Korean Journal of English Language and Linguistics, Vol 25, February 2025, pp. 203-232
DOI: 10.15738/kjell.25..202502.203

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Language Complexity of the Listening Scripts of the High School *English I* and *II* Textbooks Based on the 2015 Revised National Curriculum

Jiyong Lee (Korea Military Academy) · **Hyunoo Lee** (Inha University)

Received: January 2, 2025
Revised: January 24, 2025
Accepted: February 10, 2025

Jiyong Lee (First author)
Assistant Professor, Department of English
Korea Military Academy
574 Hwarang-Ro, Nowon-Gu
Seoul 01805, Korea
Email: jlee0123@kma.ac.kr

Hyunoo Lee (corresponding author)
Emeritus Professor, Department of English Education
Inha University
100 Inha-Ro, Michuhol-Gu,
Incheon 22212, Korea
Email: hylee@inha.ac.kr

ABSTRACT

Lee, Jiyong and Hyunoo Lee. 2025. Language complexity of the listening scripts of the high school English I and II textbooks based on the 2015 Revised National Curriculum. *Korean Journal of English Language and Linguistics* 25, 203-232.

This paper discusses the ideal characteristics of listening materials of English textbooks within the framework of the 2015 Revised National Curriculum of English, emphasizing two key principles: Tolerance (or Allowable Deviation) and Continuity. Tolerance refers to the need for textbooks to strike a balance in linguistic complexity. English textbooks should feature texts that are neither overly complex nor too simplistic, ensuring that they are appropriately challenging compared to other textbooks covering the same subject. This balance helps to maintain a consistent level of difficulty across materials, catering to diverse student capabilities without overwhelming or under-stimulating learners. Continuity suggests that English textbooks should be more linguistically complex and longer in length than those intended for lower-ranking subjects in the curriculum. This progression supports students' development, preparing them for increasingly sophisticated language use as they advance through their education. The study evaluated nine sets of the listening scripts of high school *English I* and *II* textbooks, employing two-way (M)ANOVAs to analyze the effects of the subject and the publisher on six linguistic complexity indices: AG, TTR, MCI, MLC, DP/C, and CP/C. The findings revealed that a considerable number of the nine textbooks failed to meet the established Tolerance and Continuity standards on multiple occasions. This calls for a reevaluation of how English textbooks are chosen and assessed to ensure that they effectively support student learning in line with curriculum expectations. Enhancing the criteria for textbook selection could lead to more effective teaching materials that better serve the diverse needs of students.

KEYWORDS

Continuity, *English I* and *II* textbooks, listening scripts, lexical, morphological and syntactic complexity, Tolerance

1. Introduction

Textbooks play a pivotal role in EFL (English as a Foreign Language) education, serving as essential tools for curriculum implementation, language input, and instructional support. They align with national curricula and standards (Richards 2001), providing a structured framework that ensures teaching aligns with learning objectives. Grounded in Krashen's (1985) *Input Hypothesis*, textbooks offer learners comprehensible input slightly beyond their proficiency, aiding language acquisition through gradual exposure to new forms. They also embody Vygotsky's (1978) Zone of Proximal Development (ZPD), scaffolding learner progress by presenting incremental challenges that bridge knowledge gaps. Empirical studies highlight textbooks as primary resources, especially in under-resourced settings, where they deliver vital language input, structured practice, and cultural insights (Sheldon 1988, Tomlinson 2011). For teachers, particularly those less experienced, textbooks provide professional guidance with lesson plans, activities, and assessments, promoting consistent and standardized teaching across contexts (Graves 2000, Ur 1996). As Rivers (1981) emphasized, textbooks systematically organize learning, presenting content sequentially and incrementally to build skills in reading, writing, listening, and speaking, thus remaining indispensable for effective language education worldwide.

The English textbooks developed under the 2015 Revised National Curriculum (referred to as the 2015 English textbooks) adhere to publication guidelines and certification standards set by the Ministry of Education (2016). These textbooks underwent a rigorous certification process and received high evaluations for their alignment with curriculum requirements, content selection and organization, accuracy and impartiality, as well as the appropriateness of teaching methods and evaluation strategies. Given that they are all designed based on a uniform curriculum framework and evaluated against consistent standards, the 2015 English textbooks are expected to fulfill two core expectations. First, textbooks within the same grade level and subject should maintain consistent linguistic demands by showing no statistically significant differences in critical language complexity indices. Second, they should demonstrate the curriculum's intra- and inter-subject continuity, particularly among textbooks for elementary grades 3-4, elementary grades 5-6, middle school grades 1-3, and those for *Common English*, *English I*, and *English II*. This continuity is achieved through systematic variations in language complexity that reflect a gradual progression aligned with curriculum continuity and developmental appropriateness. Such an approach ensures a seamless transition in linguistic and cognitive growth for students across different educational levels.

Building on these expectations, Lee and Lee (2024) proposed two essential principles for an ideal English textbook: Tolerance (or Allowable Deviation) and Continuity (or Connectivity). Tolerance emphasizes the need for textbooks to strike a balance in linguistic complexity, ensuring they are neither excessively difficult nor overly simplistic compared to other textbooks of the same grade and subject. This balance allows textbooks to provide a consistent level of challenge, accommodating the diverse needs of students while avoiding the risks of overwhelming or under-stimulating learners. Meanwhile, Continuity requires that textbooks exhibit a progressive increase in linguistic complexity and text length as students move up the curriculum hierarchy. This gradual progression aligns with Vygotsky's (1978) ZPD and Krashen's (1985) *Input Hypothesis*, ensuring that students receive appropriately scaffolded challenges. By offering increasingly sophisticated language input, continuity helps prepare learners for the demands of advanced levels of study, fostering their linguistic and cognitive development. Together, these two Principles create a framework for textbooks that effectively support both consistency and progression in language learning.

As highlighted by Lee and Lee (2024), there is a notable lack of empirical studies evaluating whether the texts in the 2015 English textbooks for the same grade and subject meet the Principle of Tolerance. Existing studies are

scarce and focus on specific aspects, such as lexical complexity in listening and reading materials, examined by Lee and Lee (2019) for *Middle School English* and by Lee (2021) for *Common English* textbooks. Lee (2020) investigated the morphological complexity of reading materials in *English I* textbooks, while Kang and Cha (2023) analyzed both lexical and syntactic complexity in listening scripts across all nine *English I* textbooks. Additionally, Min (2023) contributed by exploring the syntactic complexity of listening materials across textbooks for three middle school grades. In contrast, the Principle of Continuity has been the focus of more extensive research. Noteworthy studies include Hwang and Park (2024), who analyzed syntactic complexity in *Elementary Grades 5-6 English* textbooks, and Min (2023), who investigated listening materials across three middle school grades to assess progression. Ryu and Jeon (2020a) utilized Coh-Metrix measures to find whether listening materials of middle school textbooks of different grades meet the Principle of Continuity, and their subsequent study (2020b) expanded this analysis to *Common English*, *English I*, and *English II* textbooks. Yang and Bae (2022) enriched this area of research by examining lexical and syntactic complexity, as well as readability, in the reading materials of textbooks from these subjects.

The studies highlighted above face significant limitations that Lee and Lee (2024) critically examined. First, selectively analyzing textbooks from specific publishers—often chosen for their popularity in schools—while excluding others for the same subject risks distorting findings, especially in studies on Tolerance. This flaw is evident in works like Sohng (2013), Lee and Lee (2019), Lee (2020), Min (2023), and Yang and Bae (2022), as it undermines the representativeness and reliability of results. While this issue may seem minor, a more critical methodological flaw arises in studies on Continuity. Research that examines continuity between the overall language complexity of textbooks for one subject and that of another—such as Ryu and Jeon (2020a, b), Hwang and Park (2024), Sohng (2013), and Yang and Bae (2022)—is fundamentally flawed. As emphasized by Lee and Lee (2024), the principle of continuity is relevant only between textbooks for two consecutive subjects used by students and teachers in the same classrooms. Continuity in this context refers to the systematic progression of language complexity according to learners' developmental levels, focusing on the actual textbooks employed in classroom settings. Even when the overall language complexity of textbooks for one subject aligns with those for another, significant differences may exist between specific textbooks used in classrooms for these subjects. Conversely, textbooks used in consecutive classrooms may exhibit similar language complexity, even if the overall complexity of textbooks for the corresponding subjects differs. This misalignment in focus reduces the practical applicability of such studies, calling for a more refined and context-specific analysis of textbook continuity.

A recurring issue in previous textbook comparison studies, likely stemming from methodological limitations, is the improper handling of statistical analyses. When examining language complexity indices as dependent variables, the analysis must account for the influence of not only subjects like *Common English*, *English I*, and *English II* but also the specific publishers or authors of the textbooks, as these factors significantly impact the complexity indices. Despite this, many studies have employed one-way ANOVA, which fails to consider the interaction between subjects and publishers. A two-way ANOVA is necessary to incorporate both factors and capture their combined effects. Additionally, given the multivariate nature of language complexity indices—often encompassing multiple interrelated metrics—a two-way multivariate ANOVA (MANOVA) is essential for achieving statistically robust and reliable results. MANOVA, as discussed by Meyers et al. (2016), allows for the simultaneous examination of multiple dependent variables, providing a more comprehensive understanding of the data while mitigating the risk of inflated Type I error rates. Adopting this approach would enhance the methodological rigor of textbook comparison studies and yield more accurate insights into language complexity variations.

Building upon the work of Lee and Lee (2024), which focused on evaluating the reading passages in the *English I* and *II* textbooks, this study examines the listening scripts in the same textbooks to assess how well they meet the Principles of Tolerance and Continuity. Specifically, it investigates the lexical, morphological, and syntactic complexity of these listening scripts, analyzing variations across both subject areas (*English I* and *II*) and publishers. The study seeks to provide a deeper understanding of the listening materials' language complexity, addressing whether they maintain a balanced level of difficulty (Tolerance) and exhibit a systematic progression in complexity that aligns with curriculum requirements (Continuity). By focusing on these critical aspects, the research aims to shed light on the extent to which the listening scripts fulfill the pedagogical objectives outlined in the curriculum and support learners' linguistic development effectively. To this end, the study is guided by the following research questions designed to evaluate these dimensions comprehensively:

Research Question 1: Does the lexical complexity of listening scripts in *English I* and *English II* textbooks differ according to the subject and publisher?
Research Question 2: Does the morphological complexity of listening scripts in *English I* and *English II* textbooks differ according to the subject and publisher?
Research Question 3: Does the syntactic complexity of listening scripts in *English I* and *English II* textbooks differ according to the subject and publisher?

As noted in numerous studies, spoken texts differ from written texts in terms of vocabulary, grammar, and discourse (Biber 1988, Chafe and Tannen 1987, Leech and Svartvik 2002, Lim and Jeon 2013). According to Chafe and Tannen (1987), spoken language tends to present content information in a more fragmented manner compared to written language, while written language expresses content more integratively. Furthermore, spoken language is syntactically less complex and conveys content less coherently than written language. Although this study focuses solely on listening scripts, it is crucial to accurately assess their linguistic complexity, as *English I* and *II* are the most important foundational subjects for the English section of the College Scholastic Ability Test.

2. Selected Indices of Language Complexity

A considerable body of research in both L1 and L2 literature has focused on establishing and refining reliable metrics for assessing language complexity, reflecting the central role of complexity in understanding language acquisition and use. Significant progress has primarily been made in vocabulary and syntax, with recent years seeing increased attention to morphology.

Lexical complexity, a key indicator of linguistic richness, reflects the variety and range of words used in a text. It is commonly measured using units such as tokens (individual word occurrences), types (unique words), and word families (groups of related words), as described by Bauer and Nation (1993), Nation (2001), and Read (2000). A central aspect of lexical complexity is lexical diversity, which gauges the variation in word types within a text; higher lexical diversity indicates greater use of unique words with minimal repetition. The traditional measure of lexical diversity, Templin's (1957) type-token ratio (TTR), is calculated by dividing the number of word types by the number of tokens. However, TTR's reliability diminishes with longer texts, as the emergence of new word types slows with increased text length, leading to artificially lower values. To address this limitation, Guiraud (1954) proposed the Guiraud Index (GI), which divides the number of word types by the square root of the number of tokens, offering a more consistent and reliable measure of lexical diversity across texts of varying lengths.

