



**TEACHING ADVANCED DISCOURSE: A CORPUS-BASED EVALUATION OF
INTERMEDIATE BOOK 2 FOR LANGUAGE LEARNING**

Azhar Munir Bhatti – (Corresponding Author) azharmunir18@gmail.com

PhD Scholar, University of Education, Lahore

Assistant Professor of English, Higher Education Department, Punjab

Prof. Dr. Ahsan Bashir

Director, Division of Arts and Social Science, University of Education, Lahore

Abstract:

This study evaluates the pedagogical effectiveness of Intermediate English Book 2 in fostering advanced discourse competence through a corpus-based and natural language processing (NLP) framework. It systematically analyzes linguistic features such as syntactic complexity, clause distribution, tense–aspect usage, and lexical sophistication, mapping these features to proficiency descriptors from the Common European Framework of Reference for Languages (CEFR) and the ACTFL Proficiency Guidelines. Findings indicate that the textbook demonstrates strong alignment with B2–C1 (CEFR) and Advanced (ACTFL) levels, evidenced by high sentence complexity, frequent clause embedding, and diverse discourse functions including narration and description. The inclusion of varied genres—scientific, historical, and literary—supports learners’ discourse flexibility and cognitive development. However, the study identifies comparatively limited emphasis on argumentative discourse and potential cognitive overload due to high linguistic density. Overall, the textbook is deemed a pedagogically robust resource that effectively supports the transition from intermediate to advanced proficiency, while highlighting areas for improvement in task design and discourse balance.

Keywords: Corpus-based analysis; NLP in language education; textbook evaluation; CEFR; ACTFL; discourse competence; syntactic complexity; genre variation; lexical sophistication; SLA

1. Introduction

Textbooks remain central instruments in second language acquisition, particularly in formal educational settings where they function as both pedagogical guides and linguistic input sources. Rather than serving as passive repositories of content, contemporary textbooks actively structure learners’ progression toward communicative competence and proficiency benchmarks. Recent research highlights that effective textbooks integrate linguistic, cognitive, and discourse-level development, enabling learners to transition from sentence-level accuracy to extended discourse production (Tomlinson, 2015; McGrath, 2013). In upper-secondary contexts such as Grade 12, this role becomes more critical, as learners are expected to engage with complex texts, abstract ideas, and varied genres. Empirical evidence suggests that textbooks significantly influence learners’ ability to develop advanced discourse skills, including cohesion, coherence, and syntactic sophistication, especially when these materials are systematically sequenced and grounded in authentic language use (Curd-Christiansen & Weninger, 2015; Nguyen & Le, 2020; Mishan, 2015).

A key dimension of textbook evaluation in contemporary language education is alignment with internationally recognized proficiency frameworks such as the Common European Framework of Reference for Languages and the ACTFL Proficiency Guidelines. These frameworks provide empirically grounded descriptors of language ability, emphasizing not only grammatical accuracy but also functional language use across communicative contexts. The CEFR, particularly in its Companion Volume, foregrounds mediation, interaction, and plurilingual competence, thereby expanding the traditional notion of proficiency beyond structural knowledge (Piccardo, North & Goodier, 2019; Piccardo & North, 2019). Similarly, ACTFL guidelines stress the ability to sustain discourse across time frames, manage abstract



topics, and demonstrate linguistic flexibility. Alignment with these frameworks ensures that instructional materials are not arbitrarily graded but instead reflect measurable, internationally comparable standards. Recent studies further argue that such alignment enhances curricular transparency, facilitates assessment standardization, and supports learner autonomy through clearly defined proficiency goals (Piccardo, 2020; Council of Europe, 2020).

In parallel, advances in corpus linguistics and natural language processing (NLP) have transformed the evaluation of language textbooks, enabling objective, data-driven analysis of linguistic features. Corpus-based approaches allow researchers to examine syntactic complexity, lexical patterns, and discourse structures at scale, thereby overcoming the limitations of impressionistic evaluation (Biber, 2019; Friginal & Prado, 2025). This methodological shift is particularly relevant for assessing advanced discourse, where features such as clause embedding, lexical bundles, and tense-aspect variation play a crucial role in meaning-making. Studies within the last decade demonstrate that corpus-informed textbook analysis provides deeper insights into the extent to which instructional materials reflect authentic language use and support progressive language development (Staples et al., 2016; Farr, & Leńko-Szymańska, 2024; O’Keeffe & McCarthy, 2021).

