Research in English Language Pedagogy (2025) 13(2): 130206

 DOI: 10.71673/relp.2025.1195182

Original Research

A Corpus-based Comparison of Syntactic Complexity Measures as Indices of Advanced English Text Comprehension in EAP Textbooks and Academic Research Papers

Peyman Nasrabady[1], *Hooshang Khoshsima*[1*], *Amir Mohammadian*[1], *Nahid Yarahmadzehi*[1]
[1]Department of English, Chabahar Maritime University, Chabahar, Iran

Submission date: 31-12-2024 Acceptance date: 07-02-2025

Abstract

Reading comprehension is a crucial skill for language learners, underpinning their ability to understand texts and achieve academic success. Even with technological advances, written materials have remained fundamental to education—especially in higher education—by facilitating knowledge acquisition and promoting innovative ways of thinking. Many studies have examined reader-related challenges affecting reading comprehension; however, there remains a need in the literature to shed light on the factors inherent in the texts as well. To address this, the current study examined syntactic complexity across six corpora of advanced academic reading texts, which were compiled from books entitled "English for Students of Pharmacy," "English for Students of Psychology," and "English for Students of Accounting," as well as real-life academic written texts in the same disciplines. The texts were analysed using the L2 Syntactic Complexity Analyser (L2SCA), which is based on syntactic complexity measures. A MANOVA test identified significant differences among the texts. The findings further revealed that the textbooks do not align with real-life academic texts in terms of syntactic complexity, resulting in a disconnect between learners' readiness and the demands of real-world academic reading. The findings underscored the necessity for more representative and comparable syntactic features in English for university textbooks.

Keywords: EAP/ESP Textbooks, Corpus, Reading Texts, Research Papers, Syntactic Complexity

* Corresponding Author's E-mail: Khoshsima2002@yahoo.com

1. Introduction

Reading comprehension is a central aspect of academic success, particularly in higher education, where much of the learning process is mediated through written texts and materials (Hyland, 2006). Academic reading requires students not only to understand but also to evaluate and synthesise information from complex texts critically. This skill is particularly challenging for learners of English for Academic Purposes (EAP), who face specific obstacles when navigating the dense language and structures of scholarly texts and materials (Grabe, 2008). Therefore, developing strong academic literacy skills is essential for these learners to succeed in such an environment (Biber & Gray, 2010).

A primary challenge for EAP learners is the transition from simplified instructional materials, such as textbooks, to authentic academic research papers in higher education settings. EAP textbooks, as one source of learner reading input, are typically designed to scaffold learning by simplifying syntax to enhance accessibility (Flowerdew, 2005). On the other hand, academic research papers, as real-life reading input for learners, employ much more complex syntactic structures, including extended noun phrases, subordinate clauses, and passive constructions. These features convey nuanced and detailed arguments through grammatical complexities, creating a significant feature in academic texts (Biber et al., 2011). Therefore, this can create a gap between learner preparedness and the demands of real-life academic reading. Accordingly, this disparity can hinder learners' ability to engage effectively with academic texts, highlighting the need for targeted instructional strategies to address these linguistic challenges (Arya et al., 2011).

In EFL contexts, where learners have limited exposure to authentic English input, textbooks and test preparation materials are essential for supporting language acquisition and academic progress (Ali et al., 2022; Bernal & Bernal, 2020; Nuttall, 2005). However, many students struggle with reading comprehension, which is influenced by both the syntactic complexity of the texts and the students' grammatical knowledge. Readers' grammatical proficiency is essential for facilitating comprehension, while limited syntactic awareness has often created significant barriers (Tarlani-Aliabadi et al., 2022). Building on this, research has suggested that pedagogical strategies must explicitly target syntactic understanding, equipping learners with the skills to navigate complex sentence structures effectively (Ahmed & Ahmed, 2023). However, text-related factors have frequently challenged these efforts. Insufficient syntactic practice in textbooks has been identified and

reviewed in the related literature, which negatively affected students' ability to build strong comprehension skills (Alenezi, 2016; Mousavi et al., 2021). This challenge becomes even more significant in advanced academic texts, where high levels of syntactic density and structural complexity in the materials require an explicit instructional focus (Grabe & Stoller, 2019).

Syntactic complexity refers to the grammatical sophistication of sentence structures in a text, which plays a crucial role in determining reading difficulty (Lu, 2011). Texts with intricate grammatical constructions require advanced cognitive and linguistic skills to process. For many EAP learners, these complexities present substantial barriers to comprehension, emphasizing the need for instructional materials that better prepare students for academic reading challenges (Chen & Meurers, 2018).

Despite the acknowledged significance of syntactic complexity, research comparing the syntactic features of EAP textbooks and academic research papers is relatively sparse as far as the authors have been able to review the literature. This gap can limit our understanding of how well EAP materials align with the complexities of authentic academic texts (Jin et al., 2020). Addressing this gap is crucial for developing EAP curricula that bridge the divide between simplified learning materials and the linguistic demands of scholarly texts, thereby supporting learners in their transition to advanced academic literacy (Biber et al., 2011; Martinussen & Mackenzie, 2015).

This study aimed to investigate the differences in syntactic complexity between English for Academic Purposes (EAP) textbooks and academic research papers. By identifying key syntactic features, such as sentence length, clause variety, and overall sentence complexity, the research aims to provide insights that can inform the development of more effective English for Academic Purposes (EAP) instructional materials. Ultimately, this study aims to enhance the alignment between EAP curricula and the linguistic demands of academic scholarship, thereby equipping learners with the necessary tools for academic success.

2. Literature Review

Reading comprehension is a complex process that involves the dynamic interaction of reader- and text-related factors, working together to facilitate the construction of meaning. Reader-related elements, such as linguistic proficiency, background knowledge, and reading strategies, interact with text-based features like syntax, vocabulary, and genre.

Together, these components require learners to employ a range of cognitive and linguistic resources to navigate and interpret texts effectively (Grabe & Stoller, 2019; Koda, 2005).

In the context of English as a foreign language, access to the language input necessary for developing linguistic competence is often limited beyond the classroom environment (Al-Jamal & Al-Jamal, 2013). For many students, English textbooks serve as the primary—and in some cases, nearly exclusive—source of linguistic input (Charalambous & Ed, 2011; Jisu & Moongee, 2020).

In EFL settings, limited access to authentic and real-life English makes reading materials an essential means of providing language input for learners (Ali et al., 2022; Bernal & Bernal, 2020; Lak, 2017; Nuttall, 1996). Consequently, educators and textbook developers need to carefully select and design reading materials that enhance the learning environment and help students progress in their English proficiency. Whereas some textbooks demonstrate a coherent and systematic conceptualization of linguistic syntactic complexity features, others lack this awareness, which may challenge students' ability to engage effectively with more advanced academic texts (Putra & Lukmana, 2017; Yang & Bae, 2022).

