Policy Research Journal
ISSN (E): 3006-7030 ISSN (P) : 3006-7022 Volume 4, Issue 1, 2026

EXPLORING LINGUISTIC PATTERNS IN THE 7TH GRADE ENGLISH TEXTBOOK: A CORPUS-BASED STUDY WITH PEDAGOGICAL REFLECTIONS FROM TEACHERS

Saba Javaid[1], Dr. Muhammad Arfan Lodhi[*2]

[1]Research Scholar, Department of English, NCBA&E (Alhamra) University, Bahawalpur
*[*2]Higher Education Department, Punjab*

[1]sabajaved5502@gmail.com, [*2]samaritan_as@hotmail.com

DOI: https://doi.org/10.5281/zenodo.24495553

Keywords
Corpus Linguistics, Grade 7 English Textbook, Vocabulary Load, Collocations, Lexical Density, Corpus-Based Pedagogy, Teachers' Perception

Article History
Received: 25 November 2025
Accepted: 05 January 2026
Published: 19 January 2026

Copyright @Author
Corresponding Author: *
Dr. Muhammad Arfan Lodhi

Abstract

The objective of this study is to analyze the linguistic aspects of the Grade 7 English textbook by using the corpus-based approach and teachers' reflections. It intends to investigate how teachers perceive the textbook's linguistic suitability and to explore its collocational patterns and vocabulary including learner's linguistic requirements and proficiency levels; and to create and test a corpus-based pedagogical unit. The study is a mixed one consisting of three stages. The first phase entailed data collection from 60 male and 60 female Grade 7 English teachers from public schools using a questionnaire. The results of the questionnaire indicated that teachers felt that the vocabulary in the textbooks was very useful for learning but at the same time they experienced difficulties in vocabulary learning and retention as well as difficulties in explaining the vocabulary. The CEFR profiling showed differences in vocabulary levels between units; the frequency, collocation and KWIC analyses showed evidence of lexical patterns. The feedback from the workshops was mostly positive with teachers saying that the module was relevant and useful for their vocabulary selection, contextual teaching and lesson planning as well as their classroom practice. The study results show that the corpus-based analysis provides systematic evidence to be used in evaluating the language of the textbooks, as well as in guiding vocabulary teaching. It suggests corpus-resourced materials to reinforce the coherence of vocabulary in textbooks with learners' linguistic demands and teaching and learning.

1.1 Introduction

Classroom instruction using textbook was a key factor in influencing students' language input, particularly in the case of English as a foreign language (EFL), where classroom materials tend to be the main source of language input for students. (EFL Textbook 2024). The study aimed to explore the linguistic patterns of the 7th grade English textbook from the secondary education system in Pakistan using corpus-based approach and incorporating pedagogical insights from practicing English language teachers. The study revealed the frequency distribution of lexical items, frequent collocations, grammatical features and discourse structures in the textbook and benchmarked them with the norms of authentic language. Adopting a corpus-based approach enabled researchers to go beyond the descriptive

checklist and subjective judgment approach and provide an empirical study of linguistic features like lexical frequency, collocations, vocabulary coverage etc. Corpus methodologies used authentic language data to measure the conformity of the language used in textbooks to the natural language and uncover patterns that could not be found in traditional assessments (Nguyet & Van Long, 2020). The main source of language input for students was school textbooks and their analysis of linguistic patterns was necessary because the school textbooks had a significant impact on students' vocabulary, Grammer, and language skills during the critical period in which language learning takes place. The traditional textbook evaluation had been based mainly on the intuition of the experts, while the corpus-based method offered a more systematic approach of analysis. For 7th grade students, the transition from basic to intermediate proficiency, investigating lexical density or collocation could be used both for strengths and gaps in the textbook and to support fair and effective language learning approaches, using electronic collections of authentic language data to offer empirical methods for analyzing patterns of vocabulary, grammar, and discourse features in instructional materials (Khojasteh& Mukundan, 2025).