However, Treffers-Daller, Parslow, and Williams (2018) contended that the GI, due to its rigid compensatory formula, may fail to capture subtle variations in learners' lexical diversity. This limitation is particularly pronounced in textbooks that follow curriculum-based vocabulary restrictions, such as limiting the number of word families introduced. Such textbooks tend to present nearly uniform quantities of new target vocabulary, making the differences in lexical diversity across textbooks even subtler than those observed in learners' natural language use. From this perspective, the TTR offers a more reliable metric for evaluating lexical diversity in such controlled contexts. Moreover, for textbooks designed within these vocabulary constraints, the frequency with which vocabulary items are repeated—facilitating reinforcement and retention—often holds greater pedagogical significance than the breadth of lexical diversity. Given these considerations, this study employs TTR as the primary measure of lexical diversity, aligning with the specific focus on vocabulary repetition and learning efficacy in curriculum-driven textbooks.

Lexical sophistication, another essential index for evaluating learners' vocabulary usage, measures the proportion of advanced words within a text, reflecting the depth and breadth of a learner's lexical repertoire. Malvern et al. (2004) highlighted the close relationship between lexical diversity and sophistication, suggesting that an increase in lexical diversity, characterized by the introduction of new words, often coincides with the inclusion of more advanced vocabulary. Building on this concept, Daller et al. (2003) introduced the Advanced Guiraud (AG) as a metric for lexical sophistication. The AG focuses on advanced vocabulary, categorizing words from frequency bands beyond the common 1K or 2K high-frequency groups as advanced. It is calculated by dividing the number of advanced word types by the square root of the total word tokens, effectively normalizing for text length. This study adopts the AG with a slight modification, using the number of advanced word *families* instead of *types*, divided by the square root of total word tokens, providing a more nuanced method to evaluate textbooks' vocabulary usage in line with educational goals of the 2015 English curriculum.

While the study of morphological complexity has a relatively shorter history compared to research on lexical or syntactic complexity, it has gained significant traction in recent years. Researchers have predominantly focused on inflectional diversity, which distinguishes between the various inflected forms of a word and the broader category of its lemma, encompassing all its inflected and base forms (Brezina and Pallotti 2019, Malvern et al. 2004, Xanthos and Gillis 2010, Xanthos and Guex 2015). Morphological complexity, however, extends beyond inflectional diversity to include derivational diversity, reflecting a broader spectrum of linguistic richness. To capture this, the morphological complexity index (MCI) is used, defined as the ratio of the number of families to the number of types within a text. While adherence to curriculum-mandated vocabulary limits ensures similar family counts across textbooks, the variation lies in the presentation of inflected and derived forms. Textbooks incorporating a greater range of these forms typically exhibit lower MCIs, indicating a richer exposure to morphological variations, which can enhance learners' understanding of word structure and usage.

Syntactic complexity, a cornerstone of linguistic analysis, has been studied as extensively as lexical complexity, reflecting its central role in understanding language development. It pertains to the variety and sophistication of grammatical structures used in language, encompassing elements such as clause structure, subordination, coordination, and phrase complexity (Ortega 2003). As a key indicator of linguistic proficiency, syntactic complexity serves multiple purposes: it helps track developmental stages in language acquisition, facilitates comparisons of linguistic patterns across learner groups, and evaluates the efficacy of different instructional methods on language learning outcomes (Norris and Ortega 2009, Ortega 1999). By providing insights into how learners progress from simpler to more advanced syntactic constructions, it offers a robust framework for exploring the interplay between linguistic competence and pedagogical strategies. This makes syntactic complexity a critical focus for both theoretical research and practical applications in language education.

Lu (2010) introduced 14 syntactic complexity indices through the L2 Syntactic Complexity Analyzer (L2SCA), an automated tool designed to measure syntactic complexity in L2 texts. These indices are based on three key production units: sentence (S), clause (C), and T-unit. A clause (C) consists of a subject and a verb phrase, forming a complete idea. When a clause is independent, it can also function as a sentence (S). Clauses are categorized as independent or dependent (DCs), with the latter subordinated to a main clause. Importantly, clauses containing subjects and finite verbs are classified as clauses, while those with non-finite verb phrases are counted as phrases. A T-unit, as defined by Hunt (1965), is the shortest grammatically complete unit, generally a main clause not joined by a coordinating conjunction. T-units are further divided into simple T-units, which lack subordinate clauses, and complex T-units (CTs), which include them. L2SCA's syntactic complexity indices are grouped into five categories: output length, sentence complexity, subordination, coordination, and phrase sophistication. In this study, the mean length of clause (MLC) was used as a metric for syntactic complexity, calculated by dividing the total number of word tokens by the total number of clauses, offering insights into the structural depth of clauses.

Research on syntactic complexity consistently shows that learners with greater fluency tend to produce texts featuring longer and more intricate syntactic structures (Lu 2010, 2011, Ortega 2003, Wolfe-Quintero, Inagaki and Kim 1998). Lu (2011) highlights a key trend: as fluency increases, learners often produce texts with lower ratios of clause to sentence, dependent clause to clause, and dependent clause to T-unit. This pattern suggests a shift in focus among more fluent English learners from clause-level complexity to phrase-level sophistication. Such learners tend to prefer using coordinate clauses or complex phrases over dependent clauses (Ortega 2003, Wolfe-Quintero et al. 1998). Building on these insights, this study employed two specific metrics—dependent clauses per clause (DC/C) and coordinated phrases per clause (CP/C)—to evaluate the syntactic complexity of listening scripts from *English I* and *II* textbooks. These measures help capture variations in syntactic structuring, offering a better understanding of how textbooks align with learners' syntactic development across different proficiency levels.

In summary, this study employed a total of six measures to measure the language complexity of the listening scripts from *English I* and *II* textbooks. Specifically, AG and TTR were utilized to measure lexical complexity, MCI to measure morphological complexity, and MLC, DC/C, and CP/C to measure syntactic complexity.

3. Methodology

3.1 Textbooks to Analyze

As mentioned earlier, this study analyzed the language complexity of listening scripts in the *English I* and *English II* textbooks, which are based on the 2015 Revised National curriculum and are currently in use at high schools. The analysis focused only on the listening scripts related to the listening and speaking sections of each textbook. The listening scripts associated with unit assessments or those that appear intermittently in relation to culture or reading passages were not examined in the study.

There are nine *English I* textbooks published by nine different publishers, and similarly, nine *English II* textbooks published by the same publishers. The lead authors of the *English I* and *II* textbooks from the same publisher are the same, but the contributors to the textbooks published by four of the publishers have changed. Except for one publisher's textbooks, the nine *English I* and *II* textbooks consist of six units. The textbooks from the remaining publishers consist of five units each for both *English I* and *English II*.

3.2 Hypotheses

English I and *English II* are core subjects of the Elective-Based Curriculum of the 2015 Revised National Curriculum of English. To examine whether the language complexity of the listening scripts in these subjects' textbooks varies according to the subjects and the publishers (or textbook authors), the following hypotheses regarding the *English I* and *II* textbooks were tested:

Hypothesis 1: The AG of the listening scripts varies according to the subjects and publishers.
Hypothesis 2: The TTR of the listening scripts varies according to the subjects and publishers.
Hypothesis 3: The MCI of the listening scripts varies according to the subjects and publishers.
Hypothesis 4: The MLC of the listening scripts varies according to the subjects and publishers.
Hypothesis 5: The DC/C of the listening scripts varies according to the subjects and publishers.
Hypothesis 6: The CP/C of the listening scripts varies according to the subjects and publishers.

The 2015 Revised Curriculum of English specifies that *English I* should cover a vocabulary range of 2,000 words, while *English II* expands this to 2,500 words, naturally suggesting a difference in language complexity between the listening scripts of these two subjects. However, as an anonymous reviewer astutely noted, publishers producing educational materials must adhere to uniform publication guidelines and certification standards mandated by the Ministry of Education (2016), implying uniformity in language complexity across publishers' textbooks. To resolve this tension, the study prioritized the distinctions between subjects over the influence of publishers. Accordingly, it established alternative hypotheses to test differences in the language complexity of listening scripts in *English I* and *English II* textbooks.

The testing of research hypotheses 1–3, related to lexical and morphological complexity, was based on the AG, TTR, and MCI of each textbook unit. The testing of research hypotheses 4–6, related to syntactic complexity, was based on the MLC, DC/C, and CP/C of each unit.

To calculate the AG, TTR, and MCI of each textbook unit, the Basic Vocabulary Guidelines of the 2022 Revised National Curriculum of English were used. As discussed in the previous section, these guidelines recommend that the 800 vocabulary items marked with an asterisk (*) from the most basic 3,000-word group required to meet the achievement standards of the English curriculum be used in elementary school, the 1,200 vocabulary items marked with a double asterisk (**) be used in middle and high school common subjects, and the remaining 1,000 vocabulary items be used in other subjects. These guidelines also exclude from both the 3,000-word basic vocabulary or other learning vocabulary proper nouns, titles, interjections, Romanized Korean or non-English foreign words, alphabets, characters, unit names, abbreviations, chemical formulas, cardinal and ordinal numbers.

Accordingly, this study divided the vocabulary used in each textbook unit into five groups: basewrd1 to basewrd5. Basewrd1 consists of exceptional words mentioned above, basewrd2 includes vocabulary marked with an asterisk (*) in the curriculum, basewrd3 includes vocabulary marked with a double asterisk (**), basewrd4 includes the remaining vocabulary in the curriculum, and basewrd5 consists of words that are neither exceptional words nor part of the basic vocabulary from the curriculum. Among these five groups, the words in basewrd3 to basewrd5 were treated as advanced vocabulary. In other words, words that are considered higher-level than those recommended for elementary school use in the curriculum were classified as advanced vocabulary for listening.

To illustrate how the AG, TTR, and MCI are calculated for each textbook unit, an example of *English I* Textbook A, Lesson 1 is shown below.

Table 1. AG, TTR, and MCI of *English I* Textbook A, Lesson 1

Group	Token	Type	Family	AG	TTR	MCI
basewrd1	16	14	14	2.407	.402	.799
basewrd2	603	201	149			
basewrd3	70	58	53			
basewrd4	6	4	4			
basewrd5	12	7	7			
Total	707	284	227			

AG 2.407 is the value obtained by dividing the sum of the word family count of basewrd3–5 (64) by the square root of the total word token count (707), then rounding the result. TTR .402 is the value obtained by dividing the total word type count (284) by the total word token count (707), then rounding the result. MCI .799 is the value obtained by dividing the total word family count (227) by the word type count (284), then rounding the result.

The AG, TTR, and MCI values for each lesson of the *English I* and *English II* textbooks, manually calculated using the methods illustrated above, are presented in Tables 2 and 3, alongside with the MLC, DC/C, and CP/C values, which were computed using Lu's (2011) SCA.