In the case of Intermediate English Book 2, preliminary corpus-derived metrics indicate a strong orientation toward upper-intermediate to advanced proficiency levels (B2–C1), characterized by the use of complex syntactic structures, varied clause types, and extended discourse patterns. Unit-level analysis further reveals consistent syntactic sophistication, with mean sentence lengths ranging from approximately 17 to 19 words and clause density averaging around 2.1 clauses per sentence, suggesting a sustained exposure to cognitively demanding language forms. Additionally, the presence of diverse genres—including scientific exposition, historical narration, and literary humor—indicates an attempt to develop learners’ genre awareness and discourse flexibility, both of which are essential components of advanced language proficiency.

Against this backdrop, the present study aims to evaluate the pedagogical suitability of Intermediate English Book 2 through a corpus-based NLP framework. Specifically, the study seeks to determine whether the textbook effectively supports the development of advanced discourse competencies by systematically analyzing its linguistic complexity, discourse features, and alignment with CEFR and ACTFL standards. By integrating quantitative corpus metrics with pedagogical analysis, this research contributes to the growing body of data-driven language education studies and offers a more rigorous, evidence-based perspective on textbook evaluation. Ultimately, the study aims to inform curriculum design, material development, and instructional practices in contexts where achieving advanced language proficiency is a central educational objective.

2. Literature Review

The Common European Framework of Reference for Languages provides one of the most comprehensive and empirically grounded models for describing language proficiency, particularly at the upper-intermediate (B2) and advanced (C1) levels. At these stages, learners are expected to demonstrate the ability to understand complex texts, produce clear and well-structured discourse, and engage in abstract and specialized communication. The CEFR Companion Volume further extends these descriptors by emphasizing mediation, plurilingual competence, and discourse management, thereby foregrounding the importance of functional and interactional language use (Council of Europe, 2020; Piccardo, North & Goodier, 2019). Recent studies argue that CEFR-aligned materials must incorporate features such as clause complexity, cohesive devices, and genre variation to support learners’ transition from



controlled production to autonomous discourse (Piccardo & North, 2019). In the Pakistani context, the relevance of CEFR alignment has been increasingly highlighted, particularly in relation to employability and global competitiveness, where English proficiency functions as a critical skill (Abbas et al., 2021). Moreover, research on local textbooks indicates that while they often claim CEFR alignment, there remains a gap between stated objectives and actual linguistic input, especially in terms of discourse-level features and task design (Bhatti & Bashir, 2025; Bhatti & Bashir, 2025). Additional scholarship also underscores the importance of cultural representation within CEFR-aligned materials, arguing that culturally responsive content enhances learner engagement and interpretive competence (Azim et al., 2018; Bhatti & Bashir, 2025). Collectively, these studies suggest that CEFR descriptors should not only guide curriculum design but also be operationalized through measurable linguistic features, which can be effectively analyzed using corpus-based approaches.

Parallel to CEFR, the ACTFL Proficiency Guidelines provide a functional framework for assessing language proficiency, particularly emphasizing real-world communicative ability across novice, intermediate, advanced, and superior levels. At the advanced level, learners are expected to narrate and describe in major time frames, handle abstract topics, and sustain paragraph-length discourse with increasing accuracy and complexity. Recent research highlights that ACTFL-aligned instruction requires exposure to authentic language input and opportunities for extended discourse production, which are often mediated through textbooks (Kim & Isbell, 2024; VanPatten et al., 2020). Studies have further demonstrated that textbooks frequently fall short in providing sufficient opportunities for discourse-level practice, particularly in writing, where learners face persistent challenges related to coherence, organization, and linguistic accuracy (Bhatti et al., 2020). In addition, the integration of cultural content has been identified as a key factor in achieving ACTFL outcomes, as it enables learners to interpret and produce language within meaningful sociocultural contexts (Bhatti & Bashir, 2025; Bhatti & Bashir, 2025). Research focusing on exercise design in Pakistani textbooks reveals that many tasks remain form-focused rather than discourse-oriented, thereby limiting learners' ability to achieve advanced proficiency benchmarks (Bhatti & Bashir, 2025). Furthermore, vocabulary recycling and systematic exposure to lexical patterns have been identified as essential components of proficiency development, yet they are often inconsistently implemented in instructional materials (Azim et al., 2020). These findings collectively underscore the need for a more rigorous, data-driven evaluation of textbooks to ensure alignment with ACTFL proficiency expectations.

