CN/C which is 50% otherwise in MA theses this is just TU/S and CN/C have got more than 60% and 100% frequency in order. Based on the table there is no same frequency and percentages in Iranian MA and PHD theses however it is worth to be mentioned that there is a raise of frequency and percentages in all aspect

of a sentence in PHD these. Figure 1 and Figure 2 compare the two corpora with regard to various computes of syntactic complexity.

Figure 1

Complexity Measures in Local MA and PhD Theses

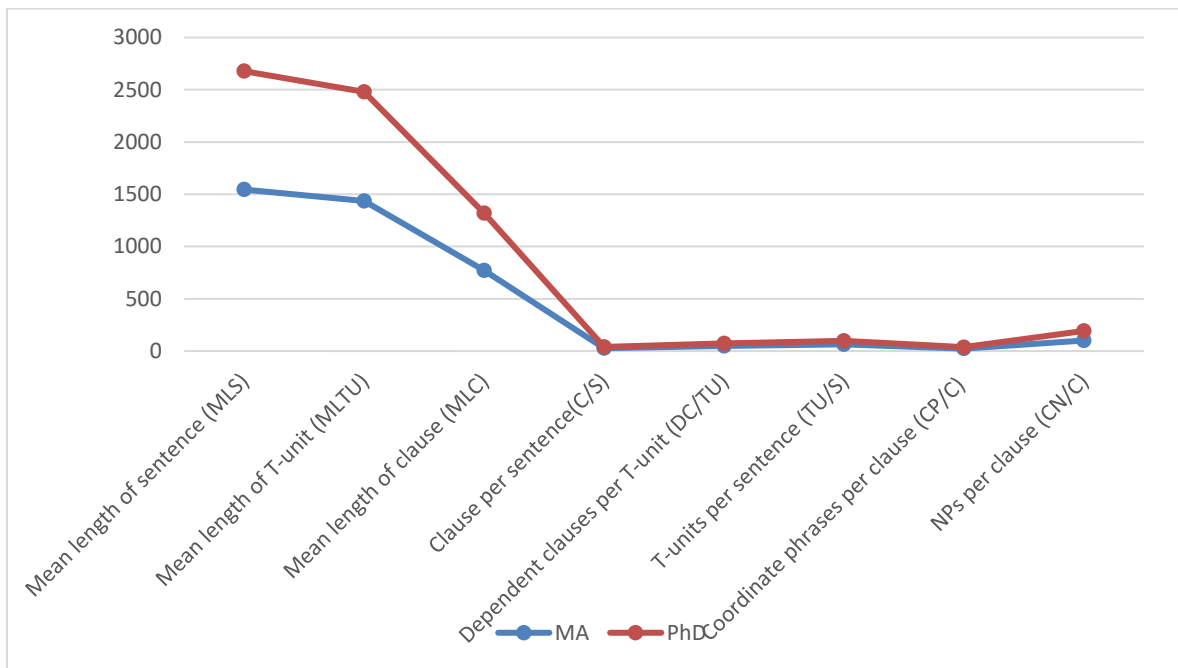


Figure 1 depicts frequency of syntactic complexity in two corpora. To begin with on an overlook on the line graph, to par of strong frequency of MLS, MLTU, and MLC is easily distinguishable from weak frequency of C/S, TU, TU/S, CPC and CN/C by lines and the process. Comparing the first most frequency, the mean length of sentence, shows 1000 frequencies in MA less than PhD in the discussion part. Based on the line graph the gap between two corpora is getting less and less by reaching the clause per sentence and falls to less than 500 frequencies. Moreover, in TU/S the graph experiences a slight difference between MA and PhD. The distance between two corpora is getting more and it shows that PhD students use CN/C more than MA students in their writing.

Due to the fact that PhD students use more complexity in large number of frequency rather than MA students. Additionally, in order to assess the two corpora in terms of complexity measures, a Chi-square test was administered. The results are shown in Table 4.

Table 4.