Table 2. Indices of Language Complexity of the *English I* Textbooks, by Publisher and Lesson

Publisher	Lesson	AG	TTR	MCI	MLC	DC/C	CP/C
A	L1	2.407	.402	.799	6.103	.336	.043
	L2	3.621	.404	.811	6.702	.371	.105
	L3	3.351	.436	.801	7.018	.279	.072
	L4	2.113	.374	.798	6.314	.234	.036
	L5	2.875	.427	.83	7.463	.288	.125
	L6	2.6	.343	.806	6.129	.258	.08
B	L1	2.673	.452	.856	6.207	.23	.103
	L2	2.661	.46	.816	7.253	.147	.187
	L3	2.875	.454	.832	6.813	.2	.133
	L4	2.349	.388	.8	6.796	.226	.118
	L5	3.264	.479	.854	7.6	.114	.186
	L6	2.252	.424	.833	5.835	.306	.047
C	L1	2.013	.456	.839	6.413	.253	.027
	L2	2.946	.494	.868	6.574	.191	.176
	L3	2.502	.524	.854	6.83	.208	.019
	L4	2.529	.491	.844	6.924	.136	.152
	L5	2.796	.487	.836	6.263	.132	.092
	L6	1.994	.476	.828	7.159	.116	.072
D	L1	2.412	.465	.833	6.721	.256	.07
	L2	3.356	.471	.866	6.889	.272	.062
	L3	3.571	.476	.851	6.917	.179	.083
	L4	2.414	.47	.838	7.026	.195	.026
	L5	3.054	.489	.826	6.793	.276	.069
	L6	3.589	.49	.847	7.282	.205	.09
E	L1	3.035	.514	.831	6.527	.284	.149
	L2	2.678	.51	.841	5.827	.296	.062
	L3	3.172	.503	.837	6.52	.2	.133
	L4	2.717	.482	.823	6.405	.241	.089
	L5	3.034	.517	.83	6.788	.288	.136

F	L1	1.436	.479	.822	5.576	.169	.051
	L2	2.007	.543	.867	7	.205	.136
	L3	1.892	.508	.841	6.288	.288	.077
	L4	2.318	.497	.842	6.765	.216	.176
	L5	1.583	.491	.825	7.658	.132	.105
	L6	2.417	.465	.874	6.246	.188	.043
G	L1	2.148	.41	.822	6.33	.28	.07
	L2	2.818	.43	.794	7.085	.287	.138
	L3	2.32	.403	.812	6.51	.26	.11
	L4	2.451	.438	.875	7.631	.274	.167
	L5	3.999	.475	.842	6.989	.213	.18
	L6	3.403	.51	.847	6.688	.338	.104
H	L1	1.83	.475	.848	5.932	.169	.034
	L2	1.62	.501	.82	6.596	.192	.077
	L3	1.756	.5	.855	5.661	.254	.034
	L4	2.943	.487	.8	6.513	.346	.051
	L5	2.133	.471	.813	6.37	.151	.123
	L6	2.713	.47	.861	7.646	.262	.046
I	L1	1.464	.45	.809	6.051	.177	.063
	L2	1.731	.456	.847	6.159	.206	.016
	L3	2.743	.501	.862	7.222	.175	.095
	L4	2.421	.475	.837	6.692	.269	.09
	L5	2.891	.485	.857	7.03	.167	.121
	L6	2.138	.511	.883	6.321	.304	.089
Mean		2.567	.468	.836	6.661	.231	.093

In Table 2, the average AG for the lessons is 2.567, with a maximum value of 3.999 in Lesson 5 of G and a minimum value of 1.436 in Lesson 1 of F. The average TTR for the lessons is .468, with a maximum value of .543 in Lesson 2 of F and a minimum value of .343 in Lesson 6 of A. The average MCI for the lessons is .836, with a maximum value of .883 in Lesson 6 of I and a minimum value of .794 in Lesson 2 of G. The average MLC for the lessons is 6.661, with a maximum value of 7.658 in Lesson 5 of F and a minimum value of 5.576 in Lesson 1 of F. The average DC/C for the lessons is .231, with a maximum value of .371 in Lesson 2 of A and a minimum value of .114 in Lesson 5 of B. The average CP/C for the lessons is .093, with a maximum value of .187 in Lesson 1 of B and a minimum value of .016 in Lesson 2 of I.

Table 3. Indices of Language Complexity of the *English II* Textbooks, by Publisher and Lesson

Publisher	Lesson	AG	TTR	MCI	MLC	DC/C	CP/C
A	L1	3.244	.345	.761	6.606	.298	.106
	L2	3.46	.325	.813	6.426	.274	.061
	L3	4.202	.351	.765	7.02	.325	.044
	L4	3.826	.342	.798	6.933	.249	.096
	L5	2.767	.328	.797	6.623	.257	.057
	L6	4.69	.374	.814	6.557	.373	.097
B	L1	3.007	.419	.835	7.232	.263	.131
	L2	3.229	.424	.82	6.533	.271	.103
	L3	3.587	.476	.853	7.275	.231	.165
	L4	2.747	.447	.784	7.044	.289	.089
	L5	2.938	.448	.827	7.636	.239	.114
	L6	3.664	.468	.838	7.62	.207	.12

C	L1	2.698	.478	.855	6.794	.265	.074
	L2	2.401	.477	.87	5.434	.265	.06
	L3	1.662	.433	.798	5.511	.25	.023
	L4	2.798	.502	.84	6.657	.271	.0
	L5	2.797	.515	.838	6.662	.191	.029
	L6	2.211	.469	.808	6.708	.154	.031
D	L1	3.065	.448	.813	6.225	.265	.069
	L2	3.959	.458	.859	7.098	.283	.065
	L3	4.17	.523	.85	6.667	.287	.08
	L4	3.531	.477	.838	6.433	.227	.072
	L5	3.891	.51	.855	6.419	.29	.161
	L6	3.846	.487	.842	6.778	.278	.111
E	L1	3.496	.513	.859	7.615	.323	.108
	L2	3.203	.472	.831	7.847	.347	.153
	L3	2.908	.491	.798	6.317	.28	.024
	L4	2.641	.477	.817	6.213	.258	.09
	L5	3.648	.508	.83	6.974	.282	.141
F	L1	2.584	.479	.82	7.961	.196	.078
	L2	2.415	.444	.848	6.013	.2	.027
	L3	2.397	.434	.811	5.632	.333	.011
	L4	2.402	.452	.845	6.972	.239	.07
	L5	3.415	.46	.833	6.725	.238	.038
	L6	3.146	.481	.839	6.683	.35	.05
G	L1	3.286	.438	.83	7.83	.216	.125
	L2	3.457	.425	.799	7.419	.229	.143
	L3	3.725	.457	.813	8.224	.282	.282
	L4	3.065	.421	.804	7.465	.287	.139
	L5	4.086	.433	.828	8.05	.24	.21
	L6	3.026	.429	.843	7.703	.231	.165
H	L1	2.916	.552	.829	8.021	.208	.125
	L2	2.503	.478	.814	6.792	.25	.097
	L3	2.948	.514	.824	7.984	.295	.131
	L4	2.465	.542	.832	7.98	.26	.06
	L5	2.238	.539	.814	7.192	.25	.115
	L6	2.233	.567	.817	7.122	.265	.061
I	L1	2.531	.445	.85	7.023	.186	.07
	L2	3.021	.466	.89	7.654	.192	.141
	L3	2.097	.447	.823	6.075	.125	.075
	L4	2.302	.436	.798	6.387	.204	.097
	L5	3.411	.487	.848	7.966	.276	.172
	L6	3.242	.43	.798	8.151	.226	.237
Mean		3.079	.457	.826	6.998	.256	.098

In Table 3, the average AG for the lessons is 3.079, with a maximum value of 4.69 in Lesson 6 of A and a minimum value of 1.662 in Lesson 3 of C. The average TTR for the lessons is .457, with a maximum value of .567 in Lesson 6 of H and a minimum value of .325 in Lesson 2 of A. The average MCI for the lessons is .826, with a maximum value of .89 in Lesson 2 of I and a minimum value of .761 in Lesson 1 of A. The average MLC for the lessons is 6.998, with a maximum value of 8.224 in Lesson 3 of G and a minimum value of 5.434 in Lesson 2 of C. The average DC/C for the lessons is .256, with a maximum value of .373 in Lesson 6 of A and a minimum value of .125 in Lesson 3 of I. The average CP/C for the lessons is .098, with a maximum value of .282 in Lesson 3 of G and a minimum value of .0 in Lesson 4 of C.

3.3 Statistical Analysis

Research Hypotheses 1–6 were tested using Analysis of Variance (ANOVA). To avoid an increase in Type I errors due to multiple tests, a multivariate analysis of variance (MANOVA) was conducted, treating the indices of the research hypotheses as multiple dependent variables. However, for MANOVA to be valid, the correlations among the dependent variables should neither be too strong nor too weak. Therefore, the strength of the correlations between the indices used in the hypotheses was examined using Pearson's bivariate correlation analysis. The correlation coefficients among the indices are shown in Table 4.

Table 4. Pearson's Bivariate Coefficient of Correlation Between Indices of Language Complexity

	AG	TTR	MCI	MLC	DC/C	CP/C
AG	1	-.157	-.017	.401**	.377**	.393**
TTR	-.157	1	.528**	.128	-.182	.050
MCI	-.017	.528**	1	.054	-.207*	.057
MLC	.401**	.128	.054	1	-.053	.594**
DC/C	.377**	-.182	-.207*	-.053	1	-.076
CP/C	.393**	.050	.057	.594**	-.076	1

* Correlation is significant at the .05 level (two-tailed).
** Correlation is significant at the .01 level (two-tailed).

Based on various interpretations of the strength of Pearson's correlation coefficient, (Chan 2003, Dancey and Reidy 2020, De Vaus 2002), this study classified the correlation as moderate if the absolute value of *r* fell between |±.3| and |±.7|. According to this criterion, the indices with moderate correlations in Table 4 were AG-MLC (*r* = .401), AG-DC/C (*r* = .377), AG-CP/C (*r* = .393), TTR-MCI (*r* = .528), and MLC-CP/C (*r* = .594). These pairs of indices alone satisfy the conditions for MANOVA dependent variables. Therefore, to minimize the number of MANOVA tests, two MANOVA analyses were conducted: one with AG, MLC, and CP/C as dependent variables, and the other with TTR and MCI as dependent variables. For the remaining index DC/C, a two-way ANOVA was performed to test the corresponding research hypothesis.

4. Result and Discussion

4.1 Testing Hypotheses 1, 4, and 6

To test Hypotheses 1, 4, and 6, a two-way MANOVA was conducted using measures of AG, MLC, and CP/C as dependent variables. This analysis employed two levels of Subject (*English I* and *English II*) and nine levels of Publisher (A, B, C, D, E, F, G, H, I) as independent variables.

No outliers were detected in the data, and all Subject or Publisher groups but two passed the normality tests. For sample sizes below 50, the Shapiro-Wilks test was used, while for sample sizes of 50 or more, the Kolmogorov-Smirnov test was applied. In the Shapiro-Wilks test, the C group of MLC had a *p*-value of .037. However, skewness (-1.245) and kurtosis (.877) were within the range of |±2|, indicating no issue with performing the MANOVA test. The D group of CP/C had a *p*-value of .048, with skewness of 1.284 (within |±2|). However, its kurtosis was 3.594, exceeding the |±2| range. Therefore, a normal Q-Q plot was inspected for further assessment. As shown in Figure

1 below, the data points appeared relatively close to the diagonal line. Thus, it was deemed appropriate to proceed with the MANOVA test while paying extra attention to subsequent statistical interpretations.

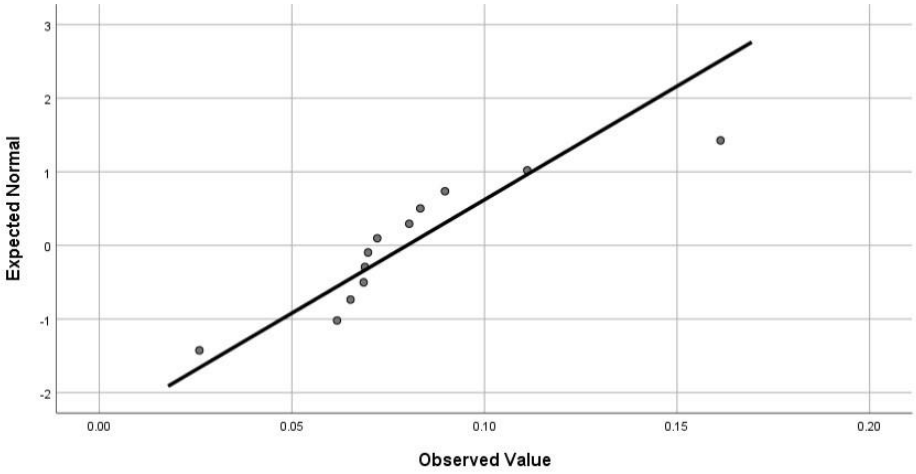

Figure 1. Normal Q-Q Plot of CP/C for Publisher D

Box's test for the equality of covariance matrices revealed that Box's M = 145.449, p > .315, indicating no statistical significance. This result suggests that the variance-covariance matrices of the dependent variables are equal across all levels of the independent variables, supporting the use of Wilks' lambda for multivariate analysis. Meanwhile, Bartlett's test of sphericity yielded an approximate chi-square value of 389.572, p < .001, which was significant. This indicates that there is sufficient correlation among the dependent variables for conducting a multivariate test.