The emergence of corpus-informed pedagogy has significantly reshaped the field of language education by introducing empirical, data-driven methods for analyzing and designing instructional materials. Corpus linguistics enables the systematic examination of authentic language use, including frequency patterns, collocations, and discourse structures, thereby providing a robust foundation for evaluating textbook content (Biber, 2019; Friginal & Prado, 2025). Recent studies emphasize that corpus-informed approaches can reveal discrepancies between textbook language and real-world usage, particularly in terms of lexical bundles, phraseology, and syntactic complexity (Farr, & Leńko-Szymańska, 2024). This is especially relevant for advanced learners, who require exposure to naturalistic discourse patterns to develop fluency and accuracy. In the context of textbook evaluation, corpus-based analyses have been used to assess vocabulary recycling, grammatical progression, and genre representation, offering insights into the pedagogical effectiveness of materials (O'Keeffe & McCarthy, 2021; McEnery & Hardie, 2011). Local research further highlights the importance of systematic vocabulary recycling in enhancing retention and language development,



particularly when supported by corpus-informed design principles (Azim et al., 2020). Additionally, the integration of NLP techniques has enabled more sophisticated analyses of linguistic features, including clause density, tense-aspect distribution, and discourse markers, thereby providing a more nuanced understanding of textbook complexity. These advancements support the argument that corpus-informed pedagogy is essential for bridging the gap between theoretical proficiency frameworks (such as CEFR and ACTFL) and actual instructional practices, ensuring that textbooks effectively facilitate the development of advanced discourse competence.

3. Methodology

This study adopts a corpus-based, quantitative research design to evaluate the pedagogical suitability of Intermediate English Book 2 for developing advanced discourse competence. The methodological approach integrates natural language processing (NLP) techniques with established linguistic frameworks to systematically analyze the structural and functional features of the textbook. By operationalizing proficiency descriptors through measurable linguistic indices, the study aims to provide an objective evaluation of how effectively the textbook aligns with advanced-level language learning outcomes. Corpus-informed methodologies have been widely recognized as robust tools for analyzing authentic language use and instructional materials, particularly in identifying patterns of syntactic complexity, lexical variation, and discourse organization (Biber, 2019; Farr, & Leńko-Szymańska, 2024; McEnery & Hardie, 2011).

3.1 Framework: Mapping Linguistic Features to Proficiency Scales

The analytical framework of this study is grounded in the alignment of observable linguistic features with proficiency descriptors derived from the Common European Framework of Reference for Languages and the ACTFL Proficiency Guidelines. Rather than relying on subjective textbook evaluation, this study operationalizes proficiency by mapping specific linguistic indicators—such as clause density, syntactic embedding, and tense-aspect variation—to the competencies described at the B2–C1 and Advanced levels, respectively.

Recent research emphasizes that proficiency at higher levels is characterized by the ability to manage complex syntactic structures, produce extended discourse, and demonstrate flexibility across communicative contexts (Piccardo & North, 2019). In line with this, the present study adopts a feature-based mapping approach, where linguistic complexity indices serve as proxies for discourse-level competence. For instance, higher mean sentence length (MLS), increased clause-per-sentence ratios (C/S), and the presence of subordinate clauses are associated with advanced proficiency, as they reflect the ability to express nuanced relationships such as causality, condition, and evaluation (Biber, Reppen, & Friginal, 2012).

The mapping process involves three key stages. First, CEFR and ACTFL descriptors relevant to advanced discourse are identified and categorized into structural (e.g., syntactic complexity), functional (e.g., narration across time frames), and lexical (e.g., range and precision of vocabulary) dimensions. Second, these dimensions are translated into quantifiable corpus metrics using NLP-based parsing techniques. Third, the extracted metrics from the textbook corpus are compared against benchmark values reported in corpus-based studies of proficient language use. This approach ensures that the evaluation is both theoretically grounded and empirically verifiable, addressing concerns raised in recent literature بشأن the lack of objective alignment between textbooks and proficiency standards (Friginal & Prado, 2025; O’Keeffe & McCarthy, 2021).