On the other hand, research on the effectiveness of English for Academic Purposes (EAP) textbooks has identified some shortcomings and areas that require improvement in these materials. In an analysis by Najjar (2020) in Iran as an EFL context, while EAP textbooks often aligned with learners' needs at a macro level, they frequently fell short in addressing specific language skills such as reading, writing, listening, and speaking. Furthermore, essential grammatical elements, such as transition markers—vital for comprehending academic texts—were inadequately addressed, leading to their overuse or misuse by learners and underscoring the need for supplementary materials and explicit instruction (Walková, 2020). Furthermore, evaluations indicated that EAP coursebooks met only about 50% of the established criteria for effective teaching materials, particularly in integrating skills such as speaking and listening, which suggested significant room for improvement (Sabiq & Muflihah, 2021). Additionally, in an EFL context such as Japan, commercially available EAP textbooks often lack cultural relevance and appropriate difficulty levels, necessitating the development of in-house materials that better suit local needs (Ruegg et al., 2018). While students generally viewed genre-oriented tasks in these textbooks positively, not all benefited equally, revealing hidden challenges; however,

tailored EAP materials were shown to enhance student motivation and proficiency (Jou, 2017; Rasekh et al., 2011).

A review of studies in the area of EAP textbooks and syntactic complexity features suggests that syntactic complexity is not merely an ornamental feature of academic language but a foundational element that underpins effective language development in EAP contexts. The studies suggested that exposure to complex sentence structures—as found in well-designed EAP materials—played a crucial role in enhancing students' ability to produce and process intricate syntactic constructions, which in turn bolstered both reading comprehension and writing proficiency (Casal & Lee, 2019; Frantz et al., 2015; Karami & Salahshoor, 2013; Ortega, 2015; Wijanti, 2017). Moreover, some advocates have proposed a balanced approach to textbook design: one that challenges students with syntactically rich texts while remaining readable enough to facilitate better academic performance (Mazgutova & Kormos, 2015; Wijanti, 2017).

Therefore, syntactic complexity has emerged as a critical determinant of comprehension, particularly for EAP materials. The ability to process and engage with complex sentence structures is essential for developing advanced reading skills. However, inconsistencies in the syntactic demands of educational materials can limit learners' exposure to the types of structures that foster proficiency. This gap highlights the importance of designing materials that align with learners' evolving linguistic capacities, ensuring they are appropriately challenged (Sun, 2020).

According to the related literature, the treatment of syntactic complexity in EAP textbooks has appeared inconsistent and misaligned with authentic academic discourse, raising concerns about their effectiveness in fostering students' academic language proficiency. The lack of consistency in syntactic complexity measures resulted in challenging research findings and failed to provide a reliable benchmark for tracking learners' linguistic development (Deng et al., 2020). Moreover, the syntactic structures presented in these materials often did not reflect those found in academic exams, creating a negative backwash effect that might have misguided students' learning priorities (Gedik & Kolsal, 2022). Another pressing issue was the failure of EAP textbooks to account for syntactic variation across academic disciplines and genres, leading to a disconnect between the structures taught and those required in different fields of study (Casal et al., 2021; Smirnova, 2022). Additionally, as mentioned earlier, transition markers—essential for

constructing syntactically coherent academic texts—were often inadequately addressed, leaving students prone to overuse and misuse, particularly due to insufficient instruction on the syntactic distinctions between conjunctions and adverbials (Walková, 2020). These findings can highlight the need for a revision of syntactic complexity if the texts in EAP materials are to effectively support learners in acquiring the advanced linguistic structures essential for academic success.

Despite these insights, and as far as the authors had reviewed in the literature, comparative studies in this area remained scarce—particularly those evaluating the syntactic complexity of reading texts in EAP textbooks against real-life academic language data. Therefore, the current study aimed to contribute to this line of research on EAP textbooks and to raise awareness of syntactic complexity measures in EAP learner input. To achieve this aim, syntactic complexity measures in a sample of EAP textbooks developed for Pharmacy, Accounting, and Psychology university students were analyzed and compared to those in research paper discussions within the same academic disciplines. Building on these considerations, the study explored the following question:

1. How do the syntactic complexity of texts in EAP textbooks and research paper discussions within the same academic disciplines compare?

3. Methodology

3.1. Design and Context of the Study

The present study employs a **corpus-based quantitative research design** to examine the syntactic complexity and readability of advanced reading comprehension texts, sourced explicitly from educational EAP textbooks and academic research papers in three distinct academic disciplines: Psychology, Accounting, and Pharmacy. Each EAP textbook was paired with research papers from the same discipline to ensure consistency and comparability across the corpora. Using the L2 Syntactic Complexity Analyzer (L2SCA) as the analytical tool, the study quantitatively evaluates key syntactic dimensions to uncover patterns and differences in linguistic structures across these genres of EFL learner input. Statistical analyses, particularly MANOVA, are applied to assess how these syntactic features vary, providing a framework for identifying meaningful relationships and contrasts.

3.2. Instruments

The primary instruments used in this study were text document data and the L2 Syntactic Complexity Analyser (L2SCA), a linguistic analysis tool. Six corpora of advanced reading texts were compiled from three academic disciplines: Psychology (EAP textbooks and research papers), Accounting (EAP textbooks and research papers), and Pharmacy (EAP textbooks and research papers). These texts were analyzed using the L2SCA to evaluate syntactic complexity.

SAMT, a prominent publisher of English for Academic Purposes (EAP) textbooks in Iran, provided the materials for this study. The aim was to assess the syntactic complexity of SAMT's EAP reading materials and compare them to that of academic research papers. EAP textbooks from SAMT were selected, each containing at least 10 units with reading passages relevant to academic disciplines. Three textbooks, serving as primary resources for "Specialized English" courses in the targeted fields (Psychology, Accounting, and Pharmacy), were chosen (SAMT, 2024a, 2024b, 2024c). Ten texts were extracted from each textbook, resulting in 30 texts, compiled into three separate EAP textbook corpora (see Appendix A).

To provide a comparison, 10 research papers from Scopus-indexed journals were selected for each academic discipline, forming the second corpus. These research papers were selected based on Scopus subject areas, which encompass 26 main categories with numerous subcategories. A total of 30 research papers—10 from each field—were compiled into three separate academic research text corpora.

3.4. Corpus Development

The process of compiling the corpora for this study was carefully aligned with its research objectives. As McEnery and Brookes (2022) explain, corpora can be classified as either general or specialized. General corpora encompass a broad spectrum of language data, whereas specialized corpora focus on specific genres or varieties of language. For this study, specialized corpora were developed by selecting advanced reading comprehension texts from EAP textbooks and academic research papers, ensuring their relevance to the study's emphasis on EFL learner input.

In preparing the reading comprehension passages for syntactic complexity analysis, the researchers undertook several steps to ensure validity. This included formatting the text

in plain text, ensuring proper punctuation, removing extraneous information, and addressing word separation issues. These measures were crucial for creating a corpus compatible with the L2SCA tool's requirements.