The shift of the study from theoretical grammar analysis to data-driven investigation of the language that has taken place in the field of corpus linguistics since the 1990s, which focuses on actual use of language instead of relying on intuition. The use of corpus tools such as AntConc with textbooks brought a new perspective to the use of corpus in textbooks by making it possible to discover hidden patterns of frequency and recurring features in learner materials. It also coincided with the Study of Systemic Functional Linguistics (SFL) dealing with the books in terms of genre structure and meaning. The recent research on the Pakistani English textbooks revealed the shortage of lexical range and authentic collocation structures, which reflect inadequacy in promoting a communicative competence (Saif et al., 2025). The new trend in education focused on communicative competence, cultural awareness and contextualization rather than on memorization of grammatical structures. Teachers' pedagogical reflections helped to move towards this by identifying strategies that helped students engage in the language in culturally relevant ways. Corpus-informed instruction was also used to promote data driven learning (DDL), where students analysed real language data and then used their own skills and expertise to identify patterns and make sense of the data. Furthermore, corpus-based assessments also gave quantitative feedback on the linguistic patterns in the text books and helped in pedagogical decision making by identifying the difference between classroom materials and the language used. These insights were very important in situations such as Pakistan where English textbooks required having a balance between both the global communicative standards and local educational objectives.

1.2 Background of the study

In the ESL classroom, textbooks were a very important part of the language input, particularly in public schools where textbooks were the main source of language input. The purpose of the grade 7 English textbook was to cultivate the middle school students' reading, writing, vocabulary and grammar. It familiarized the students with structured language use, basic communicative functions and literary and informational texts. But sometimes, the language in textbooks was created according to writers' intuition, without regard for the actual usage of language. Therefore, students may not have had enough opportunity to see and hear natural word combinations, common vocabulary, and natural grammatical patterns. The scientific approach used for the analysis of the language patterns, e.g., word frequency, collocations, lexical bundles, grammatical structures, was provided by corpus linguistics. The availability of corpus tools enabled the evaluation of the Grade 7 English textbook in order to see if the language used is authentic and useful English for learners.

Le Foll (2002) designed a study that did not involve human participants to gather and analyze textual data from two sources: the Textbook

English Corpus (TEC) containing 42 secondary-level EFL textbooks and three reference corpora of native English: spoken, youth fiction, and informative. The principal conclusions indicated that real world English was not the same as textbook English. Textbook dialogues were more dramatic, textbook fiction more realistic. The language used for instruction was also very particular. The study was in-depth and employed advanced computer analysis and a large number of original materials. It also looked at the trustworthiness of English textbooks and confirmed the differences between dialogues in the textbooks and in native English texts. The main limitation of the study was that while it demonstrated the lack of authenticity of the language in learning text books, it failed to claim that this negative impact of such language inputs on pupils. The study was relevant to the present research in the sense that it also used corpus analysis in English textbooks. As for textbooks, similarly Rosyida (2016) studied the role of textbooks as valuable teaching materials that gave structure, guidance to the syllabus and pre-made learning exercises. The study revealed that teachers were satisfied with textbooks since it ensures curriculum coverage and reduces preparation time. The study did an important job highlighting the importance of teacher adaptation. The content of textbooks was provided, modified, or deleted by the teachers to better meet the needs of their students and the context of their classrooms. The qualitative method was helpful to give information on real practices of teaching. The study, however, had a limited sample size and hence limited generalizability of the results. Moreover, the research did not mainly consider student learning outcomes but rather perceptions. In spite of these restrictions, the paper emphasized the notion that textbooks were not prescriptive scripts, and therefore that teachers' agency was necessary to successful content delivery.

Similarly, Saif et al. (2025) did a study on the Pakistani level English textbooks through the analysis of corpus-based frequency and collocation analysis. The study found that the textbooks primarily utilized a few high frequency words limiting students' access to a broad range of authentic vocabulary. Many words were used multiple times and for educational purposes rather than the way they are in natural language, leading to a decrease in lexical authenticity. Students may have had problems with language use, with low collocational criteria. The corpus approach method was used instead of the subjective approach to provide scientific and data-based results, thereby increasing the strength of the study. The study did not, however, consider the direct impact of these textbook constraints on students' communication skills. The study overall revealed that textbooks required corpus-informed design to make them more familiar to classroom materials and usage of English.