Results of Chi-square Test Comparing Complexity Measures between Local MA and PhD Theses

	Sig.	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	81.247 ^a	15	.000
Likelihood Ratio	1.735	14	1.000
Linear-by-Linear Association	.088	1	.779
N of Valid Cases	1362		

According to the statistics presented in Table 4 ($\chi^2(15) = 1.000, p < 0.05$), p value is slighter than the significance level of .05; thus, it could safely be stated that there is a meaningful difference between the discussion sections of local Iranian MA and PhD theses regarding syntactic complexity as measured by eight measures of syntactic complexity.

Syntactic Complexity in Local and International PhD Corpora

In response to the second research question, “are the use of syntactic complexity different in these two corpora from international theses?”, two corpora of PhD theses were chosen, then the discussion sections of these theses were compared. Table 5 presents the results.

Table 5

Complexity Measures in Local MA and PhD Theses

Complexity measure	Iranian		International		Total
	F	P	F	P	
Mean length of sentence (MLS)	2678.02	43.33	3501.74	56.66	6179.76
Mean length of T-unit (MLTU)	2479.57	43.49	3221.57	56.51	5701.14
Mean length of clause (MLC)	1317.99	39.95	1980.90	60.05	3298.89
Clause per sentence (C/S)	39.30	16.19	242.70	83.81	282
Dependent clauses per T-unit (DC/TU)	51.50	40.8	74.7375	59.20	126.23
T-units per sentence (TU/S)	97.462	39.95	146.49	60.05	243.95
Coordinate phrases per clause (CP/C)	37.62	40.72	54.75	62.38	92.37
NPs per clause (CN/C)	193.76	40.29	287.152	59.71	480.912

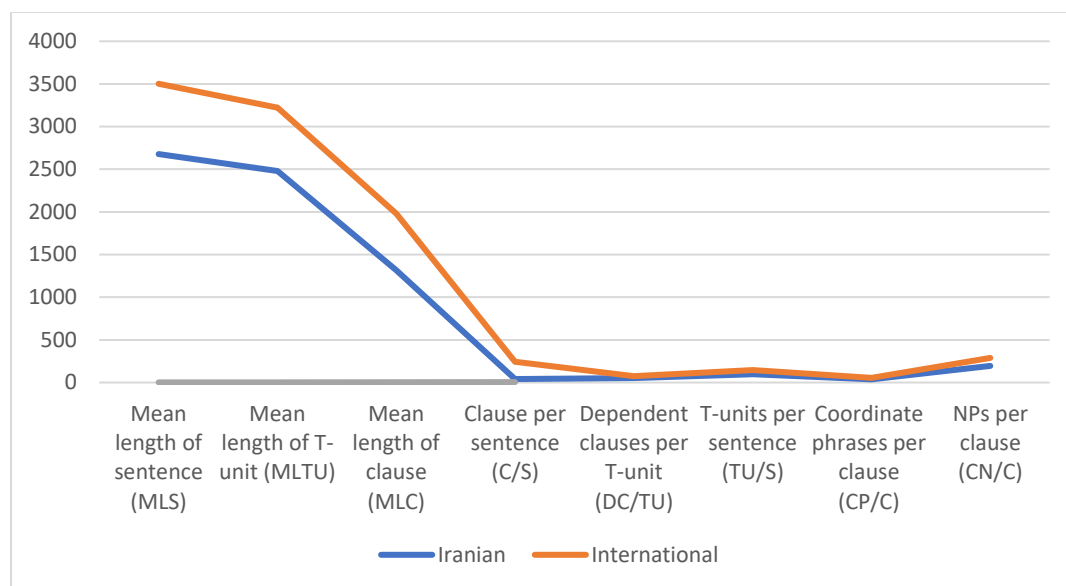
Note: F= Frequency, P=Percentage

Descriptive statistics revealed that international writers made use of more complex structures than Iranian writers. According to Table 5, the first most used complexity is MLS in Iranian theses which is almost 1000 times less than MLS in international theses. The second and the third MLTU and MLC have announced almost 2500 and 1300 frequency of use in Iranian PhD theses, respectively same as in international PhD theses the MLTU and MLC have got the second and third place however they show almost 1,000 times more use in frequency data. The most considerable difference in this table belongs to clause per sentence C/S which shows 16.19% in Iranian PHD theses but almost 84% in international PhD theses. The other understudy items (DC/TU, TU/S, CP/C and CN/C) represent frequency of almost 40% in Iranian theses and

approximately 60% in international PhD these. Figure 3 compares the two corpora in terms of three complexity measures which were employed more than 1,000.