Unlike EAP textbook reading comprehension texts, research papers contain significantly more textual content. Because the complete article data in academic research is usually much longer than in reading comprehension texts, and to ensure comparability of the textual language data in both corpora and also to enhance the validity of the corpus-based analysis, only the discussion sections of the research papers were included. About word numbers, a specific section of a research paper is typically more comparable to a reading comprehension text found in EAP textbooks. Therefore, the authors selected a specific section of the research papers (discussion sections) to develop real-life language corpora for the Pharmacy, Accounting, and Psychology disciplines. Discussion sections are recognised for their focus on interpreting findings and advancing knowledge claims, which are critical aims of research articles (Basturkmen, 2009; Le & Harrington, 2015). Detailed information on the journals and research papers is provided in Appendix B. Additionally, more information on the linguistic analysis tools used to compare the corpora is covered in the "Data Analysis" section.

Authenticity, a key consideration in corpus design, involves including language data that reflects real-world usage with minimal researcher interference (McEnery & Wilson, 2001). To maintain authenticity, this study utilised EAP textbooks and academic research papers as corpora for development and analysis, which are the types of input that EFL learners commonly encounter in academic contexts, such as universities. The chosen EAP textbooks, published by the SAMT Association, are widely used in Iran for various academic disciplines. Although not internationally renowned, these textbooks are integral to university English courses in the region. Their inclusion ensured that the texts reflected materials commonly available to university students in Iran, with which EFL learners typically engage, thereby reinforcing the authenticity of the corpora.

Representativeness was another crucial criterion, ensuring the language data reflected the variability within the target genre (Biber, 1993). To achieve this, the study included a diverse selection of advanced reading texts from three academic disciplines—Psychology, Accounting, and Pharmacy. These texts, commonly encountered by EFL learners, form a significant portion of their academic language input. Sampling from

multiple disciplines enhanced the corpora's representativeness, thereby supporting the validity of the syntactic complexity analyses.

To ensure comparability, texts within each corpus were drawn from the same academic discipline, with equal numbers of passages selected from each source. This approach facilitated valid comparisons across corpora, as texts within each discipline shared common characteristics and functions. As Hewavitharana and Vogel (2008) and Ji (2009) emphasize, comparability is critical for aligning texts in terms of content and purpose. In this study, six corpora—covering texts from three disciplines (Psychology, Accounting, and Pharmacy) and their respective research papers—each comprised 10 advanced reading passages.

3.5. Data Analysis Procedure

To address the research question of this study and assess advanced reading texts, a series of procedures was implemented. Initially, the textual data, comprising advanced reading passages from a selection of EAP textbooks and academic research papers, were collected to create individual corpora. Subsequently, an objective syntactic complexity analysis of the texts was conducted using the L2SCA tool. This involved feeding each corpus into the tool and analysing all syntactic measures of the texts, resulting in six analyses: three for the EAP textbooks individually and three for the research papers across the three academic disciplines. The derived results formed the foundation for subsequent comparisons. Each EAP textbook was juxtaposed with and compared to the research papers within its corresponding academic discipline, leveraging the syntactic complexity measures of the texts. The comparative results formed the basis for in-depth discussions and further analyses.

According to previous literature, readers process texts linearly, decoding them word by word; however, as they read, they need to compile the linguistic items into larger-scale syntactic structures (Just & Carpenter, 1987; Rayner & Pollatsek, 1996). Accordingly, the mental demands required for this operation can vary considerably depending on the complexity of the structure (Perfetti et al., 2005). Therefore, the collected texts were analysed using all 14 measures computed by the syntactic complexity measures of L2SCA in order to make the analysis comprehensive, covering each of the four core complexity measures. Additionally, the measures are divided into four core groups: length of the

production unit, amount of subordination, amount of coordination, and degree of phrasal sophistication. Table 1 presents more information on the syntactic complexity measures.

Table 1.

L2 Syntactic Complexity Measures

Category	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	C/T	Number of clauses per T-unit
	CT/T	Complex T-unit ratio
	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
	VP/T	Number of verb phrases per T-unit
Overall Sentence Complexity	C/S	Number of clauses per sentence

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

The amount of coordination is assessed through the number of coordinate phrases per clause (CP/C), the number of coordinate phrases per T-unit (CP/T), and the number of T-units per sentence (T/S) (Lu, 2010). 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) (Lu, 2010) This index reflects the complexity of noun phrases, which can impact the overall processing demands on the reader. Therefore, the L2SCA was used in the current study because it provides a comprehensive framework for analysing text complexity by examining various aspects of syntactic complexity, including the length of the production unit, the amount of subordination, the amount of coordination, and the degree of phrasal sophistication. All these 14 measures were calculated to ensure a comprehensive descriptive analysis of the variance between the corpora. However, with regard to the inferential analysis of the data, a following study by Ai and Lu (2013) provided 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 their groups, revealing how syntactic complexity varied across proficiency levels and text types. Their groupings are reviewed in Table 2.

Additionally, their study demonstrated the effectiveness of categorizing syntactic complexity measures to uncover patterns and relationships in diverse text sources. Similarly, in the current study, after calculating and analyzing the descriptive data, the syntactic complexity measures were systematically grouped into four distinct categories to enable effective statistical analysis. The length of the production unit characterises these groups**, as well as the amount of subordination, coordination, and phrasal sophistication**. Therefore, the approach aligns with established methodologies in prior research, facilitating the identification of meaningful differences and trends in syntactic complexity among the corpora under investigation.

Table 2.

Grouping Syntactic Complexity Measures Based on Ai and Lu (2013)

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

To examine variations in syntactic complexity across six corpora, a one-way multivariate analysis of variance (MANOVA) 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.

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 six 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, as described in the report presented above.

The MANOVA procedure was structured to test whether the mean vectors of the four dependent variables varied across the six corpora. **Wilks' Lambda** was employed as the primary multivariate test statistic, evaluating the significance measure of the overall effect of the independent variable on the combined dependent variables.

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 simultaneously, 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 employed to highlight the variation in syntactic features between these written genres.

4. Results and Discussion

4.1 Descriptive Analysis of the Results

Based on Table 3, the comparison between the *"Psychology books corpus"* and *the "Psychology papers corpus"* revealed several descriptive differences. The "Psychology papers corpus" contained a larger word count (12,720) compared to the "Psychology books corpus" (5,488), suggesting that the average word count in reading comprehension texts in EAP textbooks is less than half of that found in research paper discussions. While exact equivalence in word count between these texts and research paper discussions is not likely, a closer alignment in word count would likely better reflect the demands of academic

reading, potentially enhancing students' preparedness for engaging with authentic academic texts. Similar structural and text-related issues in EAP textbooks have been addressed in the previous literature as well. According to Wood and Appel (2014), EAP textbooks should ideally address vocabulary and structural constructions relevant to academic discourse; however, studies show that these are often inadequately covered, hindering students' ability to engage effectively with research discussions (Coxhead et al., 2017). Furthermore, in a study by Harwood (2005), many textbooks did not adequately incorporate research findings in their contents, creating a gap between what is taught and the real-life language used in academic written texts. This discrepancy highlights the need for textbooks that better reflect the linguistic features of authentic academic discourse.