Siddique et al. (2025) stated that the content difficulty of PTB English textbooks for intermediate students should be checked by analyzing the difficulty of the content using the corpus-based methodology. The analysis of the three linguistic features (lexical density, sentence length and readability) revealed that textbooks were not uniformly complicated, and this may have affected the growth of students' vocabulary and understanding. The purpose and reproducible methodology of the study clearly illustrated the organization of textbooks. It was, however, constrained by its reliance on intermediate level textbooks and it did not include an examination of how teachers dealt with complexity in the classroom, leading to a lesser extent of exploration of classroom pedagogical implications. For vocabulary, Yu and Renandya (2021) conducted a detailed analysis of the coverage, repetition and dispersion of the text in New Senior English for China, which was based on the English Curriculum Standard of 2017. The strength of the study was the systematic quantitative analysis, using analytical software and a full text book corpus to produce objective and repeatable results. The results indicated that a large number of these important lemmas were present in the textbooks, however some were missing, especially the vocabulary that is needed in competitive exams. Over half of the words needed were used less than 5 times, which was deemed not sufficient for effective acquisition.

Dispersion analysis revealed a reasonable evenness of distribution of the frequently repeated words, although low repetition was still a problem, which might have facilitated retention. The following limitations were identified for the study: No student data was gathered from classrooms, and it was assumed that textbook content was equal to actual exposure. In spite of this, the study offered some information for language teaching to teachers and curriculum planners.

One of the studies measured the English textbooks in the Turkish high schools and compared them with university admission examinations using the corpus-based technologies. Results indicated that the exam items had more complex words and sentences and a greater number of words than the textbooks. This showed that there was a mismatch between the pupils' learning and what they were tested on. A key strength of the study was the use of reliable linguistic analytical tools, as well as an objective and data-driven approach, which lent credibility and reproducibility to the findings. However, it is mainly focused on the analysis of the text, and does not consider the teacher perspective, student differences, or classroom teaching practices. It was suggested from the study that curriculum examination link should be improved (Gedik & Kolsal, 2002). Likewise, Bergström et al. (2025) conducted a significant study on the use of English textbooks as tools for vocabulary development and not only as teaching materials. Students' EFL textbooks contained sufficient unknown and mid-frequency lexical items, but the students' analysis of the texts revealed that these words were not adequately reused and recycled for good incidental acquisition. This indicated that students were not likely to re-encounter the words often enough in reading. The discovery contradicted the thought that the use of textbooks was an automatic process of acquiring vocabulary. The study was however limited to investigating the texts themselves and neglected to take into account what teachers might have done to expand or adapt the texts in classrooms or how they could have supplemented them with other materials, which could have influenced the actual learning opportunities. Even so, the paper offered a helpful consideration of the importance of textbook design transcending the shallow approach of vocabulary selection to meaningful lexical recycling.

The overall findings revealed that the previous studies indicated that the school-level English textbooks generally exhibited irregular language patterns, which include unequal distribution of vocabulary, less repetition of important vocabulary and the absence of natural language forms. In addition to this, a number of researchers highlighted the importance of the use of corpus-based analysis in the evaluation of textbook language as it might not have been possible to identify the actual linguistic patterns from the conventional evaluation techniques. Moreover, teachers pointed out the discrepancy between lesson plans and textbooks. Based on this, the present study combined teachers' pedagogical viewpoints with corpus analysis to investigate linguistic patterns in the English textbook used in Grade 7 and gain insight into how well the textbook met students' language-learning and teaching requirements.

Table 1.1: Review of the previous studies

No	Authors	Year	Topic	Findings	Gaps (if any)
1	Le Foll, E	2002	Corpus based analysis of the language used in secondary school EFL textbook in France, Germany, and Spain;	Authentic real-world English can be very different from textbook English. Talk in some textbooks is relatively unnatural and less realistic	There can be significant variation between textbook English and the English used in everyday situations.