Figure 2

Complexity Measures in PhD Theses by Local and International PhD Theses



As it is represented in Figure 2 the International PhD writers use all eight items of syntactic complexity with more frequency of use. As it is seen there is a clear gap between the repeated use of each item between Iranian and International PhD theses in discussion part of their theses but in DC/TU, TU/S and CP/C which is almost covered by each other and their use is so close. Based on the line representing most used item (MLS). International writers use almost 1,000 times more than Iranian writers in their discussion part which is also so similar data of first line chart about a comparative analyze between Iranian PhD and Master eve of education behalf of PhD level. One of the most significant things of this line graph is the fact that the use of the clause per sentence (C/S) has used fairly close to zero in Iranian discussion part otherwise in international PhD corpus

there are almost 250 frequencies of use. At the end of the lines the growth in use of CN/C is obvious and also there is slight gap between local and international corpora considering the point that international writers still use more even in this case even though the total use of both corpora is less than 500 but at least they are representing more numbers of use than C/S. The graph is admitting the stronger writing in the international corpus. A Chi-square test was administered between the two corpora. The results are shown in Table 6.

Table 6

Results of Chi-square

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	66.312 ^a	15	.000
Likelihood Ratio	1.625	14	.936
Linear-by-Linear Association	.085	1	.637
N of Valid Cases	1421		

According to the results presented in Table 6, there was a meaningful dissimilarity between the two corpora of Iranian and international PhD theses in terms various complexity measures ($P < .05$). This led to the refusal of the second null hypothesis of the study claiming that there is no considerable difference between the discussion sections of PhD theses written by local and international PhD theses in terms of syntactic complexity.

Discussion

The findings reveal significant patterns across the three corpora: local MA theses, local PhD corpus, and international PhD corpus. There is a clear and substantial distinction between the

authors at the MA and PhD levels of education. The discussion sections of international PhD theses exhibit the highest complexity across eight measures of syntactic variables. Notably, the mean length of sentence (MLS) is utilized most frequently in all three corpora, with frequencies of 1,500 in local MA theses, over 2,500 in local Iranian PhD theses, and 3,500 in international PhD theses within their discussion sections.

Among the three complexity measures (mean length of a sentence, mean length of T-unit, and mean length per sentence), the discussion sections of international PhD corpus employ syntactically complex structures significantly more often than those of local PhD and MA corpus. In terms of the measures of clauses per sentence and noun phrases (NPs) per clause, international PhD corpus demonstrate a higher frequency of use compared to both local Iranian PhD and MA corpus, which exhibit nearly identical frequencies. The results show that the two complexity measures—dependent clauses per T-unit and coordinated phrases per clause—have the least impact on sentence complexity, as they are employed the least frequently in the discussion sections of all three corpora.

In response to the first question—whether there are significant variations between the discussion sections of local Iranian MA and PhD theses in applied linguistics in the use of syntactic complexity—the analysis of the data indicates that the frequency of complexity in local PhD theses is higher than in local MA theses. The findings show that the measures show that all eight indices are more frequent in local PhD theses than in MA theses, although not all complexity measures reveal substantial differences. Firstly, in the discussion sections of local MA theses, the mean length of sentence (MLS) is utilized more frequently than the mean length of T-unit (MLT), which

aligns with the findings of Casal et al. (2021) that indicate a lower use of T-units in academic writing. It should be noted that this comparison has been made across different sub-genres.

The presence of significant and clear variations in syntactic complexity is not surprising. While the results indicate the higher quality of international PhD corpus in applied linguistics writing, there are notable similarities in complexity between the two corpora. For instance, in the case of noun phrases per clause (CN/C), the difference is less than 2%. In contrast, for T-units per sentence (TU/S), the frequency in local Iranian MA theses is 67.76, which is nearly half of the frequency observed in PhD theses. Given that both corpora pertain to the same disciplinary field, these differences may reflect weaknesses in the writing skills of Iranian MA students or insufficient knowledge regarding the effective use of complex structures at this educational level. As investigated by Enk and Power (2017), the progression of practitioners and individual knowledge is a valuable and important achievement in any research endeavor.