In terms of sentence structure, the mean length of sentence (MLS) in the *"Psychology books corpus"* was shorter (16.33) compared to the *"Psychology papers corpus"* (27.24). This suggested that sentences in the books were simpler, while those in the papers were longer and more information-dense. Similarly, the mean length of T-unit (MLT) and mean length of clause (MLC) were higher in the *"Psychology papers corpus,"* reflecting its use of more complex sentence structures.

Table 3.
Descriptive Results of the Syntactic Complexity Measure in Psychology EAP Textbooks versus Papers

Measure	Books Corpus	Papers Corpus	Difference
Words (nwords)	5,488	12,720	+7,232
Mean Length of Sentence (MLS)	16.3	27.2	+10.9
Mean Length of T-unit (MLT)	14.8	25.3	+10.5
Mean Length of Clause (MLC)	9.94	15.2	+5.26
Clauses per Sentence (C_S)	1.64	1.78	+0.14
Verb Phrases per T-unit (VP_T)	2.13	2.38	+0.25
Clauses per T-unit (C_T)	1.49	1.66	+0.17
Dependent Clauses per Clause (DC_C)	0.34	0.40	+0.06
Dependent Clauses per T-unit (DC_T)	0.52	0.67	+0.15
T-units per Sentence (T_S)	1.09	1.07	-0.02
Complex T-units per T-unit (CT_T)	0.38	0.47	+0.09
Complex Phrases per T-unit (CP_T)	0.60	0.87	+0.27
Complex Phrases per Clause (CP_C)	0.40	0.52	+0.12
Complex Noun Phrases per T-unit (CN_T)	1.75	3.90	+2.15
Complex Noun Phrases per Clause (CN_C)	1.17	2.34	+1.17

The papers also exhibited a greater reliance on dependent clauses and complex nominal phrases. Higher ratios of dependent clauses per clause and per T-unit (DC_C and

DC_T) in the papers pointed to their use for structuring information and enhancing cohesion. Additionally, the ratios of complex nominal phrases per T-unit and per clause (CN_T and CN_C) were also higher, indicating a stronger use of intricate nominal constructions.

These findings underscored Psychology's tendency toward detailed and intricate discourse in real-life academic written texts. The longer sentences, higher density of dependent clauses, and greater use of complex nominal constructions in papers aligned with the discipline's commitment to academic rigour and comprehensive exploration of theoretical concepts and empirical findings (Beech, 2009; Hartley, 2008). Furthermore, the larger word count in papers compared to the EAP psychology textbooks reflected the expansive nature of psychological research, incorporating detailed analyses and experimental findings that went beyond the brevity typical of textbook passage (Breakwell et al., 2020).

These findings not only highlighted the complexity of psychological writing but also emphasised the discipline's dedication to syntactical depth and richness in its research texts, as reflected in all grammatical complexity indices. However, the EAP textbook in this discipline fell short in effectively representing the syntactic features of real-life texts, such as those found in research papers.

The comparison of the "Accounting book corpus" and "Accounting papers corpus" revealed some insights into their linguistic characteristics. According to Table 4, the "Accounting papers corpus" was larger, containing 9,873 words compared to 6,387 words in the "Accounting book corpus." This indicated that the paper corpus was more extensive and likely covered a broader range of accounting topics, providing researchers and practitioners with a larger pool of information for reference and analysis.

Table 4.

Descriptive Results of the Syntactic Complexity Measure in Accounting EAP Textbooks Versus Papers

Measure	Books Corpus	Papers Corpus	Difference
Words (nwords)	6,178	15,264	+9,086
Mean Length of Sentence (MLS)	19.9	28.31	+8.41
Mean Length of T-unit (MLT)	19.1	24.94	+5.84
Mean Length of Clause (MLC)	15.7	11.7	-4.0
Clauses per Sentence (C_S)	1.26	2.41	+1.15
Verb Phrases per T-unit (VP_T)	2.07	2.92	+0.85
Clauses per T-unit (C_T)	1.21	2.13	+0.92

Dependent Clauses per Clause (DC_C)	0.21	0.51	+0.30
Dependent Clauses per T-unit (DC_T)	0.25	1.10	+0.85
T-units per Sentence (T_S)	1.04	1.13	+0.09
Complex T-units per T-unit (CT_T)	0.22	0.656	+0.436
Complex Phrases per T-unit (CP_T)	1.03	0.80	-0.23
Complex Phrases per Clause (CP_C)	0.85	0.37	-0.48
Complex Noun Phrases per T-unit (CN_T)	2.63	3.4	+0.77
Complex Noun Phrases per Clause (CN_C)	2.16	1.61	-0.55

In terms of sentence structure, the "Accounting papers corpus" exhibited longer and more complex sentences than the "Accounting book corpus." The mean length of sentence (MLS) in the papers corpus was 26.46, reflecting greater syntactic intricacy, while the books corpus had a lower MLS of 19.72, indicating shorter and more concise sentences. This difference suggested syntactic discrepancies between EAP books and academic papers in accounting texts corpora.

Both the mean length of T-unit (MLT) and mean length of clause (MLC) were higher in the "Accounting papers corpus," further demonstrating its more elaborate syntactic structures. This linguistic complexity likely stemmed from the academic rigour required to explore theoretical concepts and empirical findings in the papers.

Additionally, the "Accounting papers corpus" showed a greater prevalence of dependent clauses and complex nominal phrases. Higher ratios of dependent clauses per clause (DC_C) and per T-unit (DC_T) highlighted its reliance on these structures for cohesion and detailed information. Similarly, the higher ratios of complex nominal phrases per T-unit (CN_T) and per clause (CN_C) reflected frequent use of intricate nominal constructions, contributing to the precision and specificity expected in scholarly writing.

In the domain of accounting, these findings underscored differences between EAP books and real-life language data as well. The longer, more complex sentences in the papers, alongside a higher density of dependent clauses and complex nominal phrases, emphasized the precision and specificity characteristic of academic writing. Comparable conclusions have been drawn in previous related research as well. The discussion sections of accounting research articles are notably challenging due to their complex rhetorical structures, which integrate both the presentation of research findings and the authors' interpretative analyses, necessitating a nuanced understanding of genre-specific conventions (Amnuai, 2017). Moreover, accounting texts exhibit distinctive lexical and grammatical features that further complicate their interpretation and translation,

particularly in the context of International Financial Reporting Standards (IFRS) (Movsisyan, 2021). These linguistic peculiarities are critical for the precise transmission of accounting concepts across diverse linguistic and cultural contexts, underscoring the importance of both textual and grammatical accuracy in academic discourse (Evans, 2018).

Consistent with these findings, Davidson (2005) and Amnuai (2019) emphasised the importance of precision and specificity in scholarly writing within the field of accounting. Davidson's analysis of accounting textbooks over the past century revealed a decrease in sentence complexity and an increase in word complexity, whereas Amnuai's study of accounting research article abstracts emphasised differences in rhetorical moves and linguistic realisations. These observations supported the recommendation for EAP textbooks to align more closely with real-world academic language in accounting. Loughran and McDonald (2016), and Hussain et al. (2020) further stressed the importance of understanding textual nuances and the dichotomy in source material usage in accounting research.