2	Maufiroh, Hussen and femilia	2024	Lexical density analysis of reading text in passport to the world English textbook (8th grade)	There was medium lexical density in most passages, and a few passages had low or high lexical density, which was not balanced in text complexity.	It also offers a partial measure of readability other than lexical density and fails to explore the effect of textbook language on the learning outcomes of students.
3	MR, E.R.	2016	Teacher's perceptions toward the use of English textbooks	Textbooks are valuable tools for teachers for structured lesson planning and alignment with the curriculum. The teachers modify, add or remove information from the textbook to align with the needs of the students.	The number of samples is not large enough to ensure that the results can be generalized. Teacher-centred approach as opposed to pupil outcomes.
4	saif et al	2025	Evaluating Pakistani English textbooks through corpus-based frequency and collocation analysis	Limited use of high frequency words in textbooks, minimizing exposure to authentic words. A list of words used repeatedly for teaching reasons.	Only attend to lexical and collocational characteristics, not Grammer and discourse.
5	Siddique et al	2025	Evaluation of content complexity in PTB English textbooks for intermediate learners.	Progresses of difficulty vary across textbooks. There are some units that are complex and some that are simpler.	Only uses intermediate level; not generalizable across grade levels. Does not discuss classroom mediation (teacher adaptations to texts).
6	Yu and Renandya	2021	Corpus-based evaluation of vocabulary profile in New Senior English for China textbooks based on the 2017 English Curriculum standard.	Required lemmas are covered in textbooks at high level (approximately 89%) and there is generally good match with required curriculum standards.	Study is done without any learner performance data to confirm actual vocabulary learning.

7	Gedik and Kolsal	2022	Corpus-based analysis of Turkish high school English textbook vs. university entrance exams	Vocabulary on exams is more advanced than vocabulary seen in the textbook. Evidence of a poor match between instruction and assessment.	Carries out textual analysis primarily without paying attention to classroom teaching methods or differences among learners.
8	Bergstrom, D., Norberg, C., & Nordlund, M.	2022	Support for incidental vocabulary learning in Swedish EFL textbooks	The amount of unknown and mid-frequency (MF) vocabulary is adequate for learners in five EFL textbooks series in Swedish at intermediate level.	Give attention to the text and no other components of materials that might aid learning (e.g. exercises, word lists).
9	Jamal, Shafqat and Afzal	2021	Teachers' perceptions of incorporating a corpus-based approach in English language teaching classrooms in Karachi, Pakistan	The overwhelmingly positive responses of the teachers regarding the corpus-based approach were due to its ability to develop their students' vocabulary, grammar, reading, and writing skills through exposure to real language use.	The major weakness of the study is that the sample is relatively small and homogeneous, and that the students' perceptions were obtained through questionnaires rather than through observation of their classroom behaviors and performance.

1.3 Statement of the problem

The language instruction in public schools in Pakistan is generally taught from English textbooks to the 7th grade students. As a result, it's critical to make sure that the language used in the English textbook for seventh graders meets the linguistic needs and competence levels of the students. It is worth noting, however, that the grade 7 English reading book's vocabulary profile and collocational patterns have not been measured in any previous corpus studies. Furthermore, corpus analysis has not been studied together with the teachers' opinion towards the textbook. This means that there is limited evidence of the appropriateness of the linguistic input provided in the textbook for Grade 7 pupils. Corpus-based technologies will, therefore, be employed in this study to examine the word profile and the collocational patterns of the Grade 7 English textbook and to examine the opinions of Grade 7 English teachers on the textbook to suggest modifications.

1.4 Research Questions:

1. What are Grade 7 English teachers' perceptions of the linguistic needs, vocabulary load, and collocational patterns in the Grade 7 English textbook?
2. To what extent do the vocabulary and collocational patterns in the Grade 7 English

textbook align with the linguistic needs and proficiency levels of Grade 7 learners?
3. How can a corpus-based pedagogical module be designed based on the findings of the needs analysis and corpus analysis for the Grade 7 English textbook?
4. What is the effectiveness of the corpus-based pedagogical module in enhancing vocabulary instruction and teaching practices for the Grade 7 English textbook?
5. How do Grade 7 English teachers perceive the utilization and practicality of the corpus-based pedagogical module in teaching the Grade 7 English textbook?