3.2 Measures Used

To capture the multidimensional nature of advanced discourse, the study employs a set of interrelated linguistic measures, each corresponding to a specific aspect of language proficiency. These measures are derived from corpus linguistics and second language acquisition research, where they have been validated as reliable indicators of linguistic development.

3.2.1. Sentence Complexity

Sentence complexity is assessed through indices such as mean length of sentence (MLS), mean length of clause (MLC), and clauses per sentence (C/S). These metrics provide insights into the degree of syntactic elaboration and cognitive demand embedded in the text. Research indicates that advanced learners are typically exposed to longer and more structurally complex sentences, which require the processing of multiple clauses and hierarchical relationships (Biber, 2019; Saddique, Nadeem, & Danish, 2025; Park, 2022). Higher values of MLS and C/S are therefore interpreted as indicators of advanced-level input, particularly when accompanied by varied clause structures and coordination patterns.

3.2.2. Clause Types

Clause-level analysis focuses on the distribution and frequency of independent and dependent clauses, including adverbial, relative, nominal, and complement clauses. The presence of diverse clause types is essential for expressing complex meanings and achieving discourse coherence. Studies have shown that advanced proficiency is associated with increased use of subordination and clause embedding, which enable learners to articulate logical relationships and integrate information across sentences (Staples et al., 2016; Biber, Reppen, & Friginal, 2012). In this study, clause types are identified using rule-based parsing techniques, allowing for a detailed examination of how the textbook supports the development of syntactic sophistication.

3.2.3. Tense/Aspect Distribution

Tense and aspect usage is analyzed to evaluate the extent to which the textbook facilitates temporal and aspectual variation, a key requirement for advanced discourse. According to ACTFL guidelines, advanced learners must be able to narrate and describe events across multiple time frames, requiring control over a range of tense-aspect forms. Corpus-based research demonstrates that proficient language use involves not only the accurate use of basic tenses but also the integration of perfect and progressive aspects to convey temporal relationships and viewpoint (Friginal & Prado, 2025; Bardovi-Harlig, 2000). The present study quantifies the distribution of tense-aspect forms per thousand words, enabling a systematic comparison with established benchmarks.

3.2.4. Lexical Sophistication

Lexical sophistication is examined through measures such as lexical diversity, frequency of advanced vocabulary, and the use of lexical bundles. Advanced proficiency is characterized by the ability to employ a wide range of vocabulary with precision and appropriateness, as well as to utilize recurrent multi-word expressions that contribute to fluency and coherence (O’Keeffe & McCarthy, 2021; Kyle & Crossley, 2017; Kolahi Ahari et al., 2025). Corpus-informed studies highlight that exposure to high-frequency lexical bundles and low-frequency academic vocabulary is essential for developing discourse competence. In this study, lexical features are extracted using frequency-based analysis, with particular attention to the recycling of vocabulary and the presence of formulaic sequences.

Overall, the methodological framework combines theoretical rigor with empirical precision by linking proficiency descriptors to quantifiable linguistic features. This integration of corpus



linguistics and NLP techniques allows for a comprehensive evaluation of the textbook, ensuring that conclusions regarding its pedagogical suitability are grounded in measurable evidence rather than subjective judgment.

4. Results

The results of the corpus-based analysis provide empirical evidence regarding the alignment of Intermediate English Book 2 with advanced proficiency frameworks and its pedagogical effectiveness in fostering discourse competence. The findings are organized into three main dimensions: CEFR alignment, ACTFL alignment, and pedagogical features.

4.1 CEFR Alignment

The analysis indicates that Intermediate English Book 2 demonstrates strong alignment with B2–C1 descriptors of the Common European Framework of Reference for Languages, particularly in terms of grammatical complexity and engagement with abstract content.

Table 1: Syntactic Complexity (Overall)

Measure	Value Range	Interpretation
Mean Sentence Length (MLS)	14.2 – 20.4	Moderate to high complexity
Mean Clause Length (MLC)	5.8 – 12.6	Increasing clause elaboration
Clauses per Sentence (C/S)	1.6 – 2.4	Multi-clausal discourse
Dependent Clauses Ratio	Moderate–High	Frequent subordination

The variation between the first and second halves of the book (MLS increasing from ~14 to ~20) indicates a **progressive increase in syntactic complexity**, which is consistent with CEFR expectations for scaffolding toward C1-level discourse. The high clause density and frequent subordination reflect learners' exposure to:

- Cause–effect relations
- Conditional reasoning
- Relative clause embedding

These are core features of **advanced discourse competence**, enabling learners to process and produce extended, logically structured texts.