Teaching academic writing through contact-based instruction leads to improved writing outcomes for students, particularly when they practice the genre through written compositions. Despite the challenges learners may face, they are able to grasp the overall points of a topic (Brooks, 2017). From the findings of this study, a significant conclusion can be drawn: an increase in the use of coherence instruments in writing, especially when teaching genre-specific writing, enhances writing quality.

When comparing Iranian and international PhD theses, the knowledge gained through the process of higher education, along with extensive study and the writing of academic theses and papers, enables students to examine different clauses, thereby improving writing quality. This experience equips them with the skills necessary to produce more professional work.

In a study conducted by Crossley et al. (2017) on cohesion features and domain-independent lexical features, strong support for both macro and micro disciplinary writing was found. The study indicated that all academic articles are macro-additive and contain more frequent words, a pattern that was similarly observed in the PhD theses analyzed in this study. Additionally, Crossley et al. (2017) demonstrated that micro disciplinary writing often results from a lack of knowledge and skills, leading to the use of less complex sentences due to insufficient general education received by students.

In response to the second research question, the findings showed that overall a considerable pattern extracted from these findings at syntactic complexity. Generally speaking, there was a clear priority in international corpus in comparison with Iranian PhD corpus. Based on what has been done by Kuiken, et al (2019), syntactic complexity probable depends on proficiency level of a writer or better to say level of L2 learners. Also, in Kuiken, et al (2019) study beginner and intermediate L2 learners use complexity of coordinates more than phrasal complexity which is chosen by advanced level of students, fortunately it is a proof for the current study findings.

This study managed to compare two same level of a student but under one inevitable qualification of the writer education background. Similarly, in the study has been done by Lahuerta (2018), the finding allows us to recognize the way that syntactic complexity develops in same proficiency levels. Like what happened in this study which the chart and table in chapter four depict growth in length of linguistic units in international corpus in applied linguistics in comparison with Iranian PhD theses. In addition, in Lahuerta (2018) the results of examining general writing skill between upper-intermediate and intermediate students agree with the findings by Bulte and Housen (2015), that significant increase in compound sentences based on an intensive English writing program at

the end of the course for college-level is one of the most powerful ways of upgrading students (EFL learners)'s level of proficiency.

Academic writing skill has commonly been examined regarding enhancements in several evaluation standards, like cohesion, coherence and organization, also student's writing showed how the number of using recurrent words and complex sentence, compound complexity has been increased. The most repeated syntactic device which has been applied for changing part, thing and place in lexical compensation was the relative clause. Possibly, the restricted employment of relative clauses which has been done by the less proficient L2 learners diverged to the more proficient learners may have been related to their restricted access to task-relevant lexis.

The other significant point of this study is the large distance of using the clause per sentence, while in international the use of the clause per sentence is 83.81%, unlikely in Iranian PhD thesis it is just 16.19 use of the clause per sentence, so it could be admitted that Iranian writers need to use more clauses in the sentences to write more natives like writings. From Larsson and Kaatari, (2020) the employment of Noun Phrases in an expert writing was more expanded by adjectives and propositions therefore this result ascertains findings of this paper of more proficiency use of noun phrases as what has been found on international PhD theses. In addition to expanded adjective phrases, the frequency of coordinated phrases and use of coordinate nouns or adjectives in case of clarity should be considered too; this finding is obviously a proof of the using phrases importance in writing a native like paper as what you can see in international PhD theses in applied linguistics (Wu et al, 2020).

Conclusion

The explanation provided in this comparative research on syntactic complexity, which investigates academic theses at two different levels of education (MA and PhD) in applied linguistics, offers a perspective on skillful writing that, while not entirely new, can serve as a foundational study for other researchers pursuing the question of how to write in a more native-like manner or how to enhance the quality of writing. Therefore, it considers the fact that better writing at higher levels of education demonstrates the effect of educational background on writing quality. Furthermore, the greater complexity observed in theses written by native speakers reflects the impact of the environment on producing more syntactic complexity. The findings indicate that comparing different proficiency levels inevitably captures meaningful differences in thesis writing when complexity measures are employed (Lahuerta, 2018).