The "Pharmacy papers corpus" is, similar to other disciplines, richer in textual content, containing 8,431 words compared to 5,749 words in the "Pharmacy book corpus," suggesting that the papers corpus covers a broader range of words and structures. This comprehensive collection of academic papers offers researchers and professionals access to more specialised subjects and detailed research findings.

Table 5.

Descriptive Results of the Syntactic Complexity Measure in Pharmacy EAP Textbooks Versus Papers

Measure	Books Corpus	Papers Corpus	Difference
Words (nwords)	7,483	8,135	+652
Mean Length of Sentence (MLS)	17.90	22.66	+4.76
Mean Length of T-unit (MLT)	18.29	21.57	+3.28
Mean Length of Clause (MLC)	14.58	13.90	-0.68
Clauses per Sentence (C_S)	1.227	1.62	+0.393
Verb Phrases per T-unit (VP_T)	2.01	2.05	+0.04
Clauses per T-unit (C_T)	1.254	1.55	+0.296
Dependent Clauses per Clause (DC_C)	0.19	0.35	+0.16
Dependent Clauses per T-unit (DC_T)	0.24	0.55	+0.31
T-units per Sentence (T_S)	0.97	1.05	+0.08
Complex T-units per T-unit (CT_T)	0.21	0.40	+0.19
Complex Phrases per T-unit (CP_T)	0.98	0.58	-0.40
Complex Phrases per Clause (CP_C)	0.78	0.37	-0.41
Complex Noun Phrases per T-unit (CN_T)	2.70	3.18	+0.48
Complex Noun Phrases per Clause (CN_C)	2.15	2.04	-0.11

In terms of sentence structure, the "Pharmacy papers corpus" exhibits longer and more complex sentences. Its mean length of sentence (MLS) is 25.18, indicating greater syntactic intricacy, whereas the "Pharmacy book corpus" has a lower MLS of 21.36, reflecting shorter and more concise sentence structures. This distinction highlights variations in writing style and potentially differing purposes between academic papers and textbooks.

Further analysis shows that the mean length of T-unit (MLT) and mean length of clause (MLC) are both higher in the "Pharmacy papers corpus" than in the "Pharmacy book corpus," revealing that sentences in the papers corpus are more clause-dense and syntactically elaborate. This complexity reflects the academic rigour and detailed exploration of theoretical or empirical concepts characteristic of pharmacy papers.

Additionally, the "Pharmacy papers corpus" demonstrates a higher frequency of dependent clauses and complex nominal phrases. The ratios of dependent clauses per clause (DC_C) and per T-unit (DC_T) are greater in the papers corpus, indicating a stronger reliance on dependent clauses for cohesion and the provision of supplementary information or evidence. Similarly, the ratios of complex nominal phrases per T-unit (CN_T) and per clause (CN_C) are higher, reflecting the frequent use of intricate nominal constructions.

These findings revealed that pharmacy research papers employed more complex syntactic structures than pharmacy textbooks—characterized by longer sentences, greater T-unit lengths, and higher clause densities. While these results have contributed to the understanding of the distinct stylistic demands of authentic academic written texts, related studies have explored other dimensions of language use within the pharmaceutical field. For example, previous research has demonstrated that pharmacy texts often utilise specialised keywords and frequent text structures (lexical bundles), as well as a dedicated Pharmacy Academic Word List (PAWL), to address the domain-specific linguistic needs of academic reading in this field (Grabowski, 2015; Heidari et al., 2020). Additionally, the discussion sections of pharmacy research articles often follow a structured move analysis, similar to those in applied linguistics. This includes reporting results and commenting on them, which are crucial for constructing arguments and discussing findings (Zainal, 2015). Building upon this, these insights suggest that incorporating both the syntactic complexity and specialised lexical features observed in pharmacy research into EAP textbooks could

more comprehensively equip students to navigate the syntactic demands of real-world academic reading and writing, as identified in the field's research article discussion sections.

4.2 Inferential Statistics of the Findings

To deepen the insights from the descriptive analysis, a one-way MANOVA was conducted to evaluate linguistic differences across six corpora: Psychology books (PsychoBok), Accounting books (AccBok), Pharmacy books (PhrmBok), Psychology papers (PsychoPprs), Accounting papers (AccPprs), and Pharmacy papers (PhrmPprs). The analysis included four dependent variables: the length of the production unit, the amount of subordination, the amount of coordination, and the degree of phrasal sophistication. The results suggested a statistical foundation for the trends identified in the descriptive analysis. Additionally, the plots provided at the end of the analysis visually summarized the relationships between corpus type and dependent variables, illustrating patterns of variation for each linguistic feature across the corpora.

The MANOVA results revealed a significant multivariate effect for corpus type on the syntactic linguistic features (Wilks' Lambda = 0.000, $p < 0.001$), confirming significant linguistic differences across the corpora. As will be discussed, these results confirm discrepancies between EAP textbooks and real-life academic texts, and highlight that academic papers generally employ more complex syntactic and phrasal structures than textbooks. Table 6 presents the multivariate test results from SPSS.

Table 6.
Multivariate Tests Results

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	1.000	10682.348[b]	4.000	3.000	.000
	Wilks' Lambda	.000	10682.348[b]	4.000	3.000	.000
	Hotelling's Trace	14243.130	10682.348[b]	4.000	3.000	.000
	Roy's Largest Root	14243.130	10682.348[b]	4.000	3.000	.000
corpus	Pillai's Trace	2.904	3.179	20.000	24.000	.004
	Wilks' Lambda	.000	23.089	20.000	10.900	.000
	Hotelling's Trace	647.333	48.550	20.000	6.000	.000
	Roy's Largest Root	606.144	727.372[c]	5.000	6.000	.000

According to Table 7, the length of production units, including sentences and clauses, varied significantly across the corpora (F = 25.361, p < 0.001). Academic papers in Psychology (M = 26.31), Accounting (M = 26.63), and Pharmacy (M = 22.12) consistently demonstrated longer production units compared to textbooks. Although MANOVA found significant differences across the corpora based on the length of the production unit, other syntactic complexity measures related to subordination and coordination did not show any meaningful differences. To better evaluate and explain the significance of length of production unit evaluations, and the insignificant means for amount of subordination and coordination in the corpora, Table 7, together with the explanations, is given below. Additionally, the plot (Figure 1) for the length of production units clearly highlights the higher values for Psychology, Accounting, and Pharmacy papers compared to their corresponding book corpora, showcasing the distinction between complex academic writing and simpler textbook language. According to the statistical results, the amount of subordination and coordination did not vary significantly across the corpora. However, the descriptive results table, together with the plots, are presented to shed light on the mean differences between the corpora as well.

