

A Corpus-based Evaluation of Syntactic Complexity Measures as Indices of Advanced English Text Comprehension

¹Peyman Nasrabady

²Hooshang Khoshshima*

³Nahid Yarahmadzahi

⁴Amir Mohammadian

Research Paper

IJEAP- 2501-2112 DOR: [20.1001.1.24763187.2025.14.1.4.6](https://doi.org/10.1001.1.24763187.2025.14.1.4.6)

Received: 2025-01-13

Accepted: 2025-02-27

Published: 2025-03-01

Abstract: Reading comprehension is a vital skill for language learners, enabling text understanding and academic success. Despite technological progress, written text has remained key to learning, especially in higher education, fostering knowledge acquisition and new ways of thinking. Many studies have explored reader-related challenges affecting comprehension; however, there is a need to cast a careful light on text-related factors as well. To address this, the current study examined syntactic complexity across four corpora of advanced academic reading texts to highlight the need for greater syntactic alignment in teaching and testing materials. By analyzing 100 texts from the Vision series textbooks, Iranian M.A. TEFL entrance exams, Cambridge IELTS reading tests, and discussion sections of research papers, the study addressed and reviewed the linguistic challenges EFL learners face in comprehending advanced academic texts. While differing in purpose, these texts share a relatively advanced complexity level. Typically, strong M.A. reading exam performance aligns with IELTS reading proficiency, which is often seen as readiness for research papers. Similarly, Vision coursebooks, the most advanced in Iranian high schools, are expected to exhibit a level of syntactic complexity that, while less dense than research papers, may sufficiently prepare students for university-level reading demands. However, recent research has suggested that such assessments or textbooks, including M.A. entrance exams, may not effectively prepare students for the complexities of real academic contexts. Therefore, using L2 Syntactic Complexity Analyzer (L2SCA), the study analyzed the texts, and a MANOVA test confirmed significant differences among the corpora. The findings further revealed that standardized tests and instructional materials often underrepresent the syntactic complexity of authentic academic research, creating a gap between learners' preparedness and real-world academic reading demands. In addition to challenging traditional views on test validity, the results highlighted the need for more representative and comparable syntactic features in instructional and assessment materials.

Keywords: Advanced Reading Texts, Corpus, Corpus-based, Reading Comprehension, Syntactic Complexity, Text Evaluation

¹ PhD Candidate of TEFL, Peyman.Nasrabady1994@gmail.com; English Department, Chabahar Maritime University, Chabahar, Iran.

² Professor of TEFL, Khoshshima2002@gmail.com (Corresponding Author); English Department, Chabahar Maritime University, Chabahar, Iran.

³ Assistant Professor of Linguistics, venayarahmadi@gmail.com; English Department, Chabahar Maritime University, Chabahar, Iran.

⁴ Assistant Professor of Linguistics, amiancmu@gmail.com; English Department, Chabahar Maritime University, Chabahar, Iran.

Introduction

Reading Comprehension: Reader Characteristics as well as Textual Features

Reading comprehension is a multifaceted process influenced by the interplay between the reader's abilities and the characteristics of the text. For English as a Foreign Language (EFL) learners, this process becomes even more demanding due to the linguistic complexities of advanced texts, such as academic research papers and standardized test materials. Textual features, including syntactic complexity, cohesion, and overall readability, are useful in determining the ease with which learners can process and understand written material ([Crossley et al., 2017](#); [McNamara et al., 2010](#))

In EFL contexts, where exposure to authentic English input is often limited, educational resources like textbooks and test preparation materials are pivotal in supporting language acquisition and academic development ([Ali et al., 2022](#); [Bernal & Bernal, 2020](#); [Nuttall, 2005](#)). However, when these materials fail to adequately reflect the grammatical and structural demands of real-world texts, learners and students find it challenging to shift to higher-level and university level academic reading ([Biber et al., 2011](#); [Parkinson & Dinsmore, 2018](#)). This gap in alignment in the learners input can hinder their ability to handle advanced academic tasks.

To better understand this misalignment, examining the distribution of key textual features across related/various text genres can provide valuable insights into how educational resources can better cater to learners' needs ([Hyland, 2008](#)). By addressing discrepancies between common learner input materials and the syntactic demands of academic texts, educators and material developers can create resources that strike a balance between linguistic complexity and learner readiness. This approach not only enhances reading comprehension but also equips learners with the skills necessary for better academic performance ([Crossley & McNamara, 2011](#)).

Syntactic Complexity in Advanced Texts: Insights for Comprehension and Instruction

The ability to comprehend advanced academic reading materials depends on the interaction between reader-related factors and text characteristics. Among text factors, syntactic complexity plays a reasonable role in determining both readability and comprehension. While existing research highlights the challenges posed by syntactic complexity of the texts, much of it tends to focus on isolated learner difficulties or the analysis of specific text types, often overlooking a comprehensive comparison across different sources of reading materials ([Curran, 2020](#); [Grabe & Stoller, 2019](#))

Academic texts can serve as typical benchmarks for advanced linguistic proficiency, with syntactic complexity knowledge and grammatical mastery being important factors for both writing and reading comprehension skills in that area. Research highlights that the knowledge of and the ability to effectively practice syntactically complex texts can significantly predict writing quality and language proficiency ([Jung et al., 2019](#); [Maamuujav et al., 2021](#)). In other words, mastery and a good grasp of syntax and grammar can enhance not only writing fluency but also reading comprehension in academic contexts ([Baron, 2020](#)). Effective language instruction, language assessment, and material development should thus, particularly in EFL settings, consider these features and skills to support L2 learners' academic proficiency ([Riemenschneider et al., 2024](#); [Uccelli et al., 2015](#)). This is because those texts frequently feature complex sentence structures with multiple phrases and clauses, posing syntactic challenges to EFL learners ([Basilan & De Sagun, 2024](#); [Nergis, 2013](#)). Learners with greater syntactic awareness are better equipped to process the sophisticated structures found in research papers and standardized tests ([Arya et al., 2011](#)). This can highlight the importance of aligning textbooks and assessments with research papers to ensure that learners are adequately exposed to consistent linguistic patterns across contexts. Such alignment fosters a more reliable evaluation of reading comprehension skills while simultaneously equipping learners to meet the demands of higher-level reading tasks. Minimizing syntactic discrepancies across textbooks, tests, and research papers is crucial to preparing learners for the challenges of advanced texts. Instructional materials designed

with this alignment in mind can support a smoother transition into complex reading tasks by ensuring that learners are exposed to syntactic features reflective of authentic academic language ([Arya et al., 2011](#)).

The alignment of syntactic complexity across textbooks, standardized tests, and authentic academic texts contributes to the validity and consistency of reading materials as well. By using research papers as a benchmark for advanced syntactic structures, material developers can provide learners with a foundation for tackling the linguistic challenges of academic and professional contexts the face when entering universities. Teachers, in turn, can design instruction to address syntactic gaps and better prepare students for the demands of advanced reading comprehension ([Basilan & De Sagun, 2024](#); [Johnson, 1981](#)).

In the context of English as a Foreign Language (EFL) learning, a key challenge arises from the often limited exposure EFL learners have to the target language, which can significantly hinder their ability to process complex sentence structures in written texts ([Zuhra, 2015](#)). This limited exposure can create a mismatch between the syntactic structures learners typically encounter in their EFL contexts, and the more complex syntax often found in advanced academic materials in universities, research papers, or standardized tests. This gap can present a considerable obstacle to reading comprehension ([Nuttall, 1996](#)). Effectively addressing this challenge requires careful consideration of instructional materials. Reading materials can play an essential role in bridging this gap by providing crucial exposure to a range of syntactic structures, thereby better equipping learners to handle the demands of increasingly complex texts and ultimately fostering improved reading comprehension and academic performance ([Lak, 2017](#)).

This study aims to bridge this gap by compiling corpora from textbooks, standardized tests, and academic research papers to analyze and compare their syntactic complexity indices. Through this comparative lens, the research seeks to offer a deeper understanding of the syntactic demands imposed by advanced texts among learner input. The findings enhance general awareness on the syntactic complexity characteristics of texts, informing language instruction.

Therefore, the study provides implications for material developers and test designers by highlighting the importance of representativeness, consistency, and validity in reading materials and assessments. Accordingly, addressing syntactic discrepancies between standardized test texts and academic reading materials can offer clearer explanations for common learner difficulties. This alignment not only supports more effective teaching and assessment practices but also contributes to a more reliable evaluation of reading comprehension skills as well ([Alderson, 2000](#); [Zuhra, 2015](#)).

Literature Review

Factors Influencing Reading: Texts and Learners

Syntactic complexity in academic reading texts is often viewed as a relevant factor that can influence reading comprehension and contribute to the overall difficulty of texts. It encompasses the arrangement of clauses, phrases, and other grammatical elements, which may pose challenges for students and potentially affect their ability to grasp the author's message in reading texts ([Basilan & De Sagun, 2024](#)). As one of several components contributing to text difficulty, syntactic complexity is often discussed alongside lexical richness ([Karami & Salahshoor, 2014](#)). Academic texts tend to exhibit higher syntactic complexity, particularly in scholarly works such as research papers where clausal complexity is often greater than in learner texts, although phrasal complexity may sometimes be comparable or lower ([Vinogradova et al., 2020](#)). Additionally, the ability to process complex syntactic structures has been suggested as a potential predictor of academic reading performance, which can influence comprehension in contexts such as English for Academic Purposes (EAP) courses ([Karami & Salahshoor, 2014](#); [Wijanti, 2017](#)). However, mismatches between instructional materials or textbook grade levels and syntactic demands in real-life texts that students encounter could hinder student comprehension, underscoring the need for careful material selection and

instructional strategies that consider the syntactic demands of different genres and text sections ([Beers & Nagy, 2011](#); [Indrawan, 2018](#)). Mentioning this importance, a recent study even suggested tailored instructional interventions to potentially help students overcome difficulties associated with syntactic complexity and improve their comprehension skills ([Basilan & De Sagun, 2024](#)).