Using Wilks' lambda (see Table 5), the dependent variate was significantly affected by the main effects of Subject, Wilk's lambda = .698, F(3, 86) = 12.412, p < .001, partial η^2 = .302, and Publisher, Wilk's lambda = .296, F(24, 250.027) = 5.429, p < .001, partial η^2 = .333. The multivariate interaction effect of Subject * Publisher was also statistically significant, Wilk's lambda = .564, F(24, 250.027) = 2.276, p < .001, partial η^2 = .174.

Table 5. Multivariate and Univariate Analyses of Variance for Measures of AG, MLC, and CP/C

	Multivariate	Univariate		
Source	F [a]	AG[b]	MLC[b]	CP/C[b]
Subject	12.412*	29.452*	10.099*	.299
Publisher	5.429*	9.061*	2.468*	5.720*
Subject * Publisher	2.276*	.985	2.639*	2.707*
MSE		.231	.304	.002

Note. Multivariate f-ratios were generated from Wilks' lambda.
a. Multivariate *df* = 3, 86(subject), 24, 250.027(publisher), 24, 250.027(subject * publisher)
b. Univariate *df* = 1, 88(subject), 8, 88(publisher), 8, 88(subject * publisher)
*p < .025

The analysis revealed that most of the variance in the data (81%) could be attributed to the effects of the factors under study, with only 19% remaining unexplained or due to error. However, Levene's test indicated a violation of the equality of error variances assumption for MLC (p = .032), requiring adjustments in the analysis. Specifically,

univariate effects for AG and CP/C were tested at the standard alpha level of .05, while MLC was evaluated at a stricter alpha level of .025 (.05 divided by 2) to account for this issue. The results showed that the Subject factor significantly influenced AG ($F(1, 88) = 29.452$, $p < .001$, partial $\eta^2 = .251$) and MLC ($F(1, 88) = 10.099$, $p < .002$, partial $\eta^2 = .103$), but not CP/C. The Publisher factor significantly affected all three dependent variables: AG ($F(8, 88) = 9.061$, $p < .001$, partial $\eta^2 = .452$), MLC ($F(8, 88) = 2.468$, $p < .018$, partial $\eta^2 = .183$), and CP/C ($F(8, 88) = 5.720$, $p < .001$, partial $\eta^2 = .342$). Interaction effects between Subject and Publisher were not significant for AG but were significant for MLC ($F(8, 88) = 2.639$, $p < .012$, partial $\eta^2 = .193$) and CP/C ($F(8, 88) = 2.707$, $p < .010$, partial $\eta^2 = .198$). These findings highlight the distinct influences of Subject and Publisher on language complexity, with interaction effects present for specific indices.

Although no interaction effect was observed, the significant main effects of Subject and Publisher on AG provide support for Hypothesis 1, as AG measures in *English II* textbooks ($M = 3.081$, $SE = .066$, 95% CI = [2.950, 3.213]) were significantly higher than those in *English I* textbooks ($M = 2.573$, $SE = .066$, 95% CI = [2.442, 2.705]). A Bonferroni post hoc test was conducted to determine which publishers' textbooks differed significantly in AG measures. The results showed that textbooks from Publisher D ($M = 3.405$, $SE = .139$, 95% CI = [3.129, 3.681]), A ($M = 3.263$, $SE = .139$, 95% CI = [2.987, 3.539]), and G ($M = 3.149$, $SE = .139$, 95% CI = [2.873, 3.425]) had significantly higher AG measures than those from I ($M = 2.499$, $SE = .139$, 95% CI = [2.223, 2.775]), C ($M = 2.445$, $SE = .139$, 95% CI = [2.170, 2.721]), H ($M = 2.358$, $SE = .139$, 95% CI = [2.082, 2.634]), and F ($M = 2.344$, $SE = .139$, 95% CI = [2.058, 2.610]). Additionally, textbooks from Publisher E ($M = 3.053$, $SE = .152$, 95% CI = [2.751, 3.356]) had significantly higher AG measures than those from H and F. See Table 6 and Figure 2.

Table 6. Mean Scores and Stand Errors for Measures of AG as a Function of Publisher

Group	*M*	*SE*	*N*
Publisher D	3.405_{a}	.139	12
Publisher A	3.263_{a}	.139	12
Publisher G	3.149_{a}	.139	12
Publisher E	3.053_{b}	.152	10
Publisher I	2.499_{a}	.139	12
Publisher C	2.445_{a}	.139	12
Publisher H	$2.358_{a, b}$	.139	12
Publisher F	$2.344_{a, b}$	.139	12

Note. Means sharing subscripts are significantly different at the .05 level by means of a Bonferroni post hoc test.

The significant main effects of Subject and Publisher on MLC, along with their interaction effect, provide strong support for Hypothesis 4. Specifically, MLC scores were significantly higher in *English II* textbooks ($M = 6.998$, $SE = .076$, 95% CI = [6.847, 7.148]) than in *English I* textbooks ($M = 6.657$, $SE = .076$, 95% CI = [6.506, 6.808]), suggesting that *English II* textbooks place greater emphasis on or present content with longer clauses. However, the interaction effect reveals that the influence of Subject on MLC is not uniform across Publisher; rather, the degree to which clause length is emphasized in *English I* and *II* textbooks varies depending on the publishers of textbooks.

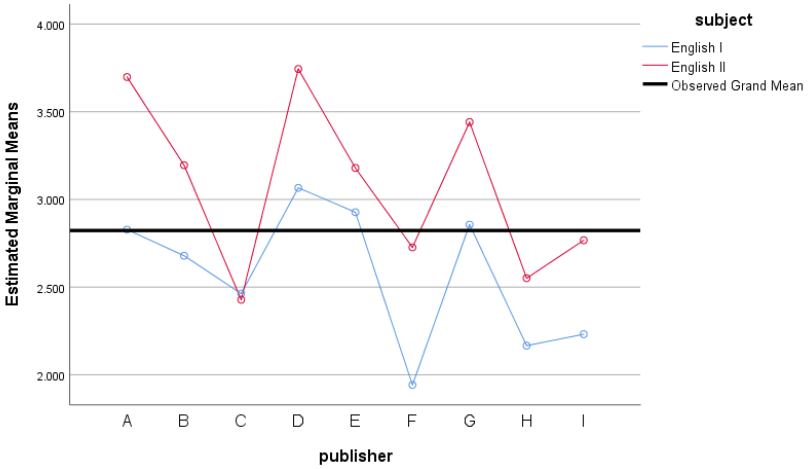

Figure 2. Estimated Marginal Means of AG

The interaction effect for MLC, illustrated in Figure 3, was further analyzed through simple effects analysis and the Tamhane T2 post hoc test, chosen due to the unequal variances among the groups. Table 7 presents the relevant means, highlighting that the significant interaction effect was primarily driven by *English II*-G (M = 7.782, SE = .130) having substantially higher MCL scores compared to *English I*-C (M = 6.694, SE = .137), *English I*-E (M = 6.413, SE = .159), *English II*-A (M = 6.694, SE = .231), and *English II*-D (M = 6.603, SE = .127). These variations point to the complexity of the interaction but also necessitate caution in interpretation due to the observed heterogeneity of variances among the groups in the current analysis.

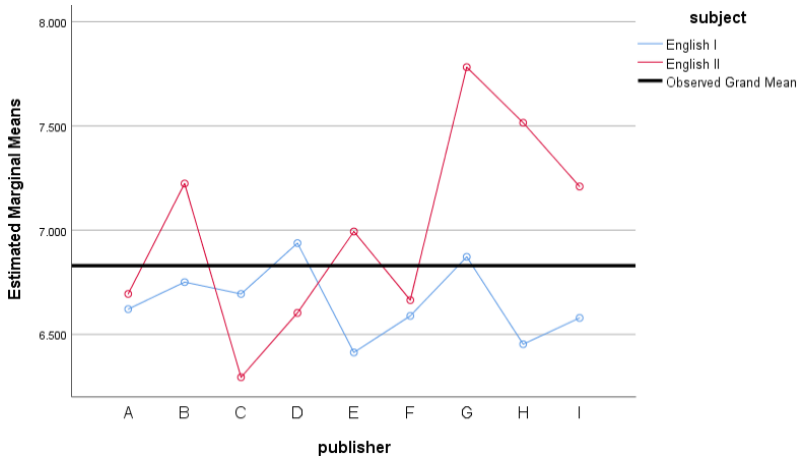

Figure 3. Estimated Marginal Means of MLC

Table 8. Homogeneous Subsets for Publisher Groups of MLC

Ryan-Einot-Gabriel-Welsch Range[a,b]			
		Subset	
Publisher	N	1	2
C	12	6.494	
F	12	6.627	
A	12	6.658	6.658
E	10	6.703	6.703
D	12	6.771	6.771
I	12	6.894	6.894
H	12	6.984	6.984
B	12	6.987	6.987
G	12		7.327
Sig.		1.000	.481

Means for groups in homogeneous subsets are displayed.
Based on observed means.
The error term is Mean Square (Error) = .304.

a. Critical values are not monotonic for these data. Substitutions have been made to ensure monotonicity. Type I error is therefore smaller.
b. Alpha = .05

Combined with the significant main effect of Publisher on CP/C, the significant interaction effect supports Hypothesis 6. The interaction effect for CP/C, visualized in Figure 4, was analyzed using simple effects analysis and the Ryan, Einot, Gabriel, Welch Studentized Range post hoc test, chosen due to the equal variances across groups. Table 9 highlights the relevant means, revealing that *English II*-G (M = .177, SE = .017) exhibited substantially higher CP/C scores compared to several other combinations, including *English II*-A (M = .077, SE = .017), *English II*-C (M = .036, SE = .017), *English II*-F (M = .046, SE = .017), as well as *English I* counterparts such as *English I*-A (M = .077, SE = .017), *English I*-D (M = .067, SE = .017), *English I*-H (M = .061, SE = .017), and *English I*-I (M = .079, SE = .017). Conversely, *English II-C* demonstrated significantly lower CP/C scores than *English I*-B (M = .129, SE = .017), *English I*-G (M = .128, SE = .017), and *English II*-I (M = .132, SE = .017). These patterns collectively contributed to the significant interaction effect, accounting for approximately 55% of the variations in CP/C with the main effect, underscoring the complexity and variability in how subject and publisher jointly influence CP/C measures.

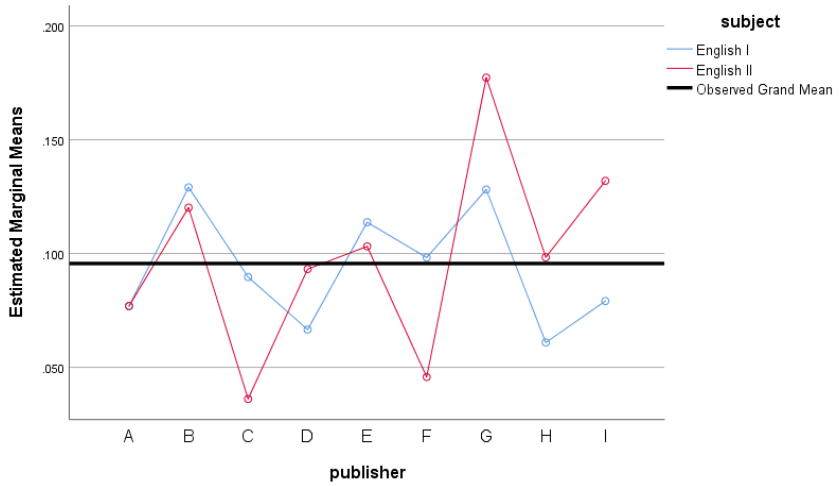

Figure 4. Estimated Marginal Means of CP/C

Table 9. Homogeneous Subsets for All CP/C Groups

Ryan-Einot-Gabriel-Welsch Range[a]				Subset for alpha = .05
Group	N	1	2	3
II-C	6	.036		
II-F	6	.046	.046	
I-H	6	.061	.061	
I-D	6	.067	.067	
I-A	6	.077	.077	
II-A	6	.077	.077	
I-I	6	.079	.079	
I-C	6	.090	.090	
II-D	6	.093	.093	.093
I-F	6	.098	.098	.098
II-H	6	.098	.098	.098
II-E	5	.103	.103	.103
I-E	5	.114	.114	.114
II-B	6	.120	.120	.120
I-G	6		.128	.128
I-B	6		.129	.129
II-I	6		.132	.132
II-G	6			.177
Sig.		.065	.056	.050

Means for groups in homogeneous subsets are displayed.

a. Critical values are not monotonic for these data. Substitutions have been made to ensure monotonicity. Type I error is therefore smaller.