Table 7.
Descriptive Results, Mean, and Standard Deviation Across All Corpora

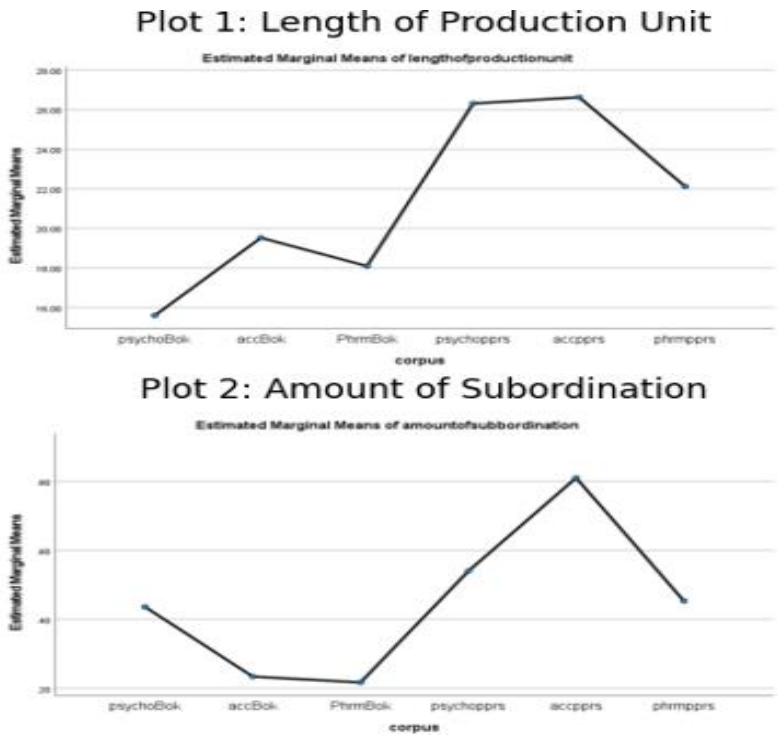
Measure	Corpus	Mean	Standard Deviation
Length of Production Unit	**PsychoBok**	15.60	1.03
	AccBok	19.53	0.57
	PhrmBok	18.10	0.28
	PsychoPprs	26.31	1.31
	AccPprs	26.63	2.39
	PhrmPprs	22.12	0.77
Amount of Subordination	**PsychoBok**	0.44	0.12
	AccBok	0.23	0.03
	PhrmBok	0.22	0.03
	PsychoPprs	0.54	0.19
	AccPprs	0.81	0.41
	PhrmPprs	0.45	0.14
Amount of Coordination	**PsychoBok**	0.51	0.14
	AccBok	0.94	0.13
	PhrmBok	0.89	0.14
	PsychoPprs	0.70	0.25
	AccPprs	0.59	0.30
	PhrmPprs	0.48	0.15
Degree of Phrasal Sophistication	**PsychoBok**	1.46	0.41
	AccBok	2.40	0.33
	PhrmBok	2.43	0.39

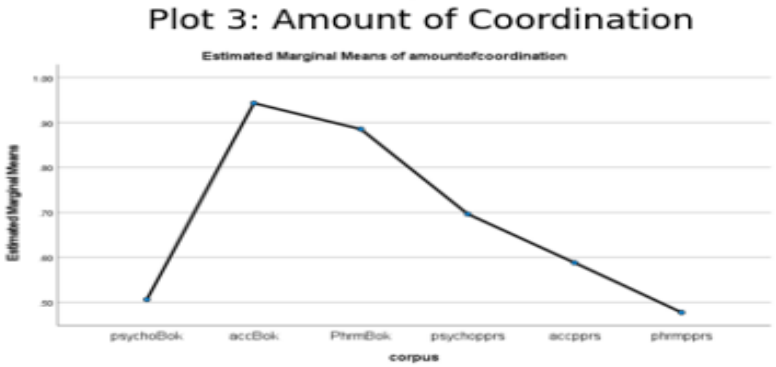
	PsychoPprs	3.12	1.10
	AccPprs	2.54	1.29
	PhrmPprs	2.62	0.80

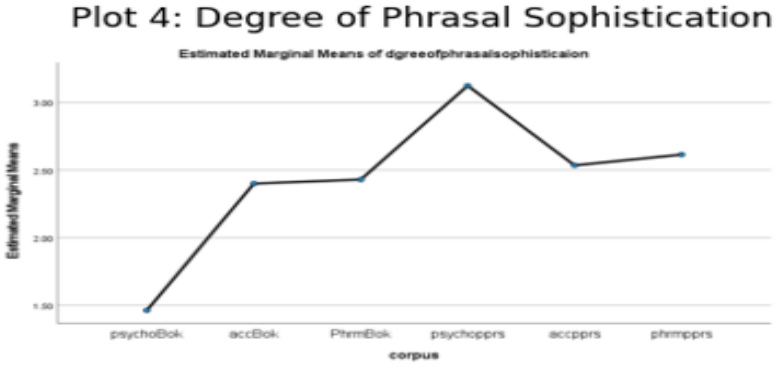
Although MANOVA did not identify statistically significant differences for subordination (F = 2.347, p = 0.164), descriptive trends indicated that papers, particularly in Psychology (M = 0.54) and Accounting (M = 0.81), employed higher ratios of dependent clauses than textbooks. This suggested that papers relied on subordination to achieve cohesion and elaborate on complex ideas. The plot for subordination revealed a slight increase in paper corpora, particularly in Psychology and Accounting papers. Although these differences between the book and paper corpora were less pronounced in the MANOVA results, a look at the right side of the plots, which represented the syntactic complexity of real-life research papers, can suggest notable discrepancies in measures such as the amount of subordination, coordination, and the degree of phrasal complexity when compared to EAP textbooks.

Figure 1.

Plots for Length of Production Unit, Amount of Subordination, Amount of Coordination, and Degree of Phrasal Sophistication Across all Corpora







The degree of phrasal sophistication did not show statistical significance in MANOVA (F = 0.586, p = 0.542). Nonetheless, descriptive trends suggested a higher degree of sophistication in papers, especially in Psychology (M = 3.12), reflecting the field's linguistic demand for precise expression of nuanced concepts. Table 7 provides data on phrasal sophistication. The plot for phrasal sophistication confirmed the descriptive results, with academic papers generally exhibiting greater sophistication than textbooks. The Psychology papers corpus stood out with the highest level of sophistication, supporting the idea that academic papers require more complex linguistic features.

5. Conclusions

The results of this study have revealed notable syntactic discrepancies between EAP textbooks and authentic academic articles across the disciplines of Psychology, Accounting, and Pharmacy. It appears that academic papers have consistently exhibited greater linguistic complexity, as evidenced by longer mean sentence lengths, T-units, and higher clause densities, alongside an increased use of dependent clauses and complex nominal phrases. In contrast, EAP textbooks have tended to employ more simplified and concise constructions that may not fully capture the syntactic features characteristic of real-world academic discourse. These findings suggest that, while it is understandable that

textbooks aim to scaffold students' learning, they may inadvertently underprepared them for the language demands of academic research written texts. This can be one of the reasons why EFL learners face linguistic challenges in academic reading performance at universities, as addressed in the literature review. Ultimately, the data suggest that there is a need for EAP instructional materials to more closely align with the syntactic realities of authentic academic texts, provide more real-life representative texts, and thereby equip students with the skills necessary for effective engagement with complex scholarly texts. Additionally, syntactic complexity studies can cast a more careful light on these differences, thereby identifying potential linguistic challenges and guiding the development of EAP materials that more authentically reflect the demands of academic discourse.