L2 reading comprehension is a multidimensional process shaped by the interaction of reader-related and text-related factors ([Grabe & Stoller, 2019](#); [Koda, 2005](#)). Reader-related factors include linguistic proficiency, such as grammar and vocabulary knowledge, as well as background knowledge ([Qian, 2002](#)). Individual differences, such as anxiety and reading fluency, further influence the reading experience ([Klauda & Guthrie, 2008](#); [Sellers, 2000](#)). These factors play a crucial role in determining how well a reader engages with and understands a text. Text-related factors, including content, length, grammar, and genre affect text comprehension ([Alderson, 2000](#); [Nation, 2006](#)). The linguistic features and structure of the text can either facilitate or hinder understanding, depending on how well they match the reader's language skills and prior knowledge.

Reading Process

Reading comprehension as a multifaceted cognitive process requires readers to access word meanings, semantic and syntactic ties, activate relevant prior knowledge, and construct coherence while reading a text ([van Dijk & Kintsch, 1983](#)). This process operates on two interrelated levels: literal comprehension, which focuses on understanding explicit textual information, and inferential comprehension, which involves drawing connections between background knowledge and implicit textual cues to construct deeper meaning ([Alptekin & Erçetin, 2010, 2011](#)). While literal comprehension establishes a foundational understanding, inferential comprehension demands advanced cognitive engagement to interpret unstated relationships and implications within the text.

Successful reading comprehension relies on the interaction between bottom-up and top-down processes. Bottom-up processes, including word recognition and word-to-text integration, involve identifying individual words, retrieving their meanings, and integrating them into larger syntactic and semantic structures in the texts ([Fender, 2001](#); [Perfetti & Hart, 2001](#)). Automatization of these lower-order processes is crucial, as it reduces cognitive load, allowing readers to allocate more mental resources to higher-order inferential tasks ([Just & Carpenter, 1992](#); [Perfetti, 1985](#)). In contrast, and after bottom-up processes, top-down processes rely on readers' background knowledge and contextual understanding to interpret meaning beyond the explicit text, enabling them to construct a situational model of the content ([Grabe & Stoller, 2019](#)). Effective comprehension emerges from the interplay between these two processes, with each reinforcing the other.

In an EFL context, Yazdi, and Mohammadian ([2022](#)) investigated the relationship between intermediate EFL students' syntactic knowledge and their proficiency in productive skills (speaking and writing). Results indicated that syntactic knowledge had no significant influence on these skills, as the correlations for both writing and speaking were weak and non-meaningful. This suggested that while syntax may hold some relevance, its impact on productive skills is minimal. Consequently, it can open the hypothesis that syntactic knowledge may play a more significant role in receptive skills, such as reading and listening. With this regard, and in the case of more advanced language learners, [Ghorbani Shemshadsara et al. \(2022\)](#) explored the effects of multi-component training involving grammatical awareness and self-regulated strategies on Iranian upper-intermediate EFL learners' receptive skill of reading comprehension. Using a cognitive load theory framework, 120 undergraduate students were divided into a control group and three experimental groups receiving different interventions over 12 weeks. Results showed that text structure/syntactic awareness along with self-regulated strategies significantly improved reading comprehension of the learners.

Based on the related literature, an essential factor influencing comprehension is syntactic complexity, particularly in advanced reading materials. This is because texts with deeper grammatical structures demand more cognitive effort and advanced syntactic knowledge, often posing challenges for readers ([Fender, 2001](#); [Nation & Snowling, 2000](#)). On the other hand, exposure to linguistically

appropriate and well-structured texts can play an appreciable role in overcoming these challenges, as regular engagement with syntactically rich input supports both lexical and syntactic development (Ellis, 2002; Grabe & Stoller, 2019). As stated earlier, this knowledge can relieve cognitive load, allowing for more inferential reading comprehension processes while reading a text. As a result, text related factors such as the level of syntactic complexity can be regarded as meaningful determinants of whether a text is effective and aligns with a learner's proficiency levels (Fulcher, 1997; Jin et al., 2020). While reader-related factors, such as motivation and prior experience, have meaningful influences on reading comprehension, they are often less controllable in empirical research. Consequently, greater focus can be placed on text-related variables, which are related to reading comprehension outcomes (Liontou, 2015).

Advanced academic texts often contain more complex syntactic structures, which may present challenges for learners who lack a solid grammatical foundation (Morvay, 2012; Zuhra, 2015). Research on textbook evaluation has suggested that many EFL textbooks may not adequately address critical local text structures, such as signalling and referential words, which can hinder the development of a solid syntactic knowledge. As a result, teachers feel the need to supplement textbooks which are careful on these aspects (Bogaerds-Hazenberg et al., 2022). This absence of complex grammar in EFL reading materials limit students' preparedness for advanced academic tasks, as observed in studies from Saudi Arabia, where learners appeared to be insufficiently exposed to the syntactic complexity potentially required for higher-level academic performance (Alenezi, 2016).

Reading Comprehension and Reading Texts through the Lens of Standardized Testing and

Instructional materials

As another type of academic text, reading comprehension exams have been the focus in literature as well. A significant body of research has examined the reading assessments, and the predictive validity of those reading tests has traditionally suggested that the texts used in them effectively assessed reading skills that equated future academic success. Hopkins and Sitkei (1969) found that reading readiness tests are more reliable predictors of academic success, showing significant correlations with end-of-year teacher marks and standardized reading test scores, thus reinforcing their critical role in early academic achievement. Similarly, Bagford (1968) established significant correlations ranging from .16 to .72 at the .01 or .05 confidence levels between reading readiness test scores and later reading success among 150 students in Iowa City Public Schools, highlighting the importance of early assessments in identifying instructional needs. Bremer (1959) found a correlation of 0.40 between first-grade Metropolitan Readiness Test scores and second-grade performance on the Gray-Votaw-Rogers General Achievement Test, while Powell and Parsley (1961) reported an even stronger correlation of 0.82 between Lee-Clark readiness scores and vocabulary-comprehension scores from the California Reading Test at the beginning of second grade. Collectively, these findings underscore the foundational role of pre-literacy skills in shaping long-term academic outcomes.

On the contrary, and more recently, the study by Liu and Li (2023) indicated that the reading texts used in M.A. entrance exams did not effectively predict success in parallel academic contexts, such as IELTS reading comprehension. Chinese students encountered specific challenges in IELTS, including nuanced language and complex sentence structures that differed from those in their entrance exams. This suggested that the skills assessed in M.A. entrance exams did not fully align with those required for IELTS proficiency. Consequently, while there might have been some overlap, the differences in text complexity and reading demands implied that performance on M.A. entrance exams might not reliably forecast success in IELTS reading comprehension.

Further, investigations into the reading components of standardized English proficiency tests appear to raise questions about their effectiveness, and reliability in forecasting real-life academic performance. Johnson and Tweedie (2021) analyzed data from 1,918 post-secondary students over seven years and found that standardized test scores explained only 4-6% of the variance in GPA, indicating weak predictive power. Other studies similarly report a small correlation between IELTS

scores and academic performance, with reading scores being slightly more predictive than other components but still offering limited insight into real-world academic success ([Gagen & Faez, 2024](#); [Kerstjens & Nery, 2000](#)). While the IELTS reading section is designed to assess comprehension and synthesis of information, its ability to predict actual academic tasks has been questioned ([Marina, 2018](#); [Mauriyat, 2021](#)). Moreover, IELTS preparation courses, though capable of improving test scores, do not necessarily equip students with the full range of skills required for academic study, as certain essential competencies, such as critical reading and deep text engagement, may not be adequately covered ([Dang & Dang, 2023](#); [Marina, 2018](#)). Although strategies like skimming and scanning are useful for the IELTS reading test and can aid in academic reading, the exam itself may not sufficiently prepare students for the depth and analytical engagement needed in university settings ([Fitria, 2024](#); [Marjerison et al., 2020](#)). Thus, while IELTS reading tests provide a structured measure of language proficiency, they remain limited in predicting students' academic success and fully preparing them for the demands of higher education.

Moreover, research on the effectiveness of English textbooks as learning materials suggests that their syntactic complexity may not adequately prepare students for the demands of academic reading. Studies indicate that high school English textbooks generally exhibit lower syntactic complexity than university entrance exams, highlighting a gap that may leave students unprepared for the linguistic challenges of higher education ([Gedik & Kolsal, 2022](#); [Kim & Oh, 2019](#)). This discrepancy can lead to a negative backwash effect, where students, accustomed to less complex texts, struggle with the advanced structures found in exam materials ([Gedik & Kolsal, 2022](#)). Additionally, inconsistencies in the progression of syntactic complexity within textbooks have been identified, with some lower-grade textbooks (e.g., grade 10) demonstrating greater complexity than those for higher grades (e.g., grade 12), suggesting a lack of systematic scaffolding in language development ([Indrawan, 2018](#); [Verdiansyah, 2020](#)). While some textbooks show a structured and gradual increase in complexity, others fail to maintain this progression, potentially hindering students' readiness for more sophisticated academic texts ([Putra & Lukmana, 2017](#); [Yang & Bae, 2022](#)). These findings underscore the need for textbook revisions that ensure a consistent and developmentally appropriate increase in syntactic complexity, aligning with students' cognitive growth and better equipping them for university-level reading ([Indrawan, 2018](#); [Verdiansyah, 2020](#)).

As far as the authors reviewed in this study, it seems that, much of the existing research has primarily focused on reader-related challenges in reading comprehension literature. And, there appears to be a relevant gap in examining text-related factors, particularly the syntactic characteristics of reading materials used in learner input. Despite the apparent importance of these factors, and the enhanced importance of reading materials in EFL contexts, there has been relatively little comparative analysis across different types of EFL learning materials, such as textbooks, standardized tests, and academic research papers. This gap could be significant, as understanding how syntactic demands vary across these sources can provide valuable insights into the challenges learners face when entering universities. The present study intends to address this gap by analyzing the syntactic complexity within frequent learner input, identifying potential inconsistencies, and exploring how these discrepancies might have challenged learners' comprehension. Ultimately, this research could contribute to a broader understanding of syntactic complexity in advanced texts, suggesting more implications for test developers and material designers. To this aim, the current study poses the following question:

Research Question One: How does the syntactic complexity of reading comprehension texts in EFL textbooks, standardized tests, and research papers compare?

Methodology

Design of the Study

The current research adopted a non-experimental quantitative design to evaluate syntactic complexity of texts in samples of EFL learner input. This quantitative design, further, aimed to discover frequent patterns of syntactic complexity in advanced texts genre through interpretation and discussion of the results.