4.2 Testing Hypotheses 2 and 3

To test Hypotheses 2 and 3, another two-way MANOVA was conducted using measures of TTR and MCI as dependent variables. As before, this analysis employed two levels of Subject (*English I* and *English II*) and nine levels of Publisher (A, B, C, D, E, F, G, H, I) as independent variables.

The Kolmogorov-Smirnov test revealed *p*-values of .006 for the *English I* group and .031 for the *English II* group, suggesting potential deviations from normality. However, further examination of skewness and kurtosis indicated no substantial departure from normality; the *English I* group had skewness and kurtosis values of -.963 and .920, respectively, while the *English II* group had skewness of -.613 and kurtosis of .647. These values fell well within the acceptable range of |±2|, implying that the data's distribution did not pose a problem for conducting the MANOVA test. Therefore, it was deemed appropriate to proceed with the analysis, while ensuring that subsequent statistical interpretations were carefully scrutinized to account for any potential sensitivity in the data.

Box's test for the equality of covariance matrices revealed that Box's $M = 56.829$, $p > .608$, indicating no statistical significance. This result suggests that the variance-covariance matrices of the dependent variables are equal across all levels of the independent variables, supporting the use of Wilks' lambda for multivariate analysis. Meanwhile, Bartlett's test of sphericity yielded an approximate chi-square value of 16.410, $p < .001$, which was significant. This indicates that there is sufficient correlation among the dependent variables for conducting a multivariate test.

Using Wilks' lambda (see Table 10), the dependent variate was significantly affected by the main effects of Subject, Wilk's lambda = .923, $F(2, 87) = 3.639$, $p < .030$, partial $\eta^2 = .077$, and Publisher, Wilk's lambda = .207,

$F(16, 174) = 13.012$, $p < .001$, partial $\eta^2 = .545$. The multivariate interaction effect of Subject * Publisher was also statistically significant, Wilk's lambda = .697, $F(16, 174) = 2.154$, $p < .008$, partial $\eta^2 = .165$.

Table 10. Multivariate and Univariate Analyses of Variance for Measures of TTR and MCI

	Multivariate	Univariate	
Source	F [a]	TTR[b]	MCI[b]
Subject	3.639*	4.578*	5.612*
Publisher	13.012*	32.614*	4.833*
Subject * Publisher	2.154*	4.168*	.164
MSE		.001	.001

Note. Multivariate f-ratios were generated from Wilks' Lambda.
a. Multivariate *df* = 2, 87(subject), 16, 174(publisher), 16, 174(subject * publisher)
b. Univariate *df* = 1, 88(subject), 8, 88(publisher), 8, 88(subject * publisher)
*$p < .05$

The analysis showed that 79% of the variance in the data could be explained by the factors under study, with the remaining 21% attributable to unexplained variance or error. Levene's test confirmed no violation of the equality of error variances assumption for either dependent variable, allowing univariate effects for TTR and MCI to be assessed at the standard alpha level of .05. The Subject factor had a significant effect on both TTR ($F(1, 88) = 4.578$, $p < .035$, partial $\eta^2 = .049$) and MCI ($F(1, 88) = 5.612$, $p < .020$, partial $\eta^2 = .060$). Similarly, the Publisher factor significantly influenced both TTR ($F(8, 88) = 32.614$, $p < .001$, partial $\eta^2 = .748$) and MCI ($F(8, 88) = 4.833$, $p < .001$, partial $\eta^2 = .305$. Interaction effects between Subject and Publisher were significant for TTR ($F(8, 88) = 4.168$, $p < .001$, partial $\eta^2 = .275$), but not for MCI ($F(8, 88) = .164$, $p < .995$, partial $\eta^2 = .015$).

The significant interaction effect of Subject and Publisher on TTR, as illustrated in Figure 5 and detailed through simple effects analysis and the Ryan, Einot, Gabriel, Welch Studentized Range post hoc test, summarized in Table 11, provides compelling evidence for Hypothesis 2 and reveals a complex pattern of differences in TTR scores across various combinations. Table 11 highlights that *English II*-A ($M = .344$, $SE = .010$) had the lowest TTR scores, significantly differing from several other combinations, including *English I*-A (M = .398, SE = .010), *English II*-G (M = .434, SE = .010), and numerous higher-scoring groups such as *English II*-H (M = .532, SE = .010), which had the highest TTR score. The results also show tiered differences, with *English I*-A and *English II*-G scoring significantly lower than mid-tier groups like *English I*-B, *English I*-G, *English II*-B, *English II*-I, and *English II*-F, which in turn scored lower than higher-tier combinations such as *English I*-D, *English II*-C, and *English I*-I. Among the highest groups, *English II*-D, *English I*-H, *English I*-C, *English II*-E, *English I*-F, and *English I*-E scored significantly lower than *English II*-H, which consistently ranked the highest. These findings, accounting for approximately 28% of the variance in TTR, reveal the complexity of the interaction between Subject and Publisher, highlighting substantial variability in how these factors influence the repetition of words, with certain publishers or subjects amplifying TTR differences more significantly than others.

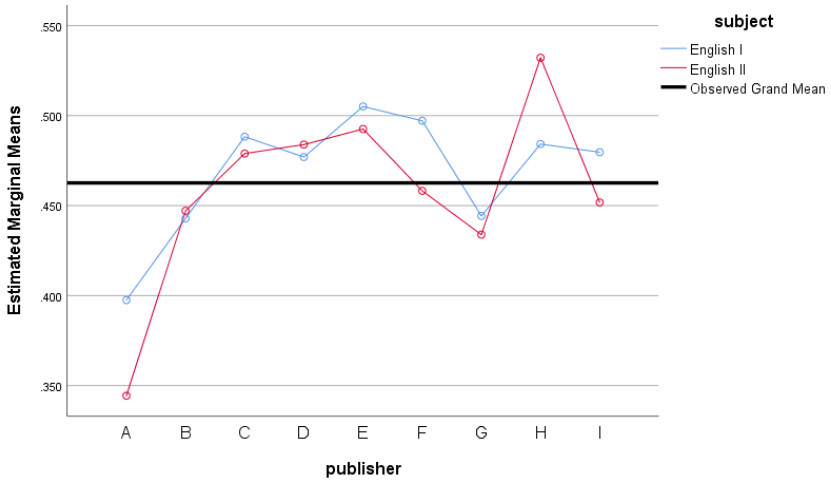

Figure 5. Estimated Marginal Means of TTR

Table 11. Homogeneous Subsets for All TTR Groups

Ryan-Einot-Gabriel-Welsch Range[a]									Subset for alpha = .05	
Group	N	1	2	3	4	5	6	7	8	9
II-A	6	.344								
I-A	6		.398							
II-G	6		.434	.434						
I-B	6			.443	.443					
I-G	6			.444	.444	.444				
II-B	6			.447	.447	.447				
II-I	6			.452	.452	.452	.452			
II-F	6			.458	.458	.458	.458	.458		
I-D	6				.477	.477	.477	.477	.477	
II-C	6				.479	.479	.479	.479	.479	
I-I	6				.480	.480	.480	.480	.480	
II-D	6					.484	.484	.484	.484	
I-H	6					.484	.484	.484	.484	
I-C	6						.488	.488	.488	
II-E	5						.493	.493	.493	.493
I-F	6							.497	.497	.497
I-E	5								.505	.505
II-H	6									.532
Sig.		1.000	.112	.900	.369	.252	.366	.294	.903	.257

Means for groups in homogeneous subsets are displayed.
a. Critical values are not monotonic for these data. Substitutions have been made to ensure monotonicity. Type I error is therefore smaller.

Although a significant interaction effect typically overshadows main effects, the main effect of Publisher on TTR warrants detailed examination due to its substantial contribution, accounting for approximately 75% of the variance in TTR—far exceeding the 4.9% explained by the main effect of Subject, which showed equivalent mean TTR scores for *English I* (M = .468, SE = .003, 95% CI = [.462, .475]) and *English II* (M = .458, SE = .003, 95% CI = [.451, .465]) textbooks. The Ryan, Einot, Gabriel, Welch Studentized Range post hoc test, summarized in

Table 12, revealed that the main effect of Publisher was primarily driven by Publisher H (*M* = .508, *SE* = .007), which had significantly higher TTR scores compared to all other publishers but C and E, that is, Publisher A (*M* = .371, *SE* = .007), G (*M* = .439, *SE* = .007), B (*M* = .445, *SE* = .007), I (*M* = .466, *SE* = .007), F (*M* = .478, *SE* = .007), and D (*M* = .480, *SE* = .007). Conversely, Publisher A had the lowest TTR scores, significantly lower than Publishers G, B, I, F, D, C (*M* = .484, *SE* = .007), E (*M* = .499, *SE* = .008), and H. Additionally, Publishers G and B scored lower than I, F, D, C, E, and H, while Publisher I had significantly lower scores than F, D, C, E, and H. Among the mid-range publishers, F, D, and C had significantly lower TTR scores than E and H. These findings underline the dominant role of Publisher in shaping TTR variations, with Publisher H markedly outperforming others and Publisher A lagging significantly behind.

Table 12. Homogeneous Subsets for Publisher Groups of TTR

Ryan-Einot-Gabriel-Welsch Range[a,b]

		Subset					
Publisher	N	1	2	3	4	5	6
A	12	.371					
G	12		.439				
B	12		.445	.445			
I	12			.466	.466		
F	12				.478	.478	
D	12				.480	.480	
C	12				.484	.484	.484
E	10					.499	.499
H	12						.508
Sig.		1.000	.975	.185	.550	.452	.130

Means for groups in homogeneous subsets are displayed.
Based on observed means.
The error term is Mean Square(Error) = .001.

a. Critical values are not monotonic for these data. Substitutions have been made to ensure monotonicity. Type I error is therefore smaller.
b. Alpha = .05

Although no interaction effect was observed for MCI, the significant main effects of Subject and Publisher provide sufficient evidence to support Hypothesis 3. MCI scores were significantly higher in *English I* textbooks (*M* = .835, *SE* = .003, 95% CI = [.830, .841]) than in *English II* textbooks (*M* = .826, *SE* = .003, 95% CI = [.820, .831]), though this accounted for only 6% of the variance in MCI. In contrast, the main effect of Publisher accounted for a substantially larger proportion of variance at 30.5%, highlighting its dominant influence. The Ryan, Einot, Gabriel, Welch Studentized Range post hoc test, as illustrated in Figure 6 and detailed in Table 13, showed that Publisher A (*M* = .799, *SE* = .139) had significantly lower MCI values than all other publishers, including G (*M* = .826, *SE* = .006), H (*M* = .827, *SE* = .006), B (*M* = .829, *SE* = .006), E (*M* = .830, *SE* = .007), F (*M* = .839, *SE* = .006), C (*M* = .840, *SE* = .006), I (*M* = .842, *SE* = .006), and D (*M* = .843, *SE* = .006). These results underscore the critical role of Publisher in determining MCI levels, with Publisher A notably underperforming, while the differences between Subject categories also contribute meaningfully to the variability in MCI measures.

4.3 Testing Hypothesis 5

To test Hypothesis 5, a two-way ANOVA was conducted using measures of DP/C as a dependent variable. This analysis employed two levels of Subject (*English I* and *English II*) and nine levels of Publisher (A, B, C, D, E, F, G, H, I) as independent variables.