Technically, the more frequent use of specific syntactic structures, such as coordinated phrases, at any educational level leads to greater proficiency in writing. Thus, the idea of teaching objectively can be concluded as a method for increasing writing complexity. For instance, teaching linking words appropriate to the learners' level results in longer sentences and, consequently, greater complexity.

The results suggest that teaching how to use noun phrases and adjective phrases to create more native-like sentences, as well as increasing sentence length by incorporating more conjunctions and linking words, helps writers develop more skillful writing. Consistent with previous studies, the teaching process has proven to be effective, as noted in various sources (Kaatari and Larsson, 2020; Lambert & Nakamura, 2018; Rahayul et al., 2021; Wu et al., 2020). Moreover, comparing different proficiency levels inevitably captures meaningful diversities in thesis writing when

complexity measures are chosen. It can be stated that syntactic complexity and writing skills tend to converge at higher levels of English competency (Lahuerta, 2018).

Technically, the more frequent use of specific syntactic structures, such as coordinated phrases, at any educational level leads to greater proficiency in writing. Additionally, teaching these structures, such as linking words appropriate to the learners' level, results in longer sentences and, consequently, increased complexity. On one hand, this research has found that while overall sentence complexity shows similarities in the use of syntactic complexity, both Iranian PhD students and native English speakers utilize complexity per sentence. On the other hand, not only is the level of education a significant factor in writing, but the writer's background also affects writing quality. There are significant differences across eight indices regarding the frequent use of syntactic complexity among EFL students including aspects such as length and coordination which were in line with Wu et al.'s (2020) studies.

To acknowledge some limitations of the present study, which may be addressed in future research, it is important to mention that.

References

- Ai, H. & Lu, X., (2015), Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds: *Journal of Second Language Writing*, 29, 16-27. <http://dx.doi.org/10.1016/j.jslw.2015.06.003>
- Arya, D. J., Hibeart, E.H., & Pearson, D. (2011). The effects of syntactic and lexical complexity on the comprehension of elementary science texts. *International Electronic Journal of Elementary Education* 4(1), 107-125.
- Bi, P., & Jiang, J. (2020). Syntactic complexity in assessing young adolescent EFL learners' writings: Syntactic elaboration and diversity. *System*, 91, 102248. <https://doi.org/10.1016/j.system.2020.102248>
- Bulte, B., & Housen, A. (2015). Evaluating short-term changes in L2 complexity development: *Círculo de Lingüística Aplicada a la Comunicación* 63, 42-76. http://dx.doi.org/10.5209/rev_CLAC.2015.v63.50169
- Bulte, 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>
- Brooks, G. (2017). Teaching academic writing in content-based classes: In P. Clements, A. Krause, & H. Brown (Eds.), *Transformation in language education* (pp. 386-391). Tokyo: JAL
- Casal, J., & Lee, J. (2019). Syntactic complexity and writing quality in assessed first-year L2 writing. *Journal of Second Language Writing*, 44, 51–62. <https://doi.org/10.1016/j.jslw.2019.03.005>
- Casal, J., Lu, X., Qiu, X., Wang, Y., & Zhang, G. (2021). Syntactic complexity across academic research article part genres: A cross-disciplinary perspective: *Journal of English for Academic Purposes*, 52, 100996. <https://doi.org/10.1016/j.jeap.2021.100996>
- Crossley, S., & McNamara, D. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners: *Journal of Second Language Writing* 26, 66–79. <http://dx.doi.org/10.1016/j.jslw.2014.09.006>
- Crossley, S., Russell, D., Kyle, K., & Romer, U. (2017). Applying natural language processing tools to a student academic writing corpus: How large are disciplinary differences across science and engineering fields?, *The Journal of Writing Analytics*, 1(1),48-81.
- Devitt, A. J. (2014). Genre performances: John Swales' genre analysis and rhetorical-linguistic genre studies: *Journal of English for Academic Purposes* 19, 44-51.
- Enk, A.V., & Power, K. (2017). What is a research article?: Genre variability and data selection in genre research: *Journal of English for Academic Purposes* 29, 1-11, <http://dx.doi.org/10.1016/j.jeap.2017.07.002>.
- Eka Sanhadi Rahayul, F., Utomo, A., & Setyowati, R. (2021). Syntactic and lexical complexity of undergraduate students' essays: A comparison study between L1 and L2 writings: *Indonesian Journal of English Language Teaching and Applied Linguistics*, 5(2), 251-263, Doi: <http://dx.doi.org/10.21093/ijeltal.v5i2.706>.
- Foster, P., & Tavakoli, P. (2009). Native speakers and task performance: Comparing effects on complexity, fluency, and lexical diversity. *Language Learning*, 59(4), 866-896, <https://doi.org/10.1111/j.1467-9922.2009.00528.x>