Instruments

The primary instruments to answer the research questions of this study were text document data, and linguistic text analysis and evaluation tool of L2 Syntactic Complexity Analyzer (Lu, 2010). Four corpora of advanced reading texts were compiled based on the sources of the texts (i.e., educational textbooks, standardized tests, and academic research papers). The texts were, then, fed into the linguistic tools for detailed syntactic complexity and readability analyses.

The Corpora Analyzed in the Study

Educational textbooks in Iran as an EFL country (First corpus): The *Prospect Series* was introduced by the Ministry of Education in Iran since 2012 and aimed at developing communicative competence of learners at the junior high school level. *Prospect Series* is part of the *English for Schools* series, prepared to be taught in junior high schools of Iran. It involves three volumes including, *Prospect 1*, *Prospect 2*, and *Prospect 3*. On the other hand, three other textbooks in the *English for Schools* series are targeted to senior high school and comparatively more advanced level which are entitled *Vision 1*, *Vision 2*, and *Vision 3* (Kheirabadi & Alavimoghaddam, 2016). *Vision* series course books are designed in 2016 to for the Iranian high school students who are in age range of 16-18. The pattern and structure of the books are similarly designed throughout the series: each book includes four lessons. Listening, speaking, reading, and writing skills are combined in these coursebooks. All of the lessons start with 'Get Ready' section which is aimed to provide introduction to the lesson. 'Conversation' and 'New Words' sections carry the new words of each lesson. Other sections in each lesson include 'Reading', 'Grammar', 'Listening', 'Speaking', 'Pronunciation' and 'Writing'. Finally, 'what you learned' section reviews the lesson, helping students internalize what they have learned (Saeedi & Shahrokhi, 2019). Since *Vision* is concerned with secondary high school studies, and considered comparatively as a type of more advanced educational textbook, reading texts in these textbooks are chosen for complexity analysis in this study. *Vision* coursebook series include *Vision 1*, *Vision 2*, and *Vision 3*. They contain 4, 3, and 3 lessons respectively. Each lesson contains one reading passage (N=10). Table 1 illustrates names of the lessons as well as reading comprehension texts included in the textbooks, which are included in the corpus. It should also be mentioned that these textbooks cannot be considered as including similarly advanced texts compared to IELTS text books or academic research papers. However, as a sample of EFL learner input in Iran, which is being taught in highest secondary high school level in this country, the texts in these books can be regarded as relatively more advanced compared to those taught in lower levels (i.e., *Prospect* coursebooks). And, therefore, they are expected to present a level of syntactic complexity, while not as dense as research papers, more representative of advanced texts. *Vision* textbooks are the most advanced English texts students read before entering universities. They are also regarded as a frequent sample of learner input in Iran as an EFL country, where, according to the literature, print material play a crucial role in language development.

Table 1

Vision Coursebook Series' Lessons and Reading Comprehension Texts

Vision Coursebook	Lesson	Reading Section
Vision 1	1: Saving Nature	Endangered Animals
	2: Wonders of Creation	A wonderful Liquid
	3: The Value of Knowledge	No Pain No Gain
	4: Travelling the World	Iran: A True Paradise
	1: Understanding People	Languages of the World

Vision 2	2: A Healthy Lifestyle	Having a Healthier and Longer Life
	3: Art and Culture	Art, Culture, and Society
	1: Sense of Appreciation	Respect your Parents
Vision 3	2: Look it Up!	How to Use a Dictionary
	3: Renewable Energy	Earth for our Children
Total	10	10

Standardized tests (Second corpus): University and higher education candidates in Iran need to take part in Iranian national university exams, which are administered every year, in order to advance their educational studies and academic career. According to Razmjoo, and Heydari Tabrizi (2010), examinees' performance on such high-stakes tests has direct impact on their future life and academic studies. M.A. Teaching English as a Foreign Language University Entrance Exams (M.A. TEFL UEE) are held since 1990 in Iran, and consist of General and Major English questions. More information on the sub-sections is presented in Tables 2, and 3, respectively.

Table 2

M.A. EFL UEE Format: General English

Section	Sub-sections	Number of Items
1. Structure and Written Expressions	1.1 Sentence Completion	10 items
	1.2 Error Recognition	10 items
2. Vocabulary		10 items
3. Cloze Test		15 items
4. Reading Comprehension		25 items / 3 passages

Note: Retrieved from "A Content Analysis of the TEFL M.A. Entrance Examinations" by S. A. Razmjoo, and H. Heydari Tabrizi, 2010, Journal of Pan-Pacific Association of Applied Linguistics, 14(1), p. 161. Copyright 2010 by the Pan-Pacific Association.

Table 3

M.A. TEFL UEE Format: Major English

Majors	Sub-parts	Number of Items
1. TEFL	1.1. Teaching Methodology	40 Items
	1.2. Testing	20 Items
	1.3. Linguistics	20 Items
2. English Literature	No-subparts	80 Items
	3.1. Theoretical principles of translation	25 Items
3. Translation	3.2. Linguistics	15 Items
	3.3. Contrastive Analysis	10 Items
	3.4. Morphology	15 Items
	3.5. Translation skill	15 Items

Note: Retrieved from "A Content Analysis of the TEFL M.A. Entrance Examinations" by S. A. Razmjoo, and H. Heydari Tabrizi, 2010, Journal of Pan-Pacific Association of Applied Linguistics, 14(1), p. 161. Copyright 2010 by the Pan-Pacific Association.

General English items are organized, as shown in Table 2, into four different sections, and major English items are specified for candidates of each major: TEFL, English Literature, and Translation programs. It should be mentioned that, although candidates of each specific major answer only respective question items in Major English section, general English items are answered by all the candidates regardless of their specific major. Reading comprehension section in these standardized

tests, according to Table 3, includes three reading passages, without a topic, followed by comprehension questions. Accordingly, and to enable a comprehensive syntactic analysis, all the reading passages in these exams during the last decade are included in the corpus (N= 30 passages).

Cambridge IELTS Textbooks (Third Corpus): As another source of advanced reading comprehension texts, these books include 4 tests, and each test includes 3 reading passages. The books are considered as including authentic practice tests. For the purposes of the current study, a random sample of reading comprehension texts (N=30) in the academic version of these tests are included in the corpora. Table 4 presents more information on the passages chosen from the books.

Table 4

Texts Chosen from Cambridge IELTS Books

Book	Passage 1	Passage 2	Passage 3	Passage 4
IELTS 9	William Henry Perkin: The man who invented synthetic dyes	Venus in Transit	Information Theory – the big idea	
IELTS 10	The psychology of innovation: Why are so few companies truly innovate?	Gifted children and learning	The Context, Meaning and Scope of Tourism	The megafires of California
IELTS 11 Academic	Research using twins	Great Migrations	Raising the <i>Mary Rose</i>	
IELTS 12 Academic	What's the purpose of gaining knowledge?	The Lost City	Flying tortoises	The Benefits of Being Bilingual
IELTS 13 Academic	Artificial artists: Can computers really create works of art?	Oxytocin	The coconut palm	
IELTS 14 Academic	The secret of staying young	Saving bugs to find new drugs	The importance of children's play	
IELTS 15 Academic	What is exploration?	Should we try to bring extinct species back to life?	Henry Moore (1898-1986)	
IELTS 16 Academic	Plant 'thermometer' triggers springtime growth by measuring night-time heat	The white horse of Uffington	Changes in reading habits	Why we need to protect polar bears
IELTS 17 Academic	The second attempt at domesticating the tomato	To catch a king	Building the skyline: The Birth and Growth of Manhattan's Skyscrapers	
Total				30

Research papers (Fourth corpus): A random sample of 30 research papers published in different disciplines (from 2012 to 2022), as another source of EFL text material, are considered for the collection of the last advanced-texts corpora. Scopus subject areas were the criteria to choose journals and research papers. There are 26 subject areas in Scopus sources with various sub-area sections. Therefore, to compile the research papers corpus, one paper was chosen from each subject area, and to maintain the comparability of the corpora, four more random papers from Scopus-indexed journals were added to the collection (N=30). It should be mentioned that, the number of words and sentences in research papers is not comparable to those in reading comprehension passages, and they tend to include much more information and pages compared to reading texts in other corpora. Therefore, to maintain the comparability, only discussion sections of the research articles were compiled into research papers corpus. Discussion section of the papers are believed to entail new knowledge claims as an important aim of research articles ([Basturkmen, 2009](#)). Additionally, in academic research papers, discussion sections are argued to play a key role in interpreting the findings and contributing to theory and practice in disciplines ([Le & Harrington, 2015](#)). More information on the journals and research papers are presented in Appendix B. Although relatively more complex than texts in other

corpora, research papers are essential reading material for M.A. students, in any discipline, who have passed entrance exams.

Data Collection and Analysis Procedures

In this study, the development of the corpus adhered to several key characteristics essential for effective corpus design. First, following points mentioned by McEnery and Brooks (2022), the corpora were classified as a specialized type, focusing exclusively on advanced reading comprehension texts. To maintain authenticity, following McEnery and Wilson (2001), the corpus included real-life language data sourced from textbooks, high-stakes standardized tests, and academic research papers, ensuring that the texts reflected the typical encounters of EFL learners. Although achieving a perfect representative corpus was challenging due to common access limitations (McEnery & Brookes, 2022), efforts were made to ensure representativeness, as defined by Biber (1993), by collecting a diverse range of advanced texts frequently available to learners. For comparability, based on definitions from Ji (2009), and Hewavitharana and Vogel (2008), all texts were drawn from the advanced reading genre, with each corpus containing 30 passages, except for one corpus (Vision Coursebooks) with only 10 texts, which warranted careful interpretation of the results. Additionally, only discussion sections of research papers published within the last decade were included to maintain time comparability with other corpora. Detailed information regarding the corpora is given in Appendix A.

The texts were analyzed using syntactic complexity measures of L2SCA. Table 3.5 shows the measures included in this tool (Lu, 2010). According to the previous literature, readers process the texts linearly, decoding it word by word; but as they read, they need to compile the linguistic items into a larger scale of syntactic structures (Just & Carpenter, 1987; Rayner & Pollatsek, 1996). Accordingly, the mental demands required for this operation can vary considerably on the basis of how complex the structure is (Perfetti et al., 2005). For these reasons, all 14 measures computed by the L2SCA are used in order to make the analysis a comprehensive one. Table 5 presents more information on the syntactic complexity measures.