1.5 Significance of the problem:

This study is important because it succeeded in filling some of the gaps in the English language learning process in middle school level. It did so by using the corpus approach and teachers' perspectives. This included an examination of the lexical coverage and collocational pattern in the prescribed English book for Grade 7 compulsory learners, emphasizing the relationship between the lexical features and the collocational pattern with the communicative needs and proficiency level of the learners. The research was focused on the analysis of the vocabulary frequency, authenticity of collocations and the development of the lexical complexity by applying corpus-based analysis tools. It also included feedback from teachers in practice, teaching Grade 7, to gain a better understanding and appreciation of the textbook from a classroom point of view. The main literary texts used in the Textbook constituted the source of language samples, which are naturally occurring texts. The study was beneficial in a number of areas. In the educational field, it may be used to improve textbooks and design lessons by identifying gaps and providing evidence-based suggestions for improvement. As a research tool, it is a creative approach which combines a corpus-based approach with teachers' feedback in the framework of data-driven learning. At the educational level, it provided useful recommendations to the educational curriculum designer who needed educational materials that could better facilitate language learning.

2. Literature Review

2.1 Corpus based pedagogy

A lot of authors emphasized the need to create a base to build corpus language materials with real language usage. The materials used in ELT classrooms as suggested by Aston (2001), Sinclair (2004), Bennet (2010) and Reppen (2010) should be encouraged for teachers and students to use the corpus resources optimally. The issue of low performance of our students is because of the inefficiency of the teaching method (Haider, 2012) and at intermediate level of college students in public colleges of Pakistan it was observed that it is difficult for students to differentiate between the key word classes (nouns, verbs) in the language. The study at Lasbela University reveals that many students are worried about the chances of learning English, since they had no experience to learn English before (Gopang et al., 2017). The situation of education in Pakistan as revealed by the research indicates that there is a great need to improve education in the country. The present theory and practice of the second and foreign language learning model is old and ineffective. This ineffective teaching method is a cause of students' poor performance (Haider, 2012). The biggest challenge for the English learners in Pakistan is their thinking process according Farooq et al. (2012). These will normally express themselves in their native tongue and then translate to English. The following are some common problems that were found in this study: Interference of L1 and limited vocabulary. These problems are primarily employed by a great number of educators nationwide. Hence, it is necessary to adopt modern methods in the classrooms of Pakistan and one such method is the Corpus-based approach. Most of the research conducted in the field of corpus linguistics in Pakistan are corpus based (Mahmood, 2009). Make decisions on materials (software and tools) with regard to the corpus-based approach, the teachers need to be well informed about the approach. The corpus-based approach for ELT, on the other hand, has

received very little research on teachers' perspectives of the approach in Pakistan. AntConc study analysed the two corpora from a syntactic perspective. To my knowledge there have been many studies on syntactic complexity and L2 writing (Kyle, 2022; Kyle and Crossley, 2018), but few studies that compare the syntactic complexity levels of exams and textbooks (Mirshojaee & Sahragard, 2015). However, these results, which are based on textbooks and exams, show that there is no correlation between the two corpora, and confirm that the aforementioned "skewed interpretation of testing and assessment is also a common problem in other parts of the world.