- Hu, G., & Liu, Y. (2018). Three minutes thesis presentations as an academic genre: Across disciplinary study of genre moves, *Journal of English for Academic Purposes*, 35, 16-30. <https://doi.org/10.1016/j.jeap.2018.06.004>
- Hyland, K. (2002). Options of identity in academic writing. *ELT Journal*, 56(4), 351-358. <https://doi.org/10.1093/elt/56.4.351>
- Heidari Kaidan, Z., Jalilifar, Z., & Don, A. (2021). On the significance of disciplinary variation in research articles: Perspectives from nominalization, *Cogent Education*, 8, 1-18. <https://doi.org/10.1080/2331186X.2021.1890872>
- Johnson, I., & William, B. (2017). Choice of academic major at a public research university: The role of gender and self-efficacy. *Research in Higher Education*, 58, 365-394. [10.1007/S11162-016-9431-1](https://doi.org/10.1007/S11162-016-9431-1)
- Jagaiah, T. (2017). *Analysis of syntactic complexity and its relationship to writing quality in argumentative essays*. Doctoral dissertations. University of Connecticut. <https://opencommons.uconn.edu/dissertations/1571>.
- Jin, T., & Lu, X. (2020). Syntactic complexity in adapted teaching materials: Differences among grade levels and implications for benchmarking. *Modern Language Journal*, 104(1), 192-208, <https://doi.org/10.1111/modl.12622>
- Kuiken, F., Vedder, I., Housen, A., & De Clercq, B. (2019). Variation in syntactic complexity: Introduction: *International Journal of Applied Linguistics*. 29, 161–170, 10.1111/ijal.12255.
- Lahuerta, A. C. (2018). *Analysis of syntactic complexity and L2 proficiency in EFL writing*. Paper presented in International Conference on Research in Education, Teaching and Learning, Paris, France.
- Lambert, C., & Nakamura, S. (2018). Proficiency-related variation in syntactic complexity: A study of English L1 and L2 oral descriptive discourse. *International Journal of Applied Linguistics*, 20(2), 1–17. <https://doi.org/10.1111/ijal.12224>
- Larsson, T., & Kaatari, H. (2020). Syntactic complexity across registers: Investigating (in)formality in second-language writing. *Journal of English for Academic Purposes* 45, 100850, <https://doi.org/10.1016/j.jeap.2020.100850>
- Mancilla, R., Polat, N., & Akcay, A. O. (2015). An investigation of native and nonnative English speakers' levels of written syntactic complexity in asynchronous online discussions. *Applied Linguistics*, 38(1), 112–134, doi: <https://doi.org/10.1093/applin/amv012>
- Mao, Z., & Jiang, L. (2017). Exploring the effects of the continuation task on syntactic complexity in second language writing. *English Language Teaching*, 10(8), 1916-4742, <http://doi.org/10.5539/elt.v10n8p100>.
- 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>
- North, S. (2005). Disciplinary variation in the use of theme in undergraduate essays: *Applied Linguistics*, 26(3) 431–452. <https://doi.org/10.1093/applin/ami023>
- Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of second language writing* 29, 82-94, <http://dx.doi.org/10.1016/j.jslw.2015.06.008>
- Ortega, L. & Norris, J. M. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–578, <https://doi.org/10.1093/applin/amp044>

- Wu, X., Mauranen, A., & Lei, L. (2020). Syntactic complexity in English as a lingua franca academic writing: *Journal of English for Specific Purposes*, 43, <https://doi.org/10.1016/j.jeap.2019.100798>.
- Wang, J., & Zeng, L. (2021). Disciplinary recognized self-presence: Self-mention used with hedges and boosters in PhD students' research writing. *SAGE Open*, 11(2), 1–13 <https://doi.org/10.1177/21582440211005454>