Table 5

L2SCA Syntactic Complexity Measures

Number	Label	Description
Length of production unit		
1	MLC	Mean length of clause
2	MLS	Mean length of sentence
3	MLT	Mean length of T-unit
Amount of subordination		
4	C/T	Number of clauses per T-unit
5	CT/T	Complex T-unit ratio
6	DC/C	Number of dependent clauses per clause
7	DC/T	Number of dependent clauses per T-unit
Amount of coordination		
8	CP/C	Number of coordinate phrases per clause
9	CP/T	Number of coordinate phrases per T-unit
10	T/S	Number of T-units per sentence
Degree of phrasal sophistication		
11	CN/C	Number of complex nominals per clause
12	CN/T	Number of complex nominals per T-unit
13	VP/T	Number of verb phrases per T-unit
Overall sentence complexity		
14	C/S	Number of clauses per sentence

Note: Retrieved from "Automatic analysis of syntactic complexity in second language writing" by X. Lu, 2010, International Journal of Corpus Linguistics, 15(4): 474-496. Copyright 2010 by John Benjamins Publishing Company.

These measures can be broadly classified into four categories: length of production unit, amount of subordination, amount of coordination, and degree of phrasal sophistication. The length of production unit is assessed through measures such as mean length of clause (MLC), mean length of sentence (MLS), and mean length of T-unit (MLT). These indices provide insights into the overall sentence complexity, as readers need to process linguistic items into larger syntactic structures. The amount of subordination is captured through measures like number of clauses per T-unit (C/T), complex T-unit ratio (CT/T), number of dependent clauses per clause (DC/C), and number of dependent clauses per T-unit (DC/T). These indices reflect the degree of syntactic embedding and the mental demands required for processing complex structures (Lu, 2010; Perfetti et al., 2005).

The amount of coordination is assessed through the number of coordinate phrases per clause (CP/C), number of coordinate phrases per T-unit (CP/T), and number of T-units per sentence (T/S). These measures provide information about the level of coordination within the text. Finally, the degree of phrasal sophistication is evaluated through the number of complex nominals per clause (CN/C). This index reflects the complexity of noun phrases, which can impact the overall processing demands on the reader. L2SCA was used in the current study because it offered a comprehensive framework for analyzing text complexity by examining various aspects of syntactic complexity, including length of production unit, amount of subordination, amount of coordination, and degree of phrasal sophistication.

All of the 14 measures in Table 5 were analyzed and treated as a single measure to present all the details with regard to the syntactic complexity and descriptive analysis of the corpora. Building on this, the study by Ai and Lu (2013) provided the next step for a structured framework for analyzing syntactic complexity by grouping measures into four distinct categories: length of production units, amount of subordination, amount of coordination, and degree of phrasal sophistication. They compared texts by examining differences in the mean values of these grouped measures across multiple writing samples. This grouping allowed for a more focused and systematic analysis of syntactic patterns. The study highlighted statistically significant differences between groups, revealing how syntactic complexity varied across proficiency levels and text types. This method demonstrates the effectiveness of categorizing syntactic complexity measures to uncover patterns and relationships inferentially in the corpora. Similarly, in the current study, the syntactic complexity measures were systematically grouped into four distinct categories to enable effective statistical analysis. The groupings are presented in Table 6. Accordingly, this approach aligns with established methodologies in prior research, facilitating a structured examination of syntactic complexity across different text sources. By categorizing measures into these specific groups, the data were prepared for comparative analysis of syntactic patterns. This strategy can enable the identification of meaningful differences and trends in syntactic complexity among the corpora under investigation.

Table 6

L2 Syntactic Complexity Measures Groupings

Group	Label	Description
Length of Production Unit	MLC	Mean length of clause
	MLS	Mean length of sentence
	MLT	Mean length of T-unit
Amount of Subordination	DC/C	Number of dependent clauses per clause
	DC/T	Number of dependent clauses per T-unit
Amount of Coordination	CP/C	Number of coordinate phrases per clause
	CP/T	Number of coordinate phrases per T-unit
	T/S	Number of T-units per sentence
Degree of Phrasal Sophistication	CN/C	Number of complex nominals per clause
	CN/T	Number of complex nominals per T-unit

Methodological Procedure for the Inferential Analyses

To examine variations in syntactic complexity across four corpora, a one-way MANOVA (Multivariate Analysis of Variance) was conducted using SPSS (Statistical Package for the Social Sciences). This analytical approach was selected for its ability to assess multiple dependent variables simultaneously while accounting for their interdependence, providing a comprehensive examination of syntactic complexity across the corpora.

Variables and Data Organization

The analysis incorporated four dependent variables, representing distinct dimensions of syntactic complexity:

1. Length of production unit, indicating overall syntactic elaboration.
2. Amount of subordination, measured as the frequency of dependent clauses relative to other units.
3. Amount of coordination, reflecting the extent of coordinate structures.
4. Degree of phrasal sophistication, which captures the intricacy of phrasal elements.

The independent variable, corpus, consisted of four categories, each corresponding to a distinct textual source. Data preparation involved calculating the syntactic complexity indices for all samples within each corpus, ensuring consistency and comparability across groups, the descriptive report of which is already presented above.

To ensure reliable results, the analysis treated each corpus as independent, with no overlap or dependency between the groups. Each corpus was carefully organized to reflect the unique characteristics of specific textbooks and research papers. MANOVA was chosen for its ability to analyse multiple related variables at once, making it ideal for exploring differences in syntactic complexity across the academic fields of Psychology, Pharmacy, and Accounting, as reflected in their textbooks and research papers corpora. This approach was used to highlight how syntactic features vary between these written genres.

Results and Discussion

Descriptive Analysis of the Results

Vision Coursebooks (Corpus A)

The reading comprehension texts from the Vision Coursebooks showed the lowest score on every aspect, not even remotely comparable to other corpora:

- Sentence and Clause Length: Sentences average 12.16 words, with T-units slightly shorter (11.62 words) and clauses even more concise (9.14 words).
- Clause and Phrase Ratios: Sentences include 1.33 clauses on average, indicating modest subordination. T-units primarily contain single clauses (1.27 clauses/T-unit).
- Dependent Clauses: Dependent clauses are infrequent (0.21 per clause, 0.27 per T-unit), reflecting limited embedding.
- Complexity Indicators: Roughly a quarter of T-units are complex (0.26 complex T-units/T-unit). Nominal structures are common, with 1.04 complex nominals per T-unit and 1.32 per clause, enhancing syntactic richness.

These measures highlighted that the Vision Coursebooks received the lowest score, indicating a level of complexity that hardly compared to that observed in other materials. Following data and comparisons together with the plots given at the end of this section can challenge Vision Coursebooks

on the basis of their weak representativeness of syntactic complexity in reading comprehension content. Although the books are developed for the school students, they are the most advanced language data they read in Iranian secondary high schools before entering universities. This discrepancy can be one of the reasons of the linguistic challenges students encounter in university studies, as reviewed in the related literature in EFL settings. Therefore, it is reasonable to expect the books to exhibit a level of syntactic complexity—understandably less dense than that of M.A. tests or research papers—yet more closely aligned with the complexity requirements of other academic sources students encounter throughout their academic careers.

M.A. TEFL University Entrance Exams in Iran (Corpus B)

The texts in Corpus B display moderate syntactic complexity:

- Sentence and Clause Length: Sentences average 19.8 words, while T-units (18.1 words) and clauses (10.9 words) are similarly extended.
- Clause and Phrase Ratios: Sentences commonly include multiple clauses (1.80 clauses/sentence), and T-units average 1.65 clauses, reflecting greater subordination.
- Dependent Clauses: Dependent clauses are more frequent (0.39 per clause, 0.66 per T-unit).
- Complexity Indicators: Nearly half of T-units are complex (0.48 complex T-units/T-unit), and nominal elaboration is pronounced (2.46 complex nominals/T-unit, 1.48 per clause).

IELTS Reading Comprehension Texts (Corpus C)

IELTS reading texts exhibit higher syntactic complexity than Corpus A and B. This suggested that M.A. reading texts in Iranian examinations are not syntactically comparable to IELTS readings:

- Sentence and Clause Length: Sentences are longer (22.9 words), with T-units (20.6 words) and clauses (11.1 words) reflecting similar complexity.
- Clause and Phrase Ratios: Sentences contain multiple clauses (2.05 clauses/sentence), and T-units frequently include subordination (1.84 clauses/T-unit).
- Dependent Clauses: Dependent clauses are moderately frequent (0.43 per clause, 0.79 per T-unit).
- Complexity Indicators: Over half of the T-units are complex (0.54 complex T-units/T-unit). Nominal structures are robust (2.83 complex nominals/T-unit, 1.53 per clause), underscoring syntactic sophistication.

Discussion Sections from Scopus Journals (Corpus D)

Corpus D displayed the highest level of syntactic complexity among the analyzed corpora, indicating that none of the previous corpora can represent the syntactic characteristics of real-life academic texts:

- Sentence and Clause Length: Sentences average 24.7 words, with T-units (23.2 words) and clauses (14.4 words) reflecting dense structures.
- Clause and Phrase Ratios: Sentences contain 1.71 clauses on average, while T-units include significant subordination (1.61 clauses/T-unit).
- Dependent Clauses: Dependent clauses are used frequently (0.37 per clause, 0.60 per T-unit).
- Complexity Indicators: Nearly half of the T-units are complex (0.44 complex T-units/T-unit), and nominal elaboration is extensive (3.52 complex nominals/T-unit, 2.18 per clause).

Table 7 is presented to show the results of the analysis in a cleared way while more appropriate for statistical analysis.