Abstract and Academic Topics

Across the dataset, lexical and structural features correspond to:

- Scientific discourse (procedural + explanatory language)
- Historical narration (temporal sequencing)
- Literary texts (stylistic variation)

The presence of such genres confirms alignment with CEFR descriptors requiring engagement with **complex, abstract, and specialized topics**. In addition, the textbook incorporates a wide range of abstract and academic topics, including scientific explanations (e.g., *The Dying Sun*), methodological discourse (e.g., *Using the Scientific Method*), and socio-historical narratives (e.g., *China's Way to Progress*). Such thematic diversity supports the development of conceptual and analytical language skills, which are central to CEFR's higher-level descriptors. The progression from moderately complex structures in earlier units to more elaborated constructions in later units further suggests a scaffolded approach to discourse development.

4.2 ACTFL Alignment

The findings also demonstrate substantial alignment with Advanced-level descriptors of the ACTFL Proficiency Guidelines, particularly in relation to discourse functions such as narration, description, and argumentation.



Table 2: Tense–Aspect Distribution (Overall)

Category	Distribution Trend	Functional Role
Past Simple	Dominant	Narrative sequencing
Present Simple	High	Generalization & explanation
Perfect Aspects	Moderate	Retrospective meaning
Progressive	Limited but present	Ongoing processes
Modals	Frequent	Evaluation, possibility
Conditionals	Present	Hypothetical reasoning

The dominance of past tense forms supports narrative discourse, enabling learners to recount events and sequences, a key requirement at the ACTFL Advanced level. The presence of perfect and progressive aspects facilitates temporal layering and aspectual nuance, allowing learners to express ongoing actions and completed events with relevance to the present.

Narration: Historical and literary texts provide rich input for narrative discourse, characterized by chronological sequencing and temporal markers.

Description: Scientific and expository passages require detailed description of processes and phenomena, supporting learners’ ability to produce precise and elaborated descriptions. The use of NPs, and PPs show enhanced elaboration techniques of the essays.

Argumentation: Texts involving scientific reasoning and socio-political themes encourage evaluative language and logical structuring, although explicit argumentative tasks may be less frequent compared to narrative and descriptive functions.

Overall, the distribution of tense-aspect forms and discourse functions indicates that the textbook provides substantial support for the development of sustained, paragraph-level discourse, consistent with ACTFL Advanced-level expectations.

4.3 Pedagogical Features

The corpus analysis further reveals that Intermediate English Book 2 incorporates diverse pedagogical features that contribute to discourse development through varied content and stylistic exposure.

Table 3: Genre and Pedagogical Exposure

Genre Type	Examples	Pedagogical Function
Scientific	<i>The Dying Sun, Scientific Method</i>	Develops analytical reasoning and technical vocabulary
Historical	<i>Mustafa Kamal</i>	Enhances narrative skills and temporal structuring
Literary/Humor	<i>My Financial Career, The Man Who Was a Hospital</i>	Introduces stylistic variation and pragmatic competence

The inclusion of **scientific reasoning** texts promotes logical structuring and cause-effect relationships, which are essential for academic discourse. These texts require learners to process complex information and engage with abstract concepts, thereby strengthening cognitive and linguistic skills simultaneously.

The presence of **historical narration** supports the development of chronological discourse and the use of varied tense-aspect forms. Such texts enable learners to construct extended narratives and understand the sequencing of events, aligning with both CEFR and ACTFL descriptors.

Finally, the incorporation of **humor and stylistic variation** introduces learners to pragmatic and rhetorical aspects of language use. Humorous texts, in particular, often involve subtle



linguistic cues, irony, and cultural references, which contribute to higher-level interpretive skills and discourse flexibility.