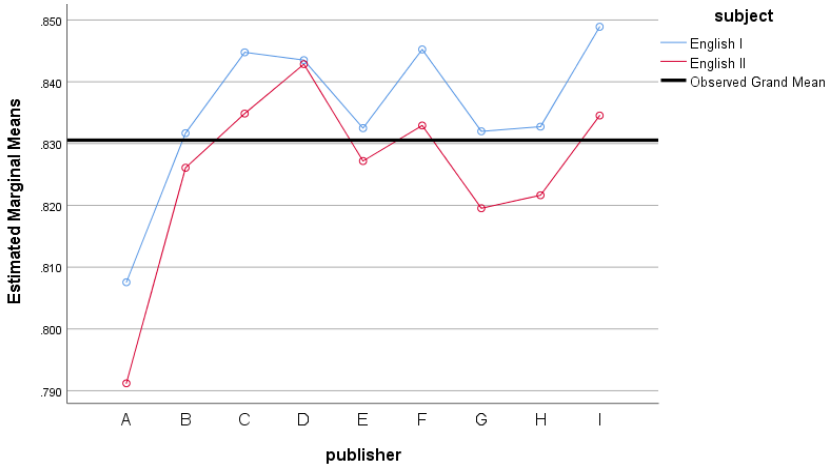

Figure 6. Estimated Marginal Means of MCI

Table 13. Homogeneous Subsets for Publisher Groups of MCI

Ryan-Einot-Gabriel-Welsch Range[a,b]

		Subset	
Publisher	N	1	2
A	12	.799	
G	12		.826
H	12		.827
B	12		.829
E	10		.830
F	12		.839
C	12		.840
I	12		.842
D	12		.843
Sig.		1.000	.481

Means for groups in homogeneous subsets are displayed.
Based on observed means.
The error term is Mean Square(Error) = .001.

a. Critical values are not monotonic for these data. Substitutions have been made to ensure monotonicity. Type I error is therefore smaller.
b. Alpha = .05

No outliers were detected in the data, and all Subject or Publisher groups but one passed the normality tests. For sample sizes below 50, the Shapiro-Wilks test was used, while for sample sizes of 50 or more, the Kolmogorov-Smirnov test was applied. In the Shapiro-Wilks test, the D group of DC/C had a *p*-value of .042 However, skewness (-.865) and kurtosis (-.814) were within the range of |±2|, indicating no issue with performing the ANOVA test. Thus, it was deemed appropriate to proceed with the ANOVA test while paying extra attention to subsequent statistical interpretations.

Although no interaction effect was found for DC/C, the significant main effects of Subject and Publisher provide strong support for Hypothesis 5. DC/C scores were significantly higher in *English II* textbooks (*M* = .257, *SE* = .007, 95% CI = [.244, .270]) compared to *English I* textbooks (*M* = .231, *SE* = .007, 95% CI = [.218, .245]), although this difference accounted for only 7.5% of the variance, indicating a relatively modest impact of Subject. In contrast, the main effect of Publisher accounted for a much larger proportion of the variance at 30.1%,

underscoring its dominant influence on DC/C scores. The Ryan, Einot, Gabriel, Welch Studentized Range post hoc test, as illustrated in Figure 7 and detailed in Table 14, revealed that Publisher A (M = .295, SE = .014) had significantly higher DC/C scores than several other publishers, including C (M = .203, SE = .014), I (M = .209, SE = .014), B (M = .227, SE = .014), and F (M = .230, SE = .014). Conversely, Publisher C, which had the lowest DC/C scores, and I were significantly lower than Publisher E (M = .499, SE = .008), further emphasizing the variability in DC/C levels across publishers. These results highlight the critical role of Publisher in shaping DC/C values, with Publishers A and E standing out as having notably higher scores, while Subject-related differences also contribute meaningfully to the overall variability in dependent clause complexity.

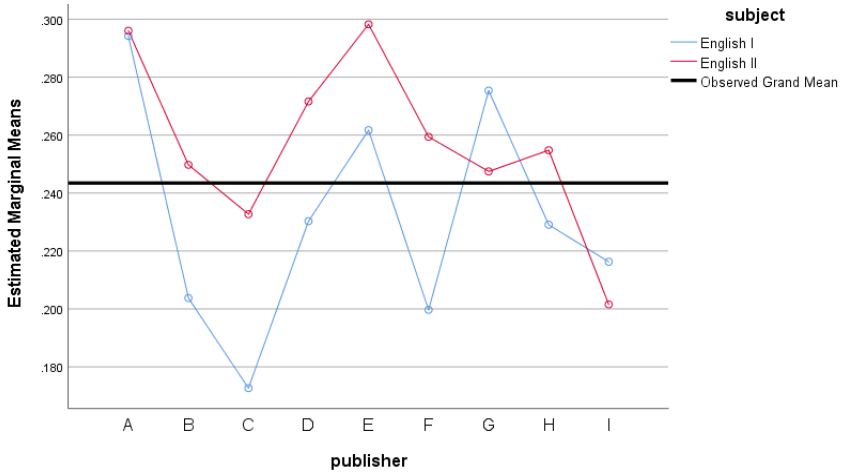

Figure 7. Estimated Marginal Means of DC/C

Table 14. Homogeneous Subsets for Publisher Groups of DC/C

Ryan-Einot-Gabriel-Welsch Range[a,b]

		Subset		
Publisher	N	1	2	3
C	12	.203		
I	12	.209		
B	12	.227	.227	
F	12	.230	.230	
H	12	.242	.242	.242
D	12	.251	.251	.251
G	12	.261	.261	.261
E	10		.280	.280
A	12			.295
Sig.		.075	.220	.115

Means for groups in homogeneous subsets are displayed.
Based on observed means.
The error term is Mean Square(Error) = .002.

a. Critical values are not monotonic for these data. Substitutions have been made to ensure monotonicity. Type I error is therefore smaller.
b. Alpha = .05

4.4 Discussion

To address Research Questions 1–3, which explores whether the language complexity of the listening scripts in *English I* and *II* textbooks varies by subjects and publishers, two two-way MANOVAs and one two-way ANOVA were performed to test Hypotheses 1–6, which are related to those research questions. The results of the (M)ANOVAs with Subject and Publisher as independent variables and the indices of language complexity of AG, TTR, MCI, MLC, DC/C, and CP/C as dependent variables are summarized in Table 15.

Table 15. Summary of the Results of (M)ANOVAs

Research Question	Hypothesis	Type of Test	DV	Significant Effect (Size %)
Q1 (Lexical)	H1	MANOVA1	AG	Subject (25.1), Publisher (45.2)
	H2	MANOVA2	TTR	Subject (4.9), Publisher (74.8), Subject * Publisher (27.5)
Q2 (Morphological)	H3	MANOVA2	MCI	Subject (6.0), Publisher (30.5)
Q3 (Syntactic)	H4	MANOVA1	MLC	Subject (10.3), Publisher (18.3), Subject * Publisher (19.3)
	H5	ANOVA	DC/C	Subject (7.5), Publisher (30.1)
	H6	MANOVA1	CP/C	Publisher (34.2), Subject * Publisher (19.8)

The findings summarized above confirm that all the alternative hypotheses were supported, indicating significant variations in the lexical, morphological, and syntactic complexity of the listening scripts in the *English I* and *II* textbooks based on both Subject (*English I* vs. *English II*) and Publisher. This result highlights the substantial role these factors play in shaping the linguistic characteristics of the textbooks. A detailed examination of the effect sizes reveals complex relationships among these variables. When both Subject and Publisher were significant, Publisher exerted a stronger influence on the complexity of the texts than Subject. Similarly, in cases where both Publisher and the interaction between Subject and Publisher were significant, Publisher had a more pronounced effect on complexity than the interaction effect for TTR and CP/C. However, for MLC, the influence of Publisher was slightly less significant than that of the interaction effect. These results highlight the variability introduced by publishers in shaping the complexity of textbook listening scripts, while also pointing to the interactive dynamics between subject matter and publication practices that contribute to variability in textbook content.

The estimated marginal means plot of AG in Figure 2 shows that the AG values for *English II* textbooks were higher than those for *English I* textbooks across all publishers, except Publisher C. Notably, the difference between *English I* and *II* was smaller for Publisher E than for other publishers. This indicates that the disparity in advanced vocabulary content between *English I* and *II* textbooks is less pronounced for Publishers C and E, potentially limiting their effectiveness in fostering incremental learning of advanced vocabulary.

The Bonferroni post hoc test results in Table 6 reveal that the AG values of textbooks from Publishers D, A, and G were significantly higher than those from Publishers I, C, H, and F. Furthermore, significant differences were noted between textbooks from Publisher E and those from Publishers H and F. While there is no universally accepted standard for the ideal proportion of advanced vocabulary, the existence of such pronounced disparities across publishers is less than desirable. Textbooks with excessively high proportions of advanced vocabulary may overwhelm learners, whereas those with overly low proportions risk failing to offer adequate vocabulary development opportunities. Consequently, reducing AG disparities among publishers would be beneficial to ensure a more balanced and effective learning experience.

A more pressing concern emerges from the observation that the AG values of *English II* textbooks from Publishers I, C, H, and F were lower than those of *English I* textbooks from Publishers D, E, G, and A. This

discrepancy raises important questions about the learning continuity for students who, due to factors such as school transfers, transition from *English I* textbooks published by D, E, G, or A to *English II* textbooks from I, C, H, or F. Such a shift could hinder students' ability to effectively acquire advanced vocabulary, potentially disrupting their progression in language learning.

The analysis of TTR emphasizes the considerable influence of Publisher, which accounts for a predominant 74.8% of the variance, while the interaction effect contributes a moderately strong 27.5%. This dynamic interplay between the two factors creates a complex pattern of variability in how TTR is affected, indicating that the repetition of words is not uniformly influenced across publishers or subjects. Instead, certain publishers or subject areas appear to magnify TTR differences more significantly than others, suggesting that specific editorial or curricular strategies may play a critical role. Figure 5 visually illustrates these disparities, highlighting the intricate relationship between publisher practices, subject content, and their combined impact on lexical repetition, ultimately shaping the educational efficacy of the textbooks.

For Publishers A, C, E, F, G, and I, the TTR values for *English II* textbooks were lower than those for their corresponding *English I* textbooks, which aligns with the observed significant yet weak main effect of Subject. This suggests a greater repetition of vocabulary in *English II* textbooks by these publishers, potentially aiding incremental vocabulary learning. However, the opposite trend was observed for Publishers B, D, and H, where the TTR values for *English II* textbooks were higher than for *English I* textbooks. This indicates a smaller proportion of repeated words in these *English II* textbooks, despite their more advanced vocabulary and longer texts. While this could signal an attempt to introduce a broader lexical range, it may inadvertently hinder students' gradual acquisition of advanced vocabulary, as reduced repetition limits opportunities for reinforcement and deeper learning of new linguistic forms.

The importance of publisher-specific approaches in determining the educational value of English textbooks is highlighted more clearly, when considering the significant main effect of Publisher and its interaction effects. The Ryan, Einot, Gabriel, Welch Studentized Range post hoc test results in Table 12 provide further insight, revealing that textbooks from Publishers A, G, and B exhibit substantially lower TTR values compared to those from Publishers C, E, and H. While there is no universally accepted benchmark for the ideal TTR in these textbooks, lower TTR values, associated with more frequent repetition of words, are generally preferable for fostering vocabulary acquisition through natural exposure. Consequently, the higher TTR values in textbooks by Publishers C, E, and H suggest a potential limitation in their ability to support students' effective vocabulary learning, emphasizing the need for careful evaluation of lexical diversity in educational materials.

The analysis of MCI highlights the significant role of Publisher, which explains 30.5% of the variance, while revealing a relatively straightforward pattern of variability in how MCI is influenced. As depicted in Figure 6, the use of inflected or derived word forms is largely consistent across subjects and publishers, with all publishers showing lower MCI values in their *English II* textbooks compared to their *English I* textbooks. This aligns with the observed significant but weak main effect of Subject (6.0%), suggesting that *English II* textbooks generally contain a higher proportion of inflected or derived word forms than *English I* textbooks. However, the disparity between the two levels is notably smaller for Publisher D, indicating a less pronounced shift in MCI values. This reduced difference may limit the ability of Publisher D's textbooks to effectively support the progressive learning of inflected or derived word forms, potentially impacting students' linguistic development compared to textbooks from other publishers.

The Ryan, Einot, Gabriel, Welch Studentized Range post hoc test results, presented in Table 13, reveal that the textbooks from Publisher A exhibit significantly lower MCI values compared to those from the other publishers. This suggests that Publisher A's textbooks include a higher proportion of inflected or derived forms of headwords,

reflecting a distinct approach to lexical presentation. Importantly, the relatively low MCI values indicate that while inflected or derived forms are emphasized, they are not excessively overrepresented, maintaining a balanced distribution. This balance ensures that the inclusion of these forms does not overwhelm students, allowing for effective incremental learning without hindering comprehension or vocabulary acquisition. As such, Publisher A's approach appears to support the development of linguistic competence without introducing unnecessary complexity.