Table 7

Syntactic Complexity Results of the Corpora

Corpus	ML S	ML T	ML C	C/ S	VP/ T	C/ T	DC/ C	DC/ T	T/S	CT/ T	CP/ T	CP/ C	CN/ T	CN/ C
IELTS	22.9 1	20.6 4	11.1 7	2.0 5	2.52	1.8 5	0.43	0.79	1.1 1	0.54	0.47	0.25	2.84	1.54
IRMA Uni	19.8 1	18.1 8	10.9 7	1.8 1	2.18	1.6 6	0.40	0.66	1.0 9	0.48	0.47	0.29	2.46	1.49
Scopus (Paper s)	24.7 0	23.2 3	14.4 0	1.7 2	2.23	1.6 1	0.37	0.60	1.0 6	0.44	0.68	0.42	3.52	2.18
Vision (Books)	12.1 6	11.6 2	9.14	1.3 3	1.61	1.2 7	0.21	0.27	1.0 5	0.26	0.40	0.31	1.32	1.04

Inferential Analysis of the Results

As explained in the methodology section, and according to the below table, the data underwent a systematic transformation to align with the newly established grouping framework. This reorganization facilitated a more targeted statistical analysis of syntactic complexity, grouping measures with similar linguistic functions to enhance the interpretation of results across different text types. Table 8 is presented to the mention the groupings:

Table 8

Grouping of the Syntactic Complexity Variables

Group	Measure	Label	Description	IELTS	IRMA Uni	Scopus (Papers)	Vision (Books)
Length of Production Unit	1	MLC	Mean length of clause	11.17	10.97	14.40	9.14
	2	MLS	Mean length of sentence	22.91	19.81	24.70	12.16
	3	MLT	Mean length of T-unit	20.64	18.18	23.23	11.62
Amount of Subordination	4	DC/C	Number of dependent clauses per clause	0.43	0.40	0.37	0.21
	5	DC/T	Number of dependent clauses per T-unit	0.79	0.66	0.60	0.27
Amount of Coordination	6	CP/C	Number of coordinate phrases per clause	0.25	0.29	0.42	0.31
	7	CP/T	Number of coordinate phrases per T-unit	0.47	0.47	0.68	0.40
Degree of Phrasal Sophistication	8	T/S	Number of T-units per sentence	1.11	1.09	1.06	1.05
	9	CN/C	Number of complex nominals per clause	2.84	2.46	3.52	1.32

10	CN/T	Number of complex nominals per T-unit	1.54	1.49	2.18	1.04
----	------	---------------------------------------	------	------	------	------

The findings of the MANOVA are presented in Table 9.

Table 9

Results of the Multivariate Tests

Effect	Multivariate Tests ^a					
		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	1.000	249813.713 ^b	4.000	1.000	.002
	Wilks' Lambda	.000	249813.713 ^b	4.000	1.000	.002
	Hotelling's Trace	999254.852	249813.713 ^b	4.000	1.000	.002
	Roy's Largest Root	999254.852	249813.713 ^b	4.000	1.000	.002
corpus	Pillai's Trace	2.399	2.997	12.000	9.000	.054
	Wilks' Lambda	.000	59.938	12.000	2.937	.003
	Hotelling's Trace	.	.	12.000	.	.
	Roy's Largest Root	50573.689	37930.266 ^c	4.000	3.000	.000

a. Design: Intercept + corpus

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

The MANOVA results indicated that the text sources (corpora) had a statistically significant effect on the combined syntactic complexity measures, as shown by Wilks' Lambda ($F(12, 2.937) = 59.938, p = .003$). This demonstrated that the linguistic characteristics of the text sources—such as, any or all, length of production units, amount of subordination, amount of coordination, and degree of phrasal sophistication—vary significantly across the groups. The findings suggested that different text sources were designed with distinct syntactic features, which may be the reason for linguistic challenges language learners and students face in EFL university studies. However, this discrepancy can partly be explained with Biber and Gray's (2010) observation that linguistic features like subordination, coordination, and phrasal sophistication vary across academic genres to suit their communicative purposes.

The univariate results also broke down the multivariate significance by testing each dependent variable separately. Table 10 provided the respective results. Further, Table 11 provided group means and standard deviations for each dependent variable across the 4 corpora. However, for space considerations, the tables are given in the Appendix C.

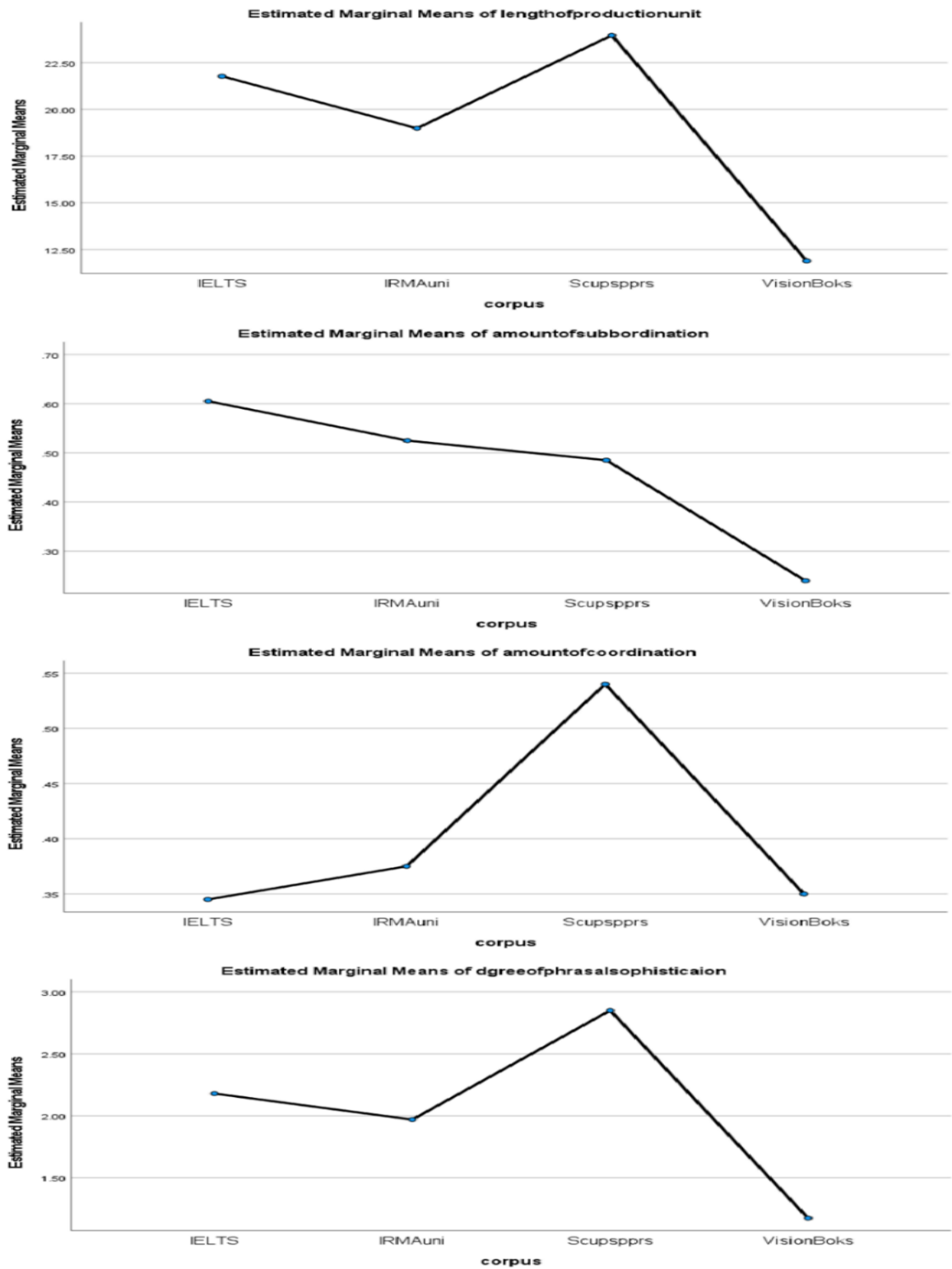
The MANOVA results showed significant and non-significant differences in syntactic complexity across the four corpora. Among the four dependent variables, length of production units demonstrated a statistically significant difference, with $F(3, 4) = 42.899, p = .002$. Descriptive statistics revealed that Scopus research papers (mean = 23.965) and IELTS texts (mean = 21.775) contained the longest production units, reflecting their reasonable alignment with advanced academic written texts. The R-squared value of 0.970 indicated that 97% of the variance in sentence length can be attributed to the text source, underscoring its relevance as a distinguishing feature of these corpora.

For the other measures—amount of subordination, amount of coordination, and degree of phrasal sophistication—no statistically significant differences were found across the corpora. This suggested that while subordination, coordination, and phrasal sophistication are important features of advanced English texts, their usage may not vary dramatically between these specific text types. Descriptive statistics, however, showed that Scopus papers and IELTS texts generally had higher levels of subordination (means = 0.485–0.605) and phrasal sophistication (means = 2.180–2.850) compared to Vision textbooks and M.A. tests. This was in line with the findings of Liu and Li (2023), who indicated that the reading texts used in M.A. entrance exams did not effectively predict success in parallel academic contexts, such as IELTS reading comprehension. However, the lack of significance might be attributed to the small sample size or overlapping variability among the groups as well. The

respective plots are given in Figure 1 which manifest the means of syntactic complexity variables across the corpora.

Figure 1

Plots for the Syntactic Variables of the Study



The first plot, which depicted the estimated marginal means of the length of production units, showed that the Scopus research papers had the highest values, followed by IELTS and IRMAuni, while VisionBoks demonstrated the lowest values. This suggested that Vision coursebooks used simpler and shorter sentence structures, whereas Scopus papers required greater syntactic complexity. The second plot illustrated the amount of subordination. Here, IELTS texts exhibited the highest level of subordinate clause use, followed by IRMAuni and Scopus papers, while VisionBoks employed the least amount of subordination. The third plot highlighted the amount of coordination. Scopus research papers exhibited a notable peak in coordination, indicating frequent use of coordinated clauses, while VisionBoks showed minimal coordination. IELTS and IRMAuni fell between these two extremes, demonstrating moderate levels of coordination.

The fourth plot focused on the degree of phrasal sophistication across the four corpora. The results indicated relatively low and stable levels of sophistication for IELTS and IRMAuni, with both maintaining similar values around 2.00. In contrast, Scopus research papers displayed a notable increase in sophistication, reaching approximately 3.00, the highest among the corpora. Vision coursebooks, however, exhibited a sharp decline, with sophistication levels dropping to significantly lower values, well below 2.00.

The comparison of syntactic complexity across the four corpora revealed that, although all were labelled as advanced, they were neither comparable nor truly representative of the academic texts students actually encountered. The notable variations in syntactic structures, levels of subordination, and nominal elaboration across the corpora suggested that these materials diverged significantly from each other. This discrepancy can raise concerns about relying on these sources as accurate benchmarks for academic performance, as they appeared to capture only a limited range of the linguistic demands present in real-world academic contexts.