Table 4: Clause Type Distribution (Full Corpus)

Clause Type	Approx. Frequency	Pedagogical Function
Independent Clauses	High	Core sentence construction
Adverbial Clauses	High	Logical relations (cause, time)
Relative Clauses	Moderate	Information expansion
Nominal Clauses	Limited–Moderate	Abstract reasoning
Complement Clauses	Low	Structural variation

Table 5: Phrase Type Distribution

Phrase Type	Frequency	Function
Noun Phrases	Very High	Information density
Verb Phrases	High	Action/process expression
Prepositional Phrases	High	Contextual elaboration
Adjective/Adverb Phrases	Moderate	Description and evaluation

Interpretation of Pedagogical Exposure

Scientific Reasoning

- Evidenced by:
 - High use of adverbial clauses (cause, condition)
 - Dense noun phrases (technical information)
- Supports **analytical thinking and academic discourse**

Historical Narration

- Supported through:
 - Past tense dominance
 - Sequential clause structures
- Develops **temporal organization skills**

Humor and Style Variation

- Literary texts show:
 - Increased coordination
 - Stylistic shifts
- Enhances:
 - Pragmatic awareness
 - Interpretive flexibility

The analysis across all lessons demonstrates that Intermediate English Book 2 provides:

- Consistent exposure to advanced syntactic structures
- Strong support for narrative and descriptive discourse
- Moderate development of argumentative competence
- Rich genre diversity supporting cognitive and linguistic growth

The progression in complexity, combined with varied discourse functions and lexical richness, confirms that the textbook is **well-aligned with B2–C1 (CEFR) and Advanced (ACTFL) proficiency levels**. However, the relative underrepresentation of explicit argumentative structures suggests an area for pedagogical enhancement.

Overall, the results indicate that Intermediate English Book 2 provides a linguistically rich and pedagogically diverse input that aligns well with advanced proficiency benchmarks. The



consistent use of complex grammatical structures, the integration of multiple discourse functions, and the exposure to varied genres collectively support the development of advanced discourse competence. However, while the textbook demonstrates strong alignment in terms of input, further analysis of task design and learner engagement would be necessary to determine the extent to which these features translate into productive language skills.

5. Discussion

The findings of this study indicate that Intermediate English Book 2 demonstrates substantial suitability for intermediate-to-advanced learners, particularly those operating within the B2–C1 range of the Common European Framework of Reference for Languages and the Advanced level of the ACTFL Proficiency Guidelines. The corpus-based results confirm that the textbook consistently exposes learners to complex syntactic structures, varied discourse functions, and genre diversity, all of which are essential for developing advanced communicative competence. These findings align with international research suggesting that progression to higher proficiency levels requires sustained engagement with authentic, linguistically rich input that reflects real-world discourse patterns (Biber, 2019; Staples et al., 2016).

One of the most significant strengths of the textbook lies in its **rich linguistic input**, as evidenced by high sentence complexity, clause density, and lexical variety. The presence of multi-clausal structures and varied clause types enables learners to process and produce complex meanings, including causality, temporality, and evaluation. This supports the development of discourse competence, which is a defining feature of advanced proficiency (Piccardo & North, 2019). Furthermore, the distribution of tense-aspect forms and the inclusion of multiple genres—scientific, historical, and literary—provide learners with opportunities to engage in narration, description, and, to a lesser extent, argumentation. Such exposure is consistent with findings from corpus-informed pedagogy, which emphasize the importance of authentic input and discourse-level variation in facilitating language development (O’Keeffe & McCarthy, 2021; Farr, & Leńko-Szymańka, 2024).

The results also address the central research objective of this study—namely, whether Intermediate English Book 2 is pedagogically suitable for fostering advanced discourse competence using NLP-based evaluation. The evidence suggests that the textbook not only aligns with CEFR and ACTFL descriptors but also operationalizes these frameworks through measurable linguistic features. The progressive increase in syntactic complexity across the text, combined with the systematic use of discourse functions, indicates that the material is designed to scaffold learners toward higher levels of proficiency. This supports the argument that corpus-based evaluation provides a reliable and objective method for assessing textbook quality, as it bridges the gap between theoretical descriptors and actual linguistic input (McEnery & Hardie, 2011; Biber, Reppen & Friginal, 2012).