2.2 Textbook Description of Grade 7

The new English curriculum for Secondary schools aims to provide a rich, engaging and enjoyable environment to enable learners of English to become independent, fluent and effective users of English. The curriculum suggests the principles of language and teaching in English as an international language, the principles of communicative competency, the principles of four primary language skills and the principles of regular teaching, instead of one method. These claims of an improvement of the educational model in the textbooks ought to be meaningful in an EFL context; as one of the most widely used EFL textbooks, it is imperative that these claims become meaningful in it (Allen, 2008). The presence of textbooks in the EFL classroom makes it imperative to discuss the issues and analysis of EFL programs' success (Choi 2008). Textbooks can be viewed as a 'route map' of the information in any English language teaching (ELT) Programme, as well as the Programme structure and aim. This is also true of governmentally ordered books and books which have limited opportunities for change (Sheldon, 1998). In EFL, and perhaps in other language learning contexts, it is common practice to produce a textbook that aims to prepare learners for a standardized test. Textbooks in various parts of the world are written with the intention of preparing learners for standardized tests; this might be a criticism of EFL culture, but textbooks have to be written for that purpose. Many countries have various university entrance tests that have multiple-choice questions. Also, the alignment to textbooks and exams for entry into university in other countries appears to be lacking. In Bangladesh Nur and Islam (2018) did a study and found that the gap exists between the intended and observed pattern in the implementation of the policy directions of English assessment in Bangladesh. The analysis of data also found that a backwash of these "disconnect between policy and practice is a significant disturbance of the overall quality of secondary English education". In the present study, the difficulty level of the textbooks and the word complexity were similar to that of Underwood (2010) in Japan, but the reading difficulty and word complexity were more oriented towards the exams over the years as was the case in Underwood's study. It does, however, note that there are still possibilities for additional development for lexical overlap between texts analyzed.

2.3 Review of the empirical studies:

As for teachers the researcher in the study by Crosthwaite et al., (2021) focused on the teachers' technological, pedagogical and content knowledge that is developed during the process of applying the implementation of corpus-based tools in language instruction. This research was carried out with the pre-service secondary school teachers who underwent intensive training in DDL and were motivated to integrate the activity of using a corpus in their EFL lesson plans. The researchers employed a qualitative method to analyze the teachers' lesson plans and determine the level of appropriateness of using corpus-based activities that are integrated with lesson objectives, target language, and teaching and learning activities. Findings showed that teachers could embed corpus-based activities into their lesson plans but the extent and appropriateness of TPACK was not the same across the lesson plans. In order to use the corpus tools effectively in the classroom, therefore, systematic training of teachers and pedagogical guidance are required. The study offers a valuable linkage between

corpus technology and teacher professional development because it shows that the knowledge of corpus tools is not enough also, the teacher should be aware of the activities that can be carried out using a corpus and how these activities relate to the learning goal and educational process at the classroom level. But the study has a few limitations with reference to the context of the present study. The first one was geared towards pre-service teachers in Indonesia but the present study deals with in-service public-school Grade 7 English teachers in Lodhran District, Pakistan. Third, the study focused mainly on the teachers' lesson planning and the development of TPACK for DDL and did not deal with the linguistic features of a particular school book by corpus analysis. It therefore does not directly address vocabulary load, lexical frequency, lexical patterns, lexical complexity, or appropriateness of content based on proficiency level in textbooks. The study, therefore, suggests teacher-oriented corpus pedagogy and also uncovers the gap between the incorporation of the textbook corpus findings, teachers' needs and context specific pedagogical development in the context of Grade 7 in Pakistan.

Likewise, Schmidt (2023) investigated the process of in-service English writing teachers acquiring knowledge they needed for using corpus-based pedagogy in their classrooms. The study was carried out via a seven-week online corpus-based pedagogy workshop featuring six second language writing teachers at the university level. The teachers created and executed activities in the workshop and used reflective diaries and revision checklists for reflection. They used a qualitative case study with analysis based on the TPACK framework to investigate the process of knowledge integration from the viewpoint of the teachers.

Corpus linguistic to analyze the lecturers' own discourse and then participated in a joint reflection based on the results. The study employed a Collaborative Micro Auto-Ethnography (CMAE) methodology for examining how micro auto-ethnography (MAE) of the type the study introduced might help teachers become attuned to various aspects of their language use and teaching. The results showed that collaborative reflection with corpus-based analysis stimulated more in-depth reflection on teacher speech in relation to the teachers' behavior and student interaction. Strengths of this article is the innovative method of corpus analysis and collaborative reflection of teachers. The study showed that corpus tools were not just for the identification of particular words or grammatical structures, but could also be used to explore teachers' discourse more generally, and to prompt pedagogical reflection. When viewed in the context of the present research, however, the study is also limited. It was not about school teachers' perceptions or the analysis of a set school textbook, but about their own university lectures and reflections. However, the study is useful to the current research as a proof of how corpus analysis can be used as evidence for fruitful pedagogical reflection and professional development. This provided a corpus-based analysis to look back at teacher talk while the present study involved corpus-based analysis of the Grade 7 English textbook and gave synoptic analysis by integrating the findings of the corpus-based analysis with teachers' perceptions and identified linguistic needs.