In contrast, Norris and Ortega (2009), similarly, highlight subordination as a critical marker of advanced writing, while McNamara et al. (2010) argue that subordination frequency alone may not reliably indicate text quality or complexity. Research has demonstrated that academic writing and reading exhibits significant variation, influenced by disciplinary conventions that shape lexicogrammatical choices, structural organization, and argumentation models (Benelhadj, 2019; Samraj, 2005). For instance, research articles tend to share similarities across disciplines, but other genres, such as PhD theses, reflect greater personal and disciplinary differences (Benelhadj, 2019). Furthermore, while subordination is often highlighted as a marker of complexity (Norris & Ortega, 2009), it is insufficient as a standalone measure, as McNamara et al. (2010) argue that subordination frequency does not reliably indicate overall text quality or complexity. In fact, disciplinary differences in argumentation models—such as the use of premise-based arguments in philosophy, hypothesis-driven arguments in computational science, and exposition-based arguments in chemistry—illustrate that syntactic complexity cannot be universally measured or benchmarked against research papers alone (Walková & Bradford, 2022). Therefore, attempts to equate only the complexity of standardized tests like IELTS or M.A. TEFL texts to research papers will overlook the broader variations in writing styles, structures, and purposes, which are central to understanding academic writing diversity (Dong et al., 2023; Moran, 2013)

Traditional literature has equated reading test readiness with academic performance, as evidenced by studies such as Hopkins and Sitkei (1969), Bagford (1968), Bremer (1959), and Powell and Parsley (1961), which have demonstrated significant correlations between early reading readiness assessments and subsequent academic success. However, more recent research has cast doubt on the validity and representativeness of tests as well as textbooks, with Liu and Li (2023) questioning the predictive power of M.A. entrance exam reading texts for IELTS success, Johnson and Tweedie (2021) demonstrating weak correlations between standardized test scores and academic performance, and studies such as Gagen and Faez (2024) and Kerstjens and Nery (2000) further challenging the predictive accuracy of IELTS reading scores. Similarly, Marina (2018) and Mauriyat (2021) have critiqued the ability of standardized reading assessments to reflect real-world academic tasks, while Dang and Dang (2023) and Fitria (2024) argue that IELTS preparation courses do not fully equip students with critical academic reading skills. Additionally, research on textbooks, including studies

by Gedik and Kolsal (2022) and Kim and Oh (2019), has highlighted gaps in syntactic complexity that may hinder students' preparedness for university-level reading, with further evidence from Indrawan (2018) and Verdiansyah (2020) suggesting inconsistencies in complexity progression across grade levels.

Collectively, these studies implied that standardized reading tests and instructional materials did not reliably measure or develop the full range of reading skills necessary for academic success. While traditional assessments had been assumed to predict future performance, recent research suggested that their validity was limited, particularly in preparing students for real-world academic tasks that required deeper comprehension, critical reading, and engagement with complex texts. Similarly, inconsistencies in textbook design contributed to a gap between secondary education and higher education reading demands, underscoring the need for revisions in both testing frameworks and instructional materials to better align with academic literacy requirements. Building on this body of research, the present study contributed to the field by evaluating the syntactic complexity of advanced reading materials encountered in schools, testing conditions, and academic studies. The findings revealed notable discrepancies, highlighting the need for greater alignment and improved representativeness in these materials. By increasing awareness of such disparities, this study underscored the importance of addressing the misrepresentation of syntactic complexity in learner input. Enhancing the alignment between instructional materials, testing texts, and academic readings could better equip students to navigate the linguistic challenges they face in both academic studies and standardized assessments.

Conclusion and Implications

The analysis of syntactic complexity across various corpora highlighted discrepancies that necessitated a re-evaluation and reconsideration of syntactic complexity across sources of learner input and testing contents. The Vision Coursebooks (Corpus A) exhibited the lowest levels of syntactic complexity, with an average sentence length of only 12.16 words and minimal subordination, indicating a stark inadequacy in preparing students for the linguistic demands of academic writing. In contrast, the M.A. TEFL University Entrance Exams (Corpus B) demonstrated moderate complexity, with an average sentence length of 19.8 words and a greater use of subordination, yet still fell short of the expectations set by sources of advanced texts. IELTS texts (Corpus C) and Scopus research papers (Corpus D) showcased significantly higher levels of syntactic sophistication, however with some differences. The notable variations in syntactic structures, levels of subordination, and nominal elaboration across the texts underscored the discrepancies among them, raising concerns about the effectiveness and representativeness of these materials as reliable benchmarks for academic performance.

The findings underscored an important gap between the types of language skills measured by standardized tests such as IELTS and TOEFL, and the authentic academic demands posed by research papers. While standardized tests evaluated proficiency through reading comprehension tasks, they might not fully encompass the range of syntactic and cognitive skills required to engage with scholarly texts. Research papers often presented more sophisticated syntactic structures and complex cognitive processes, demanding higher-order critical thinking, synthesis, and analysis. Thus, a successful performance on these proficiency tests seemed not necessarily equate to readiness for academic written texts, where learners must navigate more intricate syntactic constructions and engage in deeper levels of analysis. This discrepancy called for a rethinking of how academic proficiency is conceptualized and assessed, suggesting that test design and material development should take into account the complexities of authentic academic texts and better representativeness of the materials.

Future research could address the limitations of the current study, such as the small sample size, which may have reduced statistical power for detecting differences in measures like

subordination and phrasal sophistication. Additionally, examining complementary measures like lexical density or readability measures could provide a more comprehensive understanding of text complexity across these sources. Another limitation of this study is its focus on only a subset of English proficiency texts, which may not fully represent the range of syntactic complexity found across other text types or contexts. For example, texts from different academic disciplines were not considered, which could provide additional insights into syntactic variation. Furthermore, the analysis was restricted to specific syntactic measures, and other factors such as lexical diversity or pragmatic features may also contribute to overall text complexity and comprehension. Hulstijn (2011) similarly underscores the gap between standardized test constructs and the linguistic demands of authentic academic texts. Shaw and Liu (1998) further support the need for a scaffolded approach, transitioning from foundational materials to more complex texts.

This study's findings suggested a broader perspective on the relationship between syntactic complexity and text comprehension. They invited consideration of how syntactic complexity not only shaped reading skills but also influenced learners' broader academic development. By acknowledging the varying syntactic demands of different text types, while keeping in mind the linguistic and grammatical reading comprehension challenges in EFL settings reviewed in this study, the study emphasized the potential for educators to support students not only in their reading proficiency but also in preparing them for more complex forms of academic expression, such as writing and deeper critical analysis. The findings indicated that a focus on syntactic complexity could have been integrated into an approach to academic literacy, one that prepared learners to navigate both the structural and cognitive demands of advanced academic work, alongside the texts' content-specific demands.

EFL settings are, particularly, associated with the lack of necessary exposure to English, where reading material play a significant role in learner input (Ali et al., 2022; Bernal & Bernal, 2020; Lak, 2017; Nuttall, 1996). Therefore, it can be suggested that, syntactic inconsistencies between EFL learner input and grammatical demands of advanced texts in academic research papers or standardized tests might create linguistic challenges to their comprehension performance on respective texts.

Moreover, the current study is only concerned with syntactic features of advanced and academic discourse organization. Although the knowledge of complex syntax can significantly contribute to learners' understanding of advanced texts, other areas (vocabulary, or background knowledge, for example) can equally impact reading comprehension of the students. Therefore, in a more comprehensive study, comparatively more detailed linguistic and psychological along with contextual, and methodological variables are likely to offer inherently more objective findings.

Acknowledgement

The authors extend their sincere gratitude to the journal director and the editorial team for their patience, guidance, and thorough communication of the submission guidelines, as well as for kindly highlighting important considerations for author confidence. They also wish to acknowledge the editorial team and the anonymous reviewers of this study for generously dedicating their time to review and provide thoughtful feedback on the manuscript.

Declaration of Conflicting Interests

The authors hereby declare that they have no conflicts of interest, financial or otherwise, that could have influenced the research, authorship, or publication of this manuscript.

Funding Details

The authors affirm that no external funding was received to support the research, authorship, or publication of this manuscript.

References

- Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. *Automatic treatment and analysis of learner corpus data*, 59. <https://doi.org/https://doi.org/10.1075/scl.59.15ai>
- Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press. <https://doi.org/https://doi.org/10.1017/CBO9780511732935>
- Alenezi, S. (2016). The suitability of the EFL reading texts at the secondary and preparatory levels as a preparation for academic reading at first year university level in Saudi Arabia. <https://drepo.sdl.edu.sa/handle/20.500.14154/45778https://drepo.sdl.edu.sa/handle/20.500.14154/45778>
- Ali, Z., Palpanadan, S. T., Asad, M. M., Churi, P., & Namaziandost, E. (2022). Reading approaches practiced in EFL classrooms: a narrative review and research agenda. *Asian-Pacific Journal of Second and Foreign Language Education*, 7(1), 28. <https://doi.org/10.1186/s40862-022-00155-4>
- Alptekin, C., & Erçetin, G. (2010). The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading*, 33(2), 206-219. <https://doi.org/10.1111/j.1467-9817.2009.01412.x>
- Alptekin, C., & Erçetin, G. (2011). Effects of Working Memory Capacity and Content Familiarity on Literal and Inferential Comprehension in L2 Reading. *Tesol Quarterly*, 45(2), 235-266. <https://doi.org/https://doi.org/10.5054/tq.2011.247705>
- Arya, D. J., Hiebert, E. H., & Pearson, P. 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. <https://www.iejee.com/index.php/IEJEE/article/view/216>
- Bagford, J. (1968). Reading Readiness Scores and Success in Reading. *The reading teacher*, 21(4), 324-328. <http://www.jstor.org/stable/20195927>
- Baron, R. (2020). Implementing of academic text in advanced grammar learning. *Voices of English Language Education Society*, 4(1), 53-61. <https://doi.org/10.29408/veles.v4i1.1994>
- Basilan, M. L. J. C., & De Sagun, D. R. G. (2024). Analysis of the role of syntactic complexity in students' reading comprehension: A teacher's perspective. *Journal of Contemporary Educational Research*, 8(8). <https://doi.org/10.26689/jcer.v8i8.7650>
- Basturkmen, H. (2009). Commenting on results in published research articles and masters dissertations in Language Teaching. *Journal of English for Academic Purposes*, 8(4), 241-251. <https://doi.org/https://doi.org/10.1016/j.jeap.2009.07.001>
- Beers, S. F., & Nagy, W. E. (2011). Writing development in four genres from grades three to seven: Syntactic complexity and genre differentiation. *Reading and Writing: An Interdisciplinary Journal*, 24(2), 183-202. <https://doi.org/10.1007/s11145-010-9264-9>
- Benelhadj, F. (2019). Discipline and genre in academic discourse: Prepositional Phrases as a focus. *Journal of Pragmatics*, 139, 190-199. <https://doi.org/10.1016/j.pragma.2018.07.010>
- Bernal, M., & Bernal, P. (2020). Using reading to teach English as a foreign language. *Maskana*, 11(2), 18-26. <https://doi.org/10.18537/mskn.11.02.02>
- BIBER, D. (1993). Representativeness in Corpus Design. *Literary and linguistic computing*, 8(4), 243-257. <https://doi.org/10.1093/lc/8.4.243>