However, despite these strengths, the analysis also reveals certain limitations. A key concern is the potential for **cognitive overload**, particularly for learners transitioning from lower-intermediate levels. The high density of complex sentences, frequent subordination, and abstract content may pose processing challenges, especially in the absence of sufficient scaffolding or guided practice. Research in second language acquisition suggests that excessive linguistic complexity can hinder comprehension and reduce learning efficiency if not accompanied by appropriate instructional support (VanPatten et al., 2020; Crossley et al., 2022). Additionally, while the textbook provides substantial input for narrative and descriptive discourse, the relatively limited emphasis on explicit argumentative structures may restrict learners’ ability to develop critical writing skills, which are increasingly important at advanced levels.



In summary, the discussion confirms that Intermediate English Book 2 is largely effective in supporting advanced language development, particularly through its rich linguistic input and alignment with international proficiency frameworks. At the same time, the findings highlight the need for careful pedagogical mediation to mitigate cognitive load and to enhance the development of underrepresented discourse functions such as argumentation.

6. Conclusion

This study set out to evaluate the pedagogical suitability of Intermediate English Book 2 through a corpus-based NLP framework, focusing on its alignment with CEFR and ACTFL proficiency standards and its effectiveness in developing advanced discourse competence. The findings demonstrate that the textbook serves as a **strong bridge to advanced proficiency**, providing learners with sustained exposure to complex syntactic structures, varied discourse functions, and diverse genres.

The corpus analysis confirms that the textbook aligns well with B2–C1 descriptors of the Common European Framework of Reference for Languages and Advanced-level expectations of the ACTFL Proficiency Guidelines. Its emphasis on multi-clausal structures, tense-aspect variation, and lexical richness supports the development of key competencies required for higher-level language use, including narration, description, and analytical reasoning.

Moreover, the integration of scientific, historical, and literary texts makes the textbook particularly effective for **academic language development**, as it exposes learners to the types of discourse commonly encountered in educational and professional contexts. This aligns with contemporary views in language education that emphasize the importance of discipline-specific literacy and discourse awareness in achieving advanced proficiency.

At the same time, the study highlights areas for improvement, particularly the need to balance linguistic complexity with pedagogical accessibility and to enhance opportunities for argumentative discourse. Addressing these aspects would further strengthen the textbook's role in preparing learners for the demands of advanced language use.

In conclusion, Intermediate English Book 2 can be regarded as a pedagogically robust resource that effectively supports the transition from intermediate to advanced proficiency. Its strengths lie in its rich linguistic input and alignment with international standards, making it a valuable tool for both learners and educators aiming to achieve high levels of English language competence.

References:

- Abbas, F., Rana, A. M. K., Bashir, I., & Bhatti, A. M. (2021). The English language proficiency as a global employment skill: the viewpoint of Pakistani academia. *Humanities and Social Sciences Review*, 9(3), 1071-1077.
- Azim, M. U., Bhatti, A. M., Hussain, Z., & Iqbal, M. (2018). Culturally oriented textbooks and English language teachers. *Hamdard Islamicus*, 41(4), 301-317.
- Azim, M. U., Hussain, Z., Bhatti, A. M., & Iqbal, M. (2020). Recycling of vocabulary in English Language Teaching: From theory to practice. *Epistemology*, 7(1), 88-102.
- Bardovi-Harlig, K. (2000). Tense and aspect in second language acquisition: Form, meaning, and use. *Language Learning: A Journal of Research in Language Studies*, 50, 1.
- Bhatti, A. M., & Bashir, A. (2025). Do English Textbooks Deliver? A Critical Analysis of Exercise Design and Expected Learning Outcomes in Punjab's Intermediate English Textbook 1. *Critical Review of Social Sciences and Humanities*, 5(1), 43-62.