The analysis of MLC shows that while the main effects of Subject (10.3% of variance) and Publisher (18.3% of variance) are significant, the interaction effect between Subject and Publisher (19.3% of variance) requires deeper interpretation. Generally, the higher MLC values in *English II* textbooks compared to *English I* textbooks reflect an expected progression in linguistic complexity in line with curricular advancements. However, the interaction effect reveals that this progression is not consistent across publishers, highlighting variability in how clausal length is structured at different levels. The Tamhane T2 post hoc test results, as shown in Table 8, emphasize these discrepancies. Specifically, the *English II* textbooks from Publisher G showed significantly higher MLC values compared to the *English II* textbooks of Publishers A and D, as well as the *English I* textbooks of Publishers C and E. These findings suggest inconsistencies in publishers' approaches to developing linguistic complexity, which may have implications for the uniformity and effectiveness of curriculum delivery.

The estimated marginal means plot in Figure 3 also reveals notable inconsistencies in the progression of syntactic complexity across publishers. For instance, the MLC values for *English I* textbooks from Publishers C and D were higher than those of their corresponding *English II* textbooks, contradicting the expected increase in complexity. Publishers A and F showed minimal differences between the average MLC values of their *English I* and *II* textbooks, while Publishers G and H displayed excessively large differences. As previously noted in the discussion on AG, the MLC values for the *English I* textbooks from Publishers D and G were significantly higher than those for the *English II* textbooks from Publishers C, E, and F. These irregularities, including reversals and insufficient or excessive gaps in syntactic complexity, challenge the anticipated progression from simpler to more complex linguistic structures as students advance through the curriculum.

The Ryan, Einot, Gabriel, Welch Studentized Range post hoc test results, outlined in Table 9, indicate that the textbooks from Publisher G have significantly higher MLC values compared to those from Publishers C and F. This suggests that Publisher G's textbooks incorporate a greater number of lengthy clauses, which may effectively challenge advanced learners and support their language development. In contrast, the lower MLC values in the textbooks by Publishers C and F imply a lack of sufficient lengthy clauses, potentially limiting their capacity to meet the diverse linguistic needs of learners at higher proficiency levels.

The analysis of DC/C accentuates the significant influence of Publisher, accounting for 30.1% of the variance, and reveals a complex variability in how dependent clauses per clause are distributed. As depicted in Figure 7, the ratio of dependent clauses is largely consistent across subjects and publishers, with most publishers—except G and I—showing higher DC/C values in their *English II* textbooks compared to their *English I* counterparts. This trend aligns with the observed significant yet weak main effect of Subject (7.5%), indicating a general increase in dependent clause usage at higher levels. However, Publishers G and I deviate from this pattern, with lower DC/C values in their *English II* textbooks than in their *English I* textbooks. This suggests a reduced proportion of dependent clauses despite the inclusion of more advanced vocabulary and longer texts, potentially hindering the gradual acquisition of complex clausal structures. Additionally, the smaller disparity in DC/C values between *English I* and *II* textbooks for Publisher A suggests a less pronounced progression in syntactic complexity, which may limit their effectiveness in supporting the incremental learning of complex grammar. These inconsistencies

across publishers highlight the need for a more deliberate and structured approach to integrating dependent clauses to enhance students' linguistic development.

A particularly troubling issue emerges from the observation that the DC/C values of *English II* textbooks from Publishers C and I were lower than those of *English I* textbooks from Publishers A, E, and G. This discrepancy highlights a potential disruption in learning continuity for students transitioning between publishers, such as those who, due to school transfers or curriculum changes, move from using *English I* textbooks by Publishers A, E, or G to *English II* textbooks by Publishers C or I. Such transitions could create a mismatch in syntactic complexity, with students encountering a decrease in the proportion of dependent clauses in their *English II* textbooks, despite these materials being intended for a more advanced level. This inconsistency may hinder students' ability to develop a more sophisticated understanding of complex grammatical structures, potentially impeding their overall language learning progression and undermining the educational goals of the curriculum.

The Ryan, Einot, Gabriel, Welch Studentized Range post hoc test results in Table 14 highlight significant variability in the DC/C values across publishers, with Publisher A's textbooks showing the highest values, significantly exceeding those of Publishers C, I, B, and F. This suggests that while Publisher A emphasizes structures with more dependent clauses, the excessive proportion may overwhelm students, potentially hindering their ability to progressively build and internalize complex grammatical structures. On the other hand, the textbooks from Publisher C, which exhibit the lowest DC/C values, along with those from Publisher I, have significantly fewer dependent clauses than those from Publisher E. This lack of sufficient exposure to dependent clauses in Publishers C and I may fail to adequately challenge students or provide opportunities to practice and master more advanced syntactic structures, potentially limiting their linguistic development.

The analysis of CP/C reveals substantial variability in the use of coordinate phrases and clauses across publishers and subjects, with the main effect of Publisher explaining 34.2% of the variance and the Subject * Publisher interaction contributing an additional 19.8%. These effects highlight the complexity of interpreting CP/C trends, depicted in Figure 4 and detailed in Table 9. For instance, the *English II* textbook from Publisher G shows significantly higher CP/C values compared to *English II* textbooks from Publishers A, C, and F, as well as *English I* textbooks from Publishers A, D, H, and I. Conversely, the *English II* textbook from Publisher C has notably lower CP/C values than *English I* textbooks from Publishers B and G. This variability undermines the simplicity of analyzing main effects alone, as evidenced by the significantly higher combined CP/C values for Publisher G's *English I* and *II* textbooks compared to those from Publishers C, F, A, H, and D, and the significantly lower combined values for Publisher C compared to Publishers B and G. These findings have meaningful implications for language learning: students using Publisher G's textbooks are likely exposed to an overabundance of coordinate clauses, which may hinder syntactic variety and complexity, whereas students relying on Publisher C's textbooks may encounter insufficient opportunities to engage with coordinate phrases.

The observed CP/C values highlight notable inconsistencies in syntactic complexity progression across subject levels and publishers. The *English I* textbooks from Publishers B, C, E, and F demonstrate higher CP/C values than their corresponding *English II* textbooks, which runs counter to the expected trend of increasing syntactic complexity in advanced levels. Additionally, the CP/C values of the *English I* textbooks from Publishers B, E, and G exceed those of *English II* textbooks from Publishers C, D, F, and H, further emphasizing irregularities. The lack of a discernible difference in CP/C values between the *English I* and *II* textbooks from Publisher A raises concerns about the absence of progression, suggesting a failure to scaffold learning effectively. Conversely, the excessively large CP/C differences between the *English I* and *II* textbooks from Publishers C and G suggest abrupt shifts in complexity, which could hinder smooth linguistic development. These findings call for a more balanced

and consistent approach to integrating coordinate phrases and clauses across subject levels to ensure optimal learning outcomes.

5. Conclusion

Building on the statistical findings discussed in Section 4.4, the evaluation of *English I* and *II* textbooks by the nine publishers under study can now be framed in terms of Tolerance and Continuity, two critical dimensions for assessing a textbook's educational efficacy. Tolerance refers to the extent to which textbooks balance complexity and accessibility, ensuring that students are neither overwhelmed nor under-challenged. Continuity assesses the consistency and progression in linguistic features, such as vocabulary development and syntactic complexity, between *English I* and *II* levels. Table 16 synthesizes the results of this evaluation, providing a comparative overview of how well each set of textbooks meets these criteria. By integrating these dimensions, the analysis highlights strengths and weaknesses in the design of individual publishers' materials, offering practical insights for enhancing the pedagogical value and alignment of textbooks across grade levels. This evaluation underscores the importance of a carefully calibrated approach to curriculum design that supports incremental learning and smooth transitions for students.

Table 16. Evaluation of the *English I* and *II* Textbooks Based on Tolerance and Continuity

Index	Publisher	A	B	C	D	E	F	G	H	I
AG	Tolerance	⊥		√	⊥	⊥	√	⊥	√	√
	Continuity			√		√				
TTR	Tolerance	⊥	⊥	√		√		⊥	√	
	Continuity		√		√				√	
MCI	Tolerance	⊥								
	Continuity				√					
MLC	Tolerance			√			√	√		
	Continuity	√		√	√		√	√	√	
DC/C	Tolerance	√		√						√
	Continuity	√						√		√
CP/C	Tolerance			√			√	√		
	Continuity	√	√	√		√	√	⊥		

⊥ indicates acceptable index values and √, unacceptable values.

Table 16 reveals significant variations in publishers' adherence to the Principles of Tolerance and Continuity across multiple linguistic measures. Publishers D, E, and G exhibit higher AG values, indicating a preference for advanced vocabulary, whereas Publishers C, F, H, and I fall short, violating Tolerance.[1] For TTR, lower values are preferred, placing Publishers C, E, and G in violation of this principle. Similarly, for MCI, lower values are favorable, enabling Publisher A to meet Tolerance. The average MLC across publishers is 6.827, but Publishers C (6.494), F (6.627), and G (7.327) exceed either acceptable lower or upper limits, violating Tolerance. Minimal

[1] A reviewer questioned the relevance of Tolerance as a principle for textbook selection or evaluation, suggesting that textbooks approved through the screening process could already achieve maximum diversity. However, this assumes the screening process is flawless, which, as this paper demonstrates, is not always the case, particularly when evaluating language complexity. While textbooks should strive for maximum diversity, they must also adhere to an acceptable range of deviation in language complexity, given their role as learning materials for public education.

MLC differences between *English I* and *II* textbooks from Publishers A and F, and excessively large differences from Publishers G and H, breach Continuity. For DC/C values, Publishers A, C, and I demonstrate significant inconsistencies, violating Tolerance, while Publishers G and I, showing reversals in DC/C trends, also violate Continuity, with Publisher A's minimal inter-level difference adding to this concern. For CP/C, significant differences place Publishers C, F, and G in violation of Tolerance, while Continuity is breached by Publishers B, C, E, and F. Although Publisher A's lack of distinction between *English I* and *II* textbook averages is problematic, Publisher G's larger yet acceptable difference (.049) demonstrates an example of a tolerable deviation in Continuity.

Summarizing the findings for each textbook, Publisher A has one Tolerance violation and three Continuity violations, Publisher B has two Continuity violations, Publisher C has five Tolerance violations and three Continuity violations, Publishers D has three Continuity violations, Publisher E has one Tolerance violation and two Continuity violations, Publisher F has three Tolerance violations and two Continuity violations, Publisher G has two Tolerance violations and two Continuity violations, Publisher H has two Tolerance violations and two Continuity violations, and Publisher I has two Tolerance violations and one Continuity violation.

Meeting both Tolerance and Continuity for all six language complexity indices is practically impossible. Therefore, it is reasonable to categorize textbooks based on the number of violations: Textbooks with less than three violations can be considered excellent, those with three or four violations are deemed acceptable, those with more than four violations require improvement. Based on this criterion, textbooks from Publisher B are classified as excellent, textbooks from A, D, E, G, H and I as acceptable, and textbooks from Publisher C and F as in need of improvement.

The evaluation of the listening scripts in the *English I* and *II* textbooks, focusing on language complexity and adherence to the Principles of Tolerance and Continuity, emphasizes the efforts of textbook developers to meticulously control key language complexity indices, including AG, TTR, MLC, and DP/C. This careful calibration ensures that the linguistic demands of the scripts align with the progressive development of learners' language skills as prescribed by the national curriculum. By structuring linguistic features to increase gradually and systematically, developers should aim to provide learners with comprehensible yet sufficiently challenging input that facilitates step-by-step language acquisition. Such attention to detail reflects an intention to balance accessibility and complexity, enabling learners to develop their language proficiency in a manner consistent with curricular expectations while addressing diverse educational needs.

Starting from the 2025 academic year, new English textbooks, created based on the 2022 Revised Curriculum and approved through the screening process, will be used for grades 3–4 in elementary school, grade 1 in middle school, and all grades in high school. At this critical juncture, as the adoption of these new textbooks in schools approaches, analyzing the English textbooks based on the 2015 Revised Curriculum through indices of language complexity in relation to the principles of Tolerance and Continuity serves as a preparatory step for evaluating the new textbooks in the right direction. While the listening scripts in the newly approved textbooks may differ in detail, they are expected to yield results largely consistent with those of this study. If this expectation is confirmed, it may indicate the need to explore new methodologies for textbook screening.