- Biber, D., Gray, B., & Poonpon, K. (2011). Should We Use Characteristics of Conversation to Measure Grammatical Complexity in L2 Writing Development? *Tesol Quarterly*, 45(1), 5-35. <https://doi.org/https://doi.org/10.5054/tq.2011.244483>
- Bogaerds-Hazenberg, S. T. M., Evers-Vermeul, J., & van den Bergh, H. (2022). What textbooks offer and what teachers teach: an analysis of the Dutch reading comprehension curriculum. *Reading and writing*, 35(7), 1497-1523. <https://doi.org/10.1007/s11145-021-10244-4>
- Bremer, N. (1959). Do readiness tests predict success in reading? *The Elementary School Journal*, 59(4), 222-224. <https://doi.org/https://doi.org/10.1086/459719>
- Crossley, S., & McNamara, D. (2011). Text coherence and judgments of essay quality: Models of quality and coherence. Proceedings of the Annual Meeting of the Cognitive Science Society <https://escholarship.org/uc/item/5cp1x9r2https://escholarship.org/uc/item/5cp1x9r2>
- Crossley, S. A., Skalicky, S., Dascalu, M., McNamara, D. S., & Kyle, K. (2017). Predicting text comprehension, processing, and familiarity in adult readers: New approaches to readability formulas. *Discourse Processes*, 54(5-6), 340-359. <https://doi.org/10.1080/0163853X.2017.1296264>
- Curran, M. (2020). Complex Sentences in an Elementary Science Curriculum: A Research Note. *Lang Speech Hear Serv Sch*, 51(2), 329-335. https://doi.org/10.1044/2019_lshss-19-00064
- Dang, C. N., & Dang, T. N. Y. (2023). The Predictive Validity of the IELTS Test and Contribution of IELTS Preparation Courses to International Students' Subsequent Academic Study: Insights from Vietnamese International Students in the UK. *RELC Journal*, 54(1), 84-98. <https://doi.org/10.1177/0033688220985533>
- Dong, S., Mao, J., & Pei, L. (2023). Comparing the Writing Styles of Multiple Disciplines: A Large-Scale Quantitative Analysis. *Proceedings of the Association for Information Science and Technology*, 60(1), 941-943. <https://doi.org/https://doi.org/10.1002/praz.905>
- Ellis, N. C. (2002). Frequency Effects in Language Processing: A Review with Implications for Theories of Implicit and Explicit Language Acquisition. *Studies in second language acquisition*, 24(2), 143-188. <https://doi.org/10.1017/S0272263102002024>
- Fender, M. (2001). A Review of L1 and L2/ESL Word Integration Skills and the Nature of L2/ESL Word Integration Development Involved in Lower-Level Text Processing. *Language learning*, 51(2), 319-396. <https://doi.org/https://doi.org/10.1111/0023-8333.00157>
- Fitria, T. N. (2024). Teaching IELTS Reading Skills. *Pioneer: Journal of Language and Literature*(1), 94-111%V 116. <https://doi.org/10.36841/pioneer.v16i1.3991>
- Fulcher, G. (1997). Text difficulty and accessibility: Reading formulae and expert judgement. *System*, 25(4), 497-513. [https://doi.org/https://doi.org/10.1016/S0346-251X\(97\)00048-1](https://doi.org/https://doi.org/10.1016/S0346-251X(97)00048-1)
- Gagen, T., & Faez, F. (2024). The predictive validity of IELTS scores: a meta-analysis. *Higher Education Research & Development*, 43(4), 873-888. <https://doi.org/10.1080/07294360.2023.2280700>
- Gedik, T. A., & Kolsal, Y. S. (2022). A corpus-based analysis of high school English textbooks and English university entrance exams in Turkey. *Theory and Practice of Second Language Acquisition*, 8(1), 157-176. <https://doi.org/https://doi.org/10.31261/TAPSLA.9152>
- Ghorbani Shemshadsara, Z., Ahour, T., & Hadidi Tamjid, N. (2022). Effects of Multi-Component Training of Text Structure Intervention and Self-Regulated Strategies on Iranian Upper-Intermediate EFL Learners' Reading Comprehension (Research Paper). *Iranian Journal of English for Academic Purposes*, 11(2), 90-106. https://journalscmu.sinaweb.net/article_157885_54efc0b33797367e66e8d59aca0e6067.pdf

- Grabe, W., & Stoller, F. L. (2019). *Teaching and researching reading*. Routledge. <https://doi.org/https://doi.org/10.4324/9781315726274>
- Hewavitharana, S., & Vogel, S. (2008). Enhancing a statistical machine translation system by using an automatically extracted parallel corpus from comparable sources. Proceedings of the Workshop on Comparable Corpora, LREC'08 https://www.academia.edu/2876175/Enhancing_Statistical_Machine_Translation_with_Parallel_Data_extracted_from_Comparable_Corporahttps://www.academia.edu/2876175/Enhancing_Statistical_Machine_Translation_with_Parallel_Data_extracted_from_Comparable_Corpora
- Hopkins, K. D., & Sitkei, E. G. (1969). Predicting Grade One Reading Performance. *The Journal of Experimental Education*, 37(3), 31-33. <https://doi.org/10.1080/00220973.1969.11011127>
- Hulstijn, J. H. (2011). Language Proficiency in Native and Nonnative Speakers: An Agenda for Research and Suggestions for Second-Language Assessment. *Language Assessment Quarterly*, 8(3), 229-249. <https://doi.org/10.1080/15434303.2011.565844>
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4-21. <https://doi.org/https://doi.org/10.1016/j.esp.2007.06.001>
- Indrawan, F. (2018). *Text Readability and Syntactic Complexity in the Reading Texts of Indonesian Senior High School English Textbooks Universitas Airlangga*. <http://repository.unair.ac.id/id/eprint/70474><http://repository.unair.ac.id/id/eprint/70474>
- Ji, H. (2009). Mining name translations from comparable corpora by creating bilingual information networks. Proceedings of the 2nd Workshop on Building and Using Comparable Corpora: from Parallel to Non-parallel Corpora (BUCC) <https://aclanthology.org/W09-3107/><https://aclanthology.org/W09-3107/>
- Jin, T., Lu, X., & Ni, J. (2020). Syntactic Complexity in Adapted Teaching Materials: Differences Among Grade Levels and Implications for Benchmarking. *The Modern Language Journal*, 104(1), 192-208. <https://doi.org/https://doi.org/10.1111/modl.12622>
- Johnson, P. (1981). Effects on Reading Comprehension of Language Complexity and Cultural Background of a Text. *Tesol Quarterly*, 15(2), 169-181. <https://doi.org/https://doi.org/10.2307/3586408>
- Johnson, R. C., & Tweedie, M. G. (2021). "IELTS-out/TOEFL-out": Is the End of General English for Academic Purposes Near? Tertiary Student Achievement Across Standardized Tests and General EAP. *Interchange*, 52(1), 101-113. <https://doi.org/10.1007/s10780-021-09416-6>
- Jung, Y., Crossley, S., & McNamara, D. (2019). Predicting second language writing proficiency in learner texts using computational tools. *Journal of Asia TEFL*, 16(1), 37. <https://doi.org/http://dx.doi.org/10.18823/asiatefl.2019.16.1.3.37>
- Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Allyn & Bacon. <https://psycnet.apa.org/record/1986-98384-000>
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological review*, 99(1), 122-149. <https://doi.org/10.1037/0033-295X.99.1.122>
- Karami, M., & Salahshoor, F. (2014). The relative significance of lexical richness and syntactic complexity as predictors of academic reading performance. *International Journal of Research Studies in Language Learning*, 3(2), 17-28. <https://doi.org/10.5861/ijrse.2013.477>
- Kerstjens, M., & Nery, C. (2000). Predictive validity in the IELTS test: A study of the relationship between IELTS scores and students' subsequent academic performance. *IELTS research reports*, 3(4), 86-108. <https://doi.org/https://ielts.org/researchers/our-research/research->

[reports/predictive-validity-in-the-ielts-test-a-study-of-the-relationship-between-ielts-scores-and-students-subsequent-academic-performance](#)