The findings can, therefore, contribute to the ongoing research on EAP textbooks by revealing that these materials often employ simplified syntactic structures compared to authentic academic texts. This discrepancy may help explain why, as noted by Najjar (2020) and Walková (2020), EAP textbooks sometimes fail to address specific language skills and critical grammatical elements adequately. Moreover, previous evaluations have suggested that these textbooks meet only about 50% of effective teaching criteria (Sabiq & Muflihah, 2021) and may lack cultural relevance and appropriate difficulty levels (Ruegg et al., 2018), whereas tailored materials have been shown to enhance student motivation and proficiency (Jou, 2017; Rasekh et al., 2011). Collectively, these results underscore the need for EAP materials that more authentically reflect the syntactic complexity of academic discourse.

One limitation of this study is its reliance on a selected corpus of texts from a limited set of academic disciplines: Psychology, Accounting, and Pharmacy. The texts may not fully capture the complete range of linguistic variation inherent within these disciplines and those in other academic fields. Moreover, by focusing solely on textbooks and research articles, the study has excluded other potentially informative academic genres—such as case studies or conference proceedings—that might reveal additional views in syntactic features. It is also important to note that the findings, while informative, are context-dependent and may not be entirely generalizable to EAP textbooks published in other educational settings or disciplines. Future research could profitably address these

limitations by expanding the corpus to include a broader array of academic materials and by investigating how linguistic features evolve across diverse genres and contexts.

References

Ahmed, Z. A. D. A., & Ahmed, F. E. Y. (2023). Difficulties Encountered by Saudi EFL Learners in English Reading Comprehension. *Russian Law Journal*, *11*(8s). https://doi.org/10.52783/ rlj.v11i8s.1250

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/ 10.1075/scl.59.15ai

Al-Jamal, D. A. H., & Al-Jamal, G. A. (2013). An Investigation of the Difficulties Faced by EFL Undergraduates in Speaking Skills. *English Language Teaching*, *7*, 19-27. https://doi.org/ https://www.semanticscholar.org/paper/An-Investigation-of-the-Difficulties-Faced-by-EFL-Al-Jamal-Al-Jamal/5ed3a40b252fc17f5a22b5d56f5065c66794eba2

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* University of Essex]. https://drepo.sdl.edu.sa/handle/20.500.14154/45778https://drepo.sdl.edu.sa/handle/20.500.14154/4577 8

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

Amnuai, W. (2017). The textual organization of the discussion sections of accounting research articles. *The Kasetsart Journal Social Sciences*, *40*, 389–394-389–394. https://doi.org/10.1016/j.kjss.2017.10.007

Amnuai, W. (2019). Analyses of rhetorical moves and linguistic realizations in accounting research article abstracts published in international and Thai-based journals. *Sage open*, *9*(1), 2158244018822384. https://doi.org/10.1177/2158244018822384

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

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/10.1016/ j.jeap.2009.07.001

Beech, J. R. (2009). *How to write in psychology: A student guide*. Wiley-Blackwell Publishing. https://psycnet.apa.org/record/2007-09159-000

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/llc/8.4.243

Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, *9*(1), 2-20. https://doi.org/10.1016/ j.jeap.2010.01.001

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *Tesol Quarterly*, *45*(1), 5-35. https://doi.org/ 10.5054/tq.2011.244483

Breakwell, G. M., Barnett, J., & Wright, D. B. (2020). *Research Methods in Psychology*. SAGE Publications Ltd. http://digital.casalini.it/9781529731934

Casal, J. E., & Lee, J. J. (2019). Syntactic complexity and writing quality in assessed first-year L2 writing. *Journal of second language writing*, *44*, 51-62. https://doi.org/10.1016/j.jslw.2019.03.005

Casal, J. E., Lu, X., Qiu, X., Wang, Y., & Zhang, G. (2021). Syntactic complexity across academic research article part-genres [Article]. *A cross-disciplinary perspective*, *52*. https://doi.org/10.1016/j.jeap.2021.100996

Charalambous, A. C., & Ed. (2011). *The Role and Use of Course Books in EFL* [Report]. https://eric.ed.gov/?id=ED524247

Chen, X., & Meurers, D. (2018). Word frequency and readability: Predicting the text-level readability with a lexical-level attribute. *Journal of Research in Reading*, *41*(3), 486-510. https://doi.org/10.1111/1467-9817.12121

Coxhead, A., Dang, T. N. Y., & Mukai, S. (2017). Single and multi-word unit vocabulary in university tutorials and laboratories: Evidence from corpora and textbooks. *Journal of English for Academic Purposes*, *30*, 66-78. https://doi.org/10.1016/j.jeap.2017.11.001

Davidson, R. A. (2005). Analysis of the complexity of writing used in accounting textbooks over the past 100 years. *Accounting Education*, *14*(1), 53-74. https://doi.org/10.1080/0963928041000169524l

Deng, Y., Lei, L., & Liu, D. (2020). Calling for More Consistency, Refinement, and Critical Consideration in the Use of Syntactic Complexity Measures for Writing. *Applied linguistics*, *42*. https://doi.org/10.1093/applin/amz069

Evans, L. (2018). Language, translation and accounting: towards a critical research agenda. *Accounting, Auditing & Accountability Journal*, *31*(7), 1844-1873. https://doi.org/10.1108/AAAJ-08-2017-3055

Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: Countering criticisms against corpus-based methodologies. *English for Specific Purposes*, *24*(3), 321-332. https://doi.org/10.1016/j.esp.2004.09.002

Frantz, R. S., Starr, L. E., & Bailey, A. L. (2015). Syntactic Complexity as an Aspect of Text Complexity. *Educational Researcher*, *44*(7), 387-393. https://doi.org/10.3102/0013189x15603980

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/10.31261/TAPSLA.9152

Grabe, W. (2008). *Reading in a second language: Moving from theory to practice*. Cambridge university press. https://doi.org/10.1017/CBO9781139150484

Grabe, W., & Stoller, F. L. (2019). *Teaching and researching reading*. Routledge. https://doi.org/10.4324/9781315726274

Grabowski, Ł. (2015). Keywords and lexical bundles within English pharmaceutical discourse: A corpus-driven description. *English for Specific Purposes*, *38*, 23-33. https://doi.org/10.1016/j.esp.2014.10.004

Hartley, J. (2008). *Academic writing and publishing: A practical handbook*. Routledge. https://doi.org/10.4324/9780203927984

Harwood, N. (2005). What do we want EAP teaching materials for? *Journal of English for Academic Purposes*, *4*(2), 149-161. https://doi.org/10.1016/j.jeap.2004.07.008

Heidari, F., Jalilifar, A., & Salimi, A. (2020). Developing a Corpus-Based Word List in Pharmacy Research Articles: A Focus on Academic Culture. *International Journal of Society, Culture & Language*, *8*(1), 1-15. https://www.ijscl.com/article_38565_2dd1fbcc6d319f297ef976ed5ca73675.pdf