References

강준수·차경환(Kang, J-S. and K-W Cha). 2023. 2015 개정 교육과정 영어I 교과서 듣기활동의 어휘적·구문적 특성: 코퍼스 기반 분석(Lexical and syntax features of listening sections in English 1

textbook: Corpus-based analysis.). ≪학습자중심교과교육연구≫(*Journal of Learner-Centered Curriculum and Instruction*) 23(11), 119-132.

교육부(Ministry of Education). 2016. 『2015 개정 교육과정에 따른 교과용도서 개발을 위한 편찬상의 유의점 및 검정기준』(*Considerations and Selection Criteria for the Publication of Textbooks Based on the 2015 Revised National Curriculum*). 교육부(Ministry of Education).

교육부(Ministry of Education). 2022. 『영어과 교육과정』(*English Curriculum*), 교육부 고시 제2022-33호 [별책 14]. 교육부(Ministry of Education).

권혁승·박은성·김현정·오건석·강신자·윤소영(Kwon, H., E. Park, H. Kim, G. Oh, S. Kang and S. Yoon). 2018. *High School English I*. 동아출판(Dong-A Publishing Co., Ltd.).

권혁승·박은성·김현정·오건석·강신자·윤소영(Kwon, H., E. Park, H. Kim, G. Oh, S. Kang and S. Yoon). 2019. *High School English II*. 동아출판(Dong-A Publishing Co., Ltd.).

김길중·P. M. Aaron·임정원·장진화·김건우·김나현 외(K. Kim, P. M. Aaron, J. Lim, J. Jang, K. Kim, N. Kim et al.). 2018. *High School English I*. 다락원(Darakwon).

김길중·P. M. Aaron·임정원·장진화·김건우·김나현 외(K. Kim, P. M. Aaron, J. Lim, J. Jang, K. Kim, N. Kim et al.). 2019. *High School English II*. 다락원(Darakwon).

김성곤·윤진호·구은영·전형주·서정환·이후고 외(Kim, S., J. Yoon, E. Goo, H. Jeon, J. Seo, H. Lee et al.). 2018. *High School English I*. 엔이 능률(NE Neungyule).

김성곤·윤진호·구은영·전형주·서정환·이후고 외(Kim, S., J. Yoon, E. Goo, H. Jeon, J. Seo, H. Lee et al.). 2019. *High School English II*. 엔이 능률(NE Neungyule).

류지수·전문기(Ryu, J. and M. Jeon). 2020. Coh-Metrix 시스템을 이용한 고등학교 영어교과서의 듣기대본 복잡도 분석(An analysis of the text complexity of listening scripts of high school English textbooks using Coh-Metrix). ≪영어학≫(*Korean Journal of English Language and Linguistics*) 20, 363-383.

민나리(Min, N.). 2023. 2015 개정 영어과 교육과정에 기반한 중학교 영어 교과서 듣기 대본 분석: 통사적 복잡성을 중심으로(*Analysis of the Listening Scripts of Middle School English Textbooks Based on the 2015 Revised National Curriculum of English: With Focus on Syntactic Complexity*). 석사학위논문(Master's thesis), 인하대학교(Inha University).

민찬규·정현성·이상기·김윤규·나우철·안효선 외(Min, C., H. Jung, S. Lee, Y. Kim, W. Na, H. Ahn et al.). 2018. *High School English I*. 지학사(Jihak Publishing Co., Ltd.).

민찬규·정현성·이상기·김윤규·나우철·안효선 외(Min, C., H. Jung, S. Lee, Y. Kim, W. Na, H. Ahn et al.). 2019. *High School English II*. 지학사(Jihak Publishing Co., Ltd.).

박준언·윤병우·김선형·최선영·최지영·김진수 외(Park, J., B. Yoon, S. Kim, S. Choi, J. Choi, J. Kim et al.). 2018. *High School English I*. 와이비엠(YBM).

박준언·윤병우·김선형·최선영·최지영·김진수 외(Park, J., B. Yoon, S. Kim, S. Choi, J. Choi, J. Kim et al.). 2019. *High School English II*. 와이비엠(YBM).

송해성(Sohng, H. S.). 2013. 고등학교 영어 교과서 듣기 과업의 학년별 연계성 연구(Exploring the connectedness of listening tasks according to grades in High School English textbooks). ≪영어어문교육≫(*English Language & Literature Teaching*) 19(4), 279-305.

이소영·이현우(Lee, S. and H. Lee). 2019. 2015 개정 교육과정에 따른 중2 영어 교과서 어휘 분포 분석(An analysis of the distribution of vocabulary in Middle School English 2 textbooks based on 2015 Revised English National Curriculum). ≪중등영어교육≫(*Secondary English Education*) 12(3), 71-93.

이재영·문안나·오영일·조수경·이윤정·김정현 외(Lee, J. A. Moon, Y. Oh, S. Cho, Y. Lee, J. Kim et al.). 2018. *High School English I*. 천재교육(Chunjae Education, Inc.).

이재영·문안나·오영일·조수경·이윤정·김정현 외(Lee, J. A. Moon, Y. Oh, S. Cho, Y. Lee, J. Kim et al.). 2019. *High School English II*. 천재교육(Chunjae Education, Inc.).

이현우(Lee, H.). 2020. 2015 개정 교육과정에 기반한 영어 I 교과서의 어형론적 복잡성(Morphological complexity in English I textbooks based on the 2015 revised national curriculum). ≪중등영어교육≫(*Secondary English Education*) 13(1), 3-24.

이현우(Lee, H.). 2021. 2015 개정 교육과정에 기초한 「영어」 교과서와 고2 영어 국가수준학업성취도평가에 나타난 등급화된 단어의 분포(Distribution of graded vocabulary in the High School English textbooks and K-11 NAEA English test based on the 2015 Revised National Curriculum). ≪중등영어교육≫(*Secondary English Education*) 14(1), 3-24.

이현우·이소영(Lee, H. and S. Lee). 2024. 2015 개정 교육과정에 기반한 영어 I, II 교과서의 읽기 본문의 언어 복잡성(Language complexity of the reading sections of the High School English I and II textbooks based on the 2015 Revised National Curriculum). ≪영어학≫(*Korean Journal of English Language and Linguistics*) 24, 1,197-1,228..

임인재·전문기(Lim, I. and M. Jeon). 2013. 코퍼스에 기반한 중학교 1학년 개정 영어교과서 듣기자료 비교 분석(A corpus-based analysis of the listening materials in Middle School English 1 textbooks). ≪영어학≫(*Korean Journal of English Language and Linguistics*) 13(1), 105-134.

최인철·박리리·김근영·최수하·김주혜·손지해 외(Choi, I., L. Park, G. Kim, S. Choi, J. Kim, J. Son et al.). 2018. *High School English I*. 금성출판사(Kumsung Publishing Co., Ltd.).

최인철·서원화·이윤경·김주혜·전예지·라문선 외(Choi, I., W. Seo, Y. Lee, J. Kim, Y. Jeon, M. Rha et al.). 2019. *High School English II*. 금성출판사(Kumsung Publishing Co., Ltd.).

한상호·정은귀·김예리·김재란·이보희·이혜은(Han, S-H., E-G. Jung, Y. Kim, J. Kim, B. Lee, H. Lee). 2018. *High School English I*. 와이비엠 텍스트(YBM Text).

한상호·정은귀·김예리·김재란·이보희·이혜은(Han, S., E. Jung, Y. Kim, J. Kim, B. Lee, H. Lee). 2019. *High School English II*. 와이비엠 텍스트(YBM Text).

홍민표·안현기·박연미·김정태·장현옥·신정섭 외(Hong, M., H. Ahn, Y. Park, J. Kim, H. Jang, J. Shin et al.). 2018. *High School English I*. 비상교육(Visang Education, Inc.).

홍민표·안현기·R. Pak·O. D. Desmond·조금희(Hong, M., H. Ahn, R. Pak, O. D. Desmond, G. Cho). 2019. *High School English II*. 비상교육(Visang Education, Inc.).

황은경·박광현(Hwang, E. and K. Park). 2024. 초등학교 5·6학년 영어 교과서의 통사적 복잡성 비교 분석(A comparative analysis of syntactic complexity in 5th and 6th grade English textbooks). ≪현대영어교육≫(*Modern English Education*) 25, 100-114.

Bauer, L. and I. S. P. Nation. 1993. Word families. *International Journal of Lexicography* 6(4), 253-279.

Biber, D. 1988. *Variation across speech and writing*. Cambridge University Press.

Brezina, V. and G. Pallotti. 2019. Morphological complexity in written L2 texts. *Second Language Research* 35(1), 99-119.

Chafe, W. and D. Tannen. 1987. The relation between written and spoken language. *Annual Review of Anthropology* 16, 383-407.

Chan, Y. H. 2003. Biostatistics 104: Correlational analysis. *Singapore Medical Journal* 44(12), 614-619.

Daller, H., R. van Hout and J. Treffers-Daller. 2003. Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics* 24(2), 197-222.

Dancey, C. P. and J. Reidy. 2020. *Statistics without Maths for Psychology*, (8th ed.). Pearson.

De Vaus, D. 2002. *Analyzing Social Science Data: 50 Key Problems in Data Analysis*. Sage Publications, Ltd.

Graves, K. 2000. Designing Language Courses: A Guide for Teachers. Heinle & Heinle.

Guiraud, P. 1954. *Les Caractères Statistiques du Vocabulaire: Essai de Méthodologie*. Presses Universitaires de France.

Hunt, K. W. 1965. *Grammatical Structures Written at Three Grade Levels* (NCTE Research Report No. 3). National Council of Teachers of English.
Krashen, S. D. 1985. *The Input Hypothesis: Issues and Implications*. Longman, Ltd.
Leech, G. and J. Svartvik. 2002. *A Communicative Grammar of English*, (3rd ed.). Longman, Ltd.
Lu, X. 2010. Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics* 15(4), 474-496.
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.
Malvern, D., B. Richards, N. Chipere and P. Durán. 2004. *Lexical Diversity and Language Development: Quantification and Assessment*. Palgrave Macmillan.
Meyers, L. S., G. Gamst and A. J. Guarino. 2016. *Applied Multivariate Research: Design and Interpretation*, (3rd ed.). SAGE Publications, Inc.
Nation, I. S. P. 2001. *Learning Vocabulary in Another Language*. Cambridge University Press.
Norris, J. M. and L. Ortega. 2009. Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics* 30(4), 555-578.
Ortega, L. 1999. Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition* 21(1), 109-148.
Ortega, L. 2003. Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics* 24(4), 492-518.
Read, J. 2000. *Assessing Vocabulary*. Cambridge University Press.
Richards, J. C. 2001. Curriculum Development in Language Teaching. Cambridge University Press.
Ryu, J. and M. Jeon. 2020. An analysis of the continuity of the listening scripts of middle school English textbooks with Coh-Metrix. *English Language & Literature Teaching* 26(3), 101-121.
Sheldon, L. E. 1988. *Evaluating ELT textbooks and materials*. *ELT Journal* 42(4), 237-246.
Templin, M. 1957. *Certain Language Skills in Children: Their Development and Inter-relationships*. University of Minnesota Press.
Tomlinson, B. 2011. *Materials Development in Language Teaching*. Cambridge University Press.
Treffers-Daller, J., P. Parslow and S. Williams. 2018. Back to basics: how measures of lexical diversity can help discriminate between CEFR levels. *Applied Linguistics* 39(3), 302-327.
Ur, P. 1996. A Course in Language Teaching: Practice and Theory. Cambridge University Press.
Vygotsky, L. S. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
Wolfe-Quintero, K., S. Inagaki and H.-Y. Kim. 1998. *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. University of Hawai'i Press.
Xanthos, A. and S. Gillis. (2010). Quantifying the development of inflectional diversity. *First Language* 30(2), 175-198.
Xanthos, A. and G. Guex. (2015). On the robust measurement of inflectional diversity. In A. Tuzzi, M. Benesova and J. Macutek, eds., *Recent Contributions to Quantitative Linguistics*, 241-254. De Gruyter.
Yang, K. and J. Bae. 2022. A continuity and difficulty analysis of the reading texts in Korean high school English textbooks with the 2015 Revised National Curriculum. *English Teaching* 77(sup. 1), 43-62.

Examples in: English
Applicable Languages: English
Applicable Level: All