- Kheirabadi, R., & Alavimoghaddam, S. B. (2016). Evaluation of Prospect series: A paradigm shift from GTM to CLT in Iran. *Journal of Language Teaching and Research*, 7(3), 619-624. <https://doi.org/10.17507/jltr.0703.26>
- Kim, E., & Oh, J. (2019). 수능 영어 독해 지문과 「영어 II」 교과서에 나타난 통사적 복잡도에 관한 연구 [A Corpus-based Analysis of the Syntactic Complexity Levels of Reading Passages in the College Entrance English Examination and English II]. *STUDIES IN ENGLISH EDUCATION*, 24(3), 399-418. <https://doi.org/10.22275/SEE.24.3.03>
- Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational psychology*, 100(2), 310-321. <https://doi.org/10.1037/0022-0663.100.2.310>
- Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524841>
- Lak, M., Soleimani, H., & Parvaneh, F. (2017). The effect of teacher centeredness method vs. learner-centeredness method on reading comprehension among Iranian EFL learners. *Advances in English Language Teaching*, 5(1), 1-10. <http://european-science.com/jaelt/article/view/4886>
- Le, T. N. P., & Harrington, M. (2015). Phraseology used to comment on results in the Discussion section of applied linguistics quantitative research articles. *English for Specific Purposes*, 39, 45-61. <https://doi.org/10.1016/j.esp.2015.03.003>
- Liontou, T. (2015). *Computational text analysis and reading comprehension exam complexity: towards automatic text classification* (Vol. 36). <https://doi.org/10.3726/978-3-653-04944-2>
- Liu, C., & Li, Y. (2023). A Research on the Predicament and Improving Strategies of China College Entrance Examination Thinking in IELTS Reading—Comparative Analysis of Chinese College Entrance Examination English Reading and IELTS Reading. *Journal of Advanced Research in Education*, 2(4), 11-19. <https://doi.org/10.56397/jare.2023.07.03>
- 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>
- Maamujav, U., Olson, C. B., & Chung, H. (2021). Syntactic and lexical features of adolescent L2 students' academic writing. *Journal of second language writing*, 53, 100822. <https://doi.org/10.1016/j.jslw.2021.100822>
- Marina, K. (2018). The validation process in the IELTS reading component: Reading requirements for preparing international students. *Journal of Language and Education*, 4(1 (13)), 63-78. <https://doi.org/10.17323/2411-7390-2018-4-1-63-78>
- Marjerison, R. K., Liu, P., Duffy, L. P., & Chen, R. (2020). An Exploration of the Relationships Between Different Reading Strategies and IELTS Test Performance: IELTS Test Taking Strategies - Chinese Students. *International Journal of Translation, Interpretation, and Applied Linguistics (IJTIAL)*, 2(1), 1-19. <https://doi.org/10.4018/IJTIAL.2020010101>
- Mauriyat, A. (2021). Authenticity and validity of the IELTS writing test as predictor of academic performance. *PROJECT (Professional Journal of English Education)*, 4(1), 105-115. <https://doi.org/10.22460/project.v4i1.p105-115>
- McEnery, T., & Brookes, G. (2022). Building a written corpus: what are the basics? In *The Routledge handbook of corpus linguistics* (pp. 35-47). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780367076399-4/building-written-corpus-basics-tony-mcenery-gavin-brookes>

- McEnery, T., & Wilson, A. (2001). *Corpus Linguistics: An Introduction*. Edinburgh University Press. <http://www.jstor.org/stable/10.3366/j.ctvxcrjmp>
- 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>
- Moran, K. E. (2013). *Exploring Undergraduate Disciplinary Writing: Expectations and Evidence in Psychology and Chemistry* Georgia State University]. <https://doi.org/10.57709/3589615>
- Morvay, G. (2012). The relationship between syntactic knowledge and reading comprehension in EFL learners. *Studies in Second Language Learning and Teaching*, 2(3), 415-438. <https://doi.org/https://doi.org/10.14746/ssllt.2012.2.3.8>
- Nation, I. (2006). How large a vocabulary is needed for reading and listening? *Canadian modern language review*, 63(1), 59-82. <https://doi.org/10.1353/cml.2006.0049>
- Nation, K., & Snowling, M. J. (2000). Factors influencing syntactic awareness skills in normal readers and poor comprehenders. *Applied psycholinguistics*, 21(2), 229-241. <https://doi.org/10.1017/S0142716400002046>
- Nergis, A. (2013). Exploring the factors that affect reading comprehension of EAP learners. *Journal of English for Academic Purposes*, 12(1), 1-9. <https://doi.org/https://doi.org/10.1016/j.jeap.2012.09.001>
- Norris, J. M., & Ortega, L. (2009). Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity. *Applied linguistics*, 30(4), 555-578. <https://doi.org/10.1093/applin/amp044>
- Nuttall, C. (1996). *Teaching reading skills in a foreign language*. ERIC. <https://eric.ed.gov/?id=ED399531>
- Nuttall, C. (2005). *Teaching Reading Skills in a Foreign Language*. Macmillan Education. https://books.google.com/books/about/Teaching_Reading_Skills_in_a_Foreign_Lan.html?id=C1szNwAACAAJ
- Parkinson, M. M., & Dinsmore, D. L. (2018). Multiple aspects of high school students' strategic processing on reading outcomes: The role of quantity, quality, and conjunctive strategy use. *British Journal of Educational Psychology*, 88(1), 42-62. <https://doi.org/10.1111/bjep.12176>
- Perfetti, C. A. (1985). *Reading ability*. Oxford University Press. <https://psycnet.apa.org/record/1985-97290-000>
- Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity*. (pp. 67-86). American Psychological Association. <https://doi.org/10.1037/10459-004>
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The Acquisition of Reading Comprehension Skill. In *The science of reading: A handbook*. (pp. 227-247). Blackwell Publishing. <https://doi.org/10.1002/9780470757642.ch13>
- Powell, M., & Parsley, K. M. (1961). The Relationships between First Grade Reading Readiness and Second Grade Reading Achievement. *The Journal of Educational Research*, 54(6), 229-233. <https://doi.org/10.1080/00220671.1961.10882715>
- Putra, D. A., & Lukmana, I. (2017). Text complexity in senior high school English textbooks: A systemic functional perspective. *Indonesian Journal of Applied Linguistics*, 7(2), 436-444. <https://doi.org/https://doi.org/10.17509/ijal.v7i2.8352>
- Qian, D. D. (2002). Investigating the Relationship Between Vocabulary Knowledge and Academic Reading Performance: An Assessment Perspective. *Language learning*, 52(3), 513-536. <https://doi.org/https://doi.org/10.1111/1467-9922.00193>

- Rayner, K., & Pollatsek, A. (1996). Reading unspaced text is not easy: Comments on the implications of Epelboim et al.'s (1994) study for models of eye movement control in reading. *Vision research*, 36(3), 461-465. [https://doi.org/https://doi.org/10.1016/0042-6989\(95\)00132-8](https://doi.org/https://doi.org/10.1016/0042-6989(95)00132-8)
- Razmjoo, S. A., & Heydari Tabrizi, H. (2010). A Content Analysis of the TEFL MA Entrance Examinations (Case Study: Majors Courses). *Journal of Pan-Pacific Association of Applied Linguistics*, 14(1), 159-170. https://www.kci.go.kr/kciportal/landing/article.kci?arti_id=ART002404223
- Riemenschneider, A., Weiss, Z., Schröter, P., & Meurers, D. (2024). The Interplay of Task Characteristics, Linguistic Complexity, and Language Proficiency in High-Stakes English as a Foreign Language Writing. *Tesol Quarterly*, 58(2), 775-801. <https://doi.org/https://doi.org/10.1002/tesq.3254>
- Saeedi, Z., & Shahrokhi, M. (2019). Cultural content analysis of Iranian ELT coursebooks: A comparison of Vision I & II with English for Pre-university students I & II. *International Journal of Foreign Language Teaching and Research*, 7(27), 107-124. https://journals.iau.ir/article_628140.html
- Samraj, B. (2005). An exploration of a genre set: Research article abstracts and introductions in two disciplines. *English for Specific Purposes*, 24(2), 141-156. <https://doi.org/https://doi.org/10.1016/j.esp.2002.10.001>
- Sellers, V. D. (2000). Anxiety and Reading Comprehension in Spanish as a Foreign Language. *Foreign Language Annals*, 33(5), 512-520. <https://doi.org/https://doi.org/10.1111/j.1944-9720.2000.tb01995.x>
- Shaw, P., & Ting-Kun Liu, E. (1998). What Develops in the Development of Second-language Writing? *Applied linguistics*, 19(2), 225-254. <https://doi.org/10.1093/applin/19.2.225>
- Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. *Reading Research Quarterly*, 50(3), 337-356. <https://psycnet.apa.org/record/2015-28801-005>
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press. https://discourses.org/wp-content/uploads/2022/06/Teun-A-van-Dijk-Walter-Kintsch-1983-Strategies-Of-Discourse-Comprehension.pdf&ved=2ahUKEwilpJ_Lz5aLAXWKcPEDHVcpIp4QFnoECBwQAQ&usq=AOvVaw2fsavU9ZDzGvmI2McZTzgK
- Verdiansyah, M. Z. (2020). Text complexity in reading texts of Indonesian senior high school English textbooks using Coh- metrix 3.0. <https://doi.org/10.26594/diglossia.v12i1.1925>
- Vinogradova, O., Smirnova, E., Viklova, A., & Panteleeva, I. (2020). Syntactic complexity of academic text: A corpus study of written production by learners of English with Russian L1 in comparison with expert texts of English authors. *RSUH/RGGU Bulletin: "Literary Theory. Linguistics. Cultural Studies"*, 7, 107-129. <https://www.semanticscholar.org/paper/SYNTACTIC-COMPLEXITY-OF-ACADEMIC-TEXT%3A-A-CORPUS-OF-Vinogradova-Smirnova/81947995bef74de420c09003a6b66c5e0fa0f32f>
- Walková, M., & Bradford, J. (2022). Constructing an argument in academic writing across disciplines. *ESP Today*, 10(1), 22-42. <https://doi.org/10.18485/esptoday.2022.10.1.2>
- Wijanti, W. (2017). Syntactic Complexity in the Reading Materials of English for Academic Purposes Levels 1–3. *LLT Journal: A Journal on Language and Language Teaching*, 20(2), 102-115. <https://doi.org/10.24071/llt.v20i2.737>

- Yang, K., & Bae, J. (2022). A Continuity and Difficulty Analysis of the Reading Texts in Korean High School English Textbooks with the 2015 Revised National Curriculum. *English Teaching, 11*, 43-62. <https://doi.org/https://doi.org/10.15858/engtea.77.s1.202209.43>
- Yazdi, V., & Mohammadian, A. (2022). The Relationship Between Syntactic Knowledge and Speaking and Writing Proficiency Among Iranian Intermediate EFL Learners (Research Paper). *Iranian Journal of English for Academic Purposes, 11*(3), 84-96. https://journalscmu.sinaweb.net/article_162119_d9e6328e15485e1868056dcb9b39cfed.pdf
- Zuhra, Z. (2015). Senior high school students' difficulties in reading comprehension. *English Education Journal, 6*(3), 430-441. <https://jurnal.usk.ac.id/EEJ/article/view/2584>