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

Hussain, S., Liu, L. Y. J., & Miller, A. D. (2020). Accounting as a dichotomised discipline: An analysis of the source materials used in the construction of accounting articles. *Critical Perspectives on Accounting*, *66*, 102086. https://doi.org/10.1016/j.cpa.2019.04.007

Hyland, K. (2006). *English for Academic Purposes: An Advanced Resource Book*. Routledge. https://doi.org/10.4324/9780203006603

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/10.1111/modl.12622

Jisu, R., & Moongee, J. (2020). An analysis of the text complexity of listening scripts of high school English textbooks using Coh-Metrix . *Korea Journal of English Language and Linguistics*, *20*(1), 363-383. https://doi.org/ 10.15738/kjell.20..202008.363

Jou, Y. (2017). Hidden challenges of tasks in an EAP writing textbook: EAL graduate students' perceptions and textbook authors' responses. *Journal of English for Academic Purposes*, *30*, 13-25. https://doi.org/10.1016/j.jeap.2017.10.001

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Allyn & Bacon. https://psycnet.apa.org/record/1986-98384-000

Karami, M., & Salahshoor, F. (2013). 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

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

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, *54*(4), 1187-1230. https://doi.org/10.1111/1475-679X.12123

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International journal of corpus linguistics*, *15*(4), 474-496. https://doi.org/10.1075/ijcl.15.4.02lu

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

Martinussen, R., & Mackenzie, G. (2015). Reading comprehension in adolescents with ADHD: Exploring the poor comprehender profile and individual differences in vocabulary and executive functions. *Research in developmental disabilities*, *38*, 329-337. https://doi.org/10.1016/j.ridd.2014.12.007

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

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 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. https://doi.org/https://www.jstor.org/stable/10.3366/j.ctvxcrjmp

Mousavi, M., Ganji, M., & Kordi Tamandani, K. (2021). Let's Judge a Book by its Cover: The Case Study of Active Skills for Reading. *Journal of Modern Research in English Language Studies*, *8*(4), 183-206. https://doi.org/10.30479/jmrels.2021.15774.1919

Movsisyan, A. (2021). Lexical and Grammatical Peculiarities of Accounting Texts and Their Translation. *Translation Studies: Theory and Practice*, *1*(1 (1)), 37-45. https://doi.org/10.46991/TSTP/2021.1.1.45

Najjar, R. (2020). EAP Learners' Perception of Language Needs and its Matches/Mismatches with Textbook Materials. *Iranian Journal of Applied Language Studies*, *12*(2), 161-184. https://doi.org/10.22111/ijals.2020.5968

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

Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of second language writing*, *29*, 82-94. https://doi.org/10.1016/j.jslw.2015.06.008

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

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/10.17509/ijal.v7i2.8352

Rasekh, A. E., Jafarzadeh, M., & Simin, S. (2011). Changing general to specific: on the outcome of using special English materials in university English courses. *English for Specific Purposes World*, *10*(31), 1-18. http://utr.spb.ru/ESP-World/Articles_31/Eslam_Jafarzadeh_Simin.pdf

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/10.1016/0042-6989(95)00132-8

Ruegg, R., Williams, C., & Araki, N. (2018). Developing EAP Materials: Transforming Learning Through In-House Textbooks. In R. Ruegg & C. Williams (Eds.), *Teaching English for Academic Purposes (EAP) in Japan: Studies from an English-medium University* (pp. 167-182). Springer Singapore. https://doi.org/10.1007/978-981-10-8264-19 https://doi.org/10.1007/978-981-10-8264-1910.1007/978-981-10-8264-1_9

Sabiq, A. H. A., & Muflihah, M. (2021). Content Analysis of an EAP Coursebook in an Islamic University. *JETAL: Journal of English Teaching & Applied Linguistic*, *3*(1), 12-20. https://doi.org/10.36655/jetal.v3i1.598

SAMT. (2024a). English for the Students of Accounting (I). SAMT. Retrieved 2024, Jan. 24 from https://samt.ac.ir/fa/book/421/%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C-%D8%A8%D8%B1%D8%A7%DB%8C-%D8%AF%D8%A7%D9%86%D8%B4%D8%AC%D9%88%DB%8C%D8%A7%D9%86-%D8%B1%D8%B4%D8%AA%D9%87-%D8%AD%D8%B3%D8%A7%D8%A8%D8%AF%D8%A7%D8%B1%DB%8C-%DB%B2-

SAMT. (2024b). English for the Students of Pharmacy. SAMT. Retrieved 2024, Jan. 24 from https://samt.ac.ir/fa/book/441/%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C-%D8%A8%D8%B1%D8%A7%DB%8C-%D8%AF%D8%A7%D9%86%D8%B4%D8%AC%D9%88%DB%8C%D8%A7%D9%86-%D8%B1%D8%B4%D8%AA%D9%87-%D8%AF%D8%A7%D8%B1%D9%88%D8%B3%D8%A7%D8%B2%DB%8C

SAMT. (2024c). English for the Students of Psychology. SAMT. Retrieved 2024, Jan. 24 from https://samt.ac.ir/fa/book/1581/%D8%A7%D9%86%DA%AF%D9%84%DB%8C%D8%B3%DB%8C-%D8%A8%D8%B1%D8%A7%DB%8C-%D8%AF%D8%A7%D9%86%D8%B4%D8%AC%D9%88%DB%8C%D8%A7%D9%86-%D8%B1%D8%B4%D8%AA%D9%87-%D8%B1%D9%88%D8%A7%D9%86-%D8%B4%D9%86%D8%A7%D8%B3%DB%8C

Smirnova, E. (2022). Clausal complexity of expert and student writing: a corpus-based analysis of papers in social sciences. *Language Learning in Higher Education*, *12*(2), 453-475. https://doi.org/doi:10.1515/cercles-2022-2052

Sun, H. (2020). *Studies in Applied Linguistics and TESOL*, *20*(2). https://doi.org/10.7916/salt.v20i2.7098

Tarlani-Aliabadi, H., Tazik, K., & Azizi, Z. (2022). Exploring the role of language knowledge and background knowledge in reading comprehension of specific-purpose tests in higher education. *Language Testing in Asia*, *12*(1), 1-23. https://doi.org/10.1186/s40468-022-00198-x

Walková, M. (2020). Transition markers in EAP textbooks. *Journal of English for Academic Purposes*, *46*, 100874. https://doi.org/ 10.1016/j.jeap.2020.100874

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

Wood, D. C., & Appel, R. (2014). Multiword constructions in first year business and engineering university textbooks and EAP textbooks. *Journal of English for Academic Purposes*, *15*, 1-13. https://doi.org/ 10.1016/j.jeap.2014.03.002

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/10.15858/engtea.77.s1.202209.43

Zainal, L. (2015). *Move analysis of pharmacy research article discussion sections* Universiti Teknologi MARA]. https://ir.uitm.edu.my/id/eprint/15430https://ir.uitm.edu.my/id/eprint/15430