- Bhatti, A. M., & Bashir, A. (2025). Exploring the representation of Pakistani culture in intermediate English poetry: A critical analysis of selected poems through CEFR. *Contemporary Journal of Social Science Review*, 3(4), 1-20.
- Bhatti, A. M., & Bashir, A. (2025). From Textbooks to Tests: The Role of Cultural Integration in Shaping Outcomes of English Language Assessment in Punjab, Pakistan.
- Bhatti, A. M., & Bashir, A. (2025). Literary Language Meets CEFR: How Mr. Chips Develops Upper-Intermediate Reading Competence in Pakistani Classrooms. *AL-HAYAT Research Journal (AHRJ)*, 3(1), 100-109.
- Bhatti, A. M., Hussain, Z., Azim, M. U., & Gulfam, G. Q. (2020). Perceptions of ESL learners and teachers on writing difficulties in English language learning in Lahore. *International Bulletin of Linguistics and Literature (IBLL)*, 3(3), 11-24.
- Biber, D. (2019). Text-linguistic approaches to register variation. *Register studies*, 1(1), 42-75.
- Biber, D., Reppen, R., & Friginal, E. (2012). Research in Corpus Linguistics. In Robert B. K. (ed.), *The Oxford Handbook of Applied Linguistics*, (2nd ed.), Oxford Handbooks, <https://doi.org/10.1093/oxfordhb/9780195384253.013.0038>.
- Council of Europe (COE). (2020). *Common European framework of reference for languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing. www.coe.int/lang-cefr
- Crossley, S., Skalicky, S., Berger, C., & Heidari, A. (2022, July). Assessing readability formulas in the wild. In *Conference on Smart Learning Ecosystems and Regional Development* (pp. 91-101). Singapore: Springer Nature Singapore.
- Curdt-Christiansen, X. L., & Weninger, C. (2015). Introduction: Ideology and the politics of language textbooks. In X. L. Curdt-Christiansen, & C. Weninger (Eds.), *Language, ideology and education. The politics of textbooks in language education* (pp. 1–8). Routledge.
- Farr, F., & Leńko-Szymańska, A. (2024). Corpora in English language teacher education: Research, integration, and resources. *TESOL Quarterly*, 58(3), 1181-1192.
- Kyle, K., & Crossley, S. A. (2017). Assessing syntactic sophistication in L2 writing: A usage-based approach. *Language Testing*, 34(4), 513-535. <https://doi.org/10.1177/0265532217712554>
- Kolahi Ahari, M., Ghonsooly, B., Ghapanchi, Z., & Soodmand Afshar, H. (2025). Examining the Role of Lexical Sophistication, Lexical Diversity, Syntactic Sophistication, Syntactic Complexity, and Cohesion in L2 Speaking Proficiency Assessment. *International Journal of Language Testing*, 15(2), 43-70.
- Piccardo, E. (2020). *The CEFR and language education: Past, present, and future*. Language Education in Review Series.
- Saddique, A., Nadeem, M. A., & Danish, M. (2025). Syntactic Complexity and Lexical Density in EFL Academic Writing: A Corpus-Based Investigation of Developmental Patterns. *Liberal Journal of Language & Literature Review*, 3(4), 1173-1190.
- Park, S. (2022). A corpus-based comparison of syntactic complexity in spoken and written learner language. *Canadian Journal of Applied Linguistics*, 25(2), 47-70.
- McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
- McGrath, I. (2013). *Teaching materials and the roles of EFL/ESL teachers*. Bloomsbury.



- Mishan, F. (2015). *Materials development for TESOL*. Edinburgh University Press.
- Nguyen, C. T., & Le, D. T. K. (2020). Communicative Language Teaching: Do Tasks and Activities in School Textbooks Facilitate Learners' Development of Communicative Competence?. *Journal of Language Teaching & Research*, 11(5).
- Piccardo, E., North, B., & Goodier, T. (2019). Broadening the scope of language education: Mediation, plurilingualism, and collaborative learning: The CEFR companion volume. *Journal of e-learning and knowledge society*, 15(1).
- O'Keeffe, A., & McCarthy, M. (2021). *The Routledge handbook of corpus linguistics*. Routledge.
- Piccardo, E., & North, B. (2019). *The action-oriented approach: A dynamic vision of language education* (Vol. 72). Multilingual Matters.
- Friginal, E., & Prado, M. (2025). Corpus-informed materials in language teaching. In *The Palgrave encyclopedia of computer-assisted language learning* (pp. 1-5). Cham: Springer Nature Switzerland.
- Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written communication*, 33(2), 149-183.
- Tomlinson, B. (2015). *Materials development for language teaching* (2nd ed.). Cambridge University Press.
- Kim, J., & Isbell, D. R. (2024). Test Review: ACTFL Assessment of Performance Toward Proficiency in Languages (AAPPL). *Language Assessment Quarterly*, 21(4-5), 516-527.
- VanPatten, B., Smith, M., & Benati, A. G. (2020). *Key questions in second language acquisition: An introduction*. Cambridge University Press.