Murphy (2015) explored critical reflective practice of early-career teachers in two different contexts in which students were being taught, and investigate the nature of teachers' reflections of their professional experiences and how cultural, social and institutional dimensions shaped and influenced the reflections. To this end, he analysed two small written sub-corpora from the more extensive Corpus of Reflective Practice (CoRP), which has 200,000 words. The researcher employed corpus-based techniques and tools to analyze the patterns of teachers' reflective writing and examined the patterns that were related to their understanding of professional reflection. The results indicated that there was a variance in teachers' conception and documentation of reflective practice in the two contexts. This study underscored the need to take into account the differences of context and culture in the process of designing teacher training programs. The study also argued that

reflective practice should be seen as a learning trajectory and that teachers need suitable guidance based upon their own professional needs. One important advantage of this study is the application of corpus-based approaches for a systematic study of the teachers' reflective writing. The study instead investigated linguistic patterns rather than simply relying on general descriptions of teachers' experiences in reflecting on those experiences in their authentic writing. The study has however its own drawbacks.

Table 2. Review of empirical studies

Author(s) & Year	Title	Purpose of Study	Methodology & Sample	Key Findings	Pedagogical Implications
Crosthwaite, Luciana & Wijaya (2021)	Exploring Language Teachers' Lesson Planning for Corpus-Based Language Teaching	To examine how teachers design corpus-based lessons	Qualitative analysis of lesson plans	Teachers improved their technological, pedagogical, and content knowledge (TPACK) when planning corpus-based lessons.	Corpus-based instruction should be incorporated into teacher education curricula.
Schmidt (2022)	Unpacking Second Language Writing Teacher Knowledge through Corpus-Based Pedagogy Training	To investigate teacher learning through corpus-based writing pedagogy	Teacher training intervention; L2 writing teachers	Corpus training improved teachers' awareness of language patterns and instructional strategies.	Corpus tools can enhance writing instruction and teacher professional development.
Sert, Sahlén & Schröter (2025)	Corpus-Based Reflective Practice for Professional Development	To explore corpus-based reflective practices among teachers	Collaborative micro auto-ethnography; teacher reflections analyzed through corpus tools	Corpus analysis promoted deeper reflection on teacher talk, engagement, and instructional practices.	Corpus-informed reflection can be used as a professional development strategy.
Murphy (2015)	A Corpus-Based Investigation of Critical Reflective Practice	To analyze reflective practices of novice teachers using corpus methods	Corpus analysis of reflective writings	Distinct patterns of reflection emerged across teaching contexts.	Corpus analysis can help teacher educators understand reflective thinking.

3. Research Methodology

3.1 Research Design

This study was a mixed-methods research design, where the trends of the language of the Grade 7 English textbook were explored, and teachers' pedagogical reflections were made. The first phase of the study was related to teachers' perceptions and linguistic needs that were

assessed on a questionnaire which was distributed to 120 grade 7 English teachers across public schools of the city and villages of Lodhran District. This questionnaire will investigate teachers' attitudes toward the textbook, vocabulary load, collocational patterns, teaching practices, and difficulties in classroom teaching. The collected data was analysed through descriptive statistics to determine teachers' needs and perceptions. The second phase was involved corpus-based analysis of the complete Grade 7 English textbook. The textbook will be digitized to create a digital corpus and use Voyant Tools, AntConc and Text Inspector to analyze the text. Lexical frequency, level of vocabulary, collocation, lexical complexity and other linguistic aspects were analyzed. Some of the results were useful in assessing the appropriateness of the vocabulary and linguistic patterns that are used in the textbook for Grade 7 learners. The third phase aimed at the creation of a corpus-based pedagogical module based on the results obtained in the needs analysis and the textbook corpus analysis.

3.2 Population of the study

The population of the study was described based on the three phases of the study. For the first phase the population consisted of male and female Grade 7 English teachers in public schools of urban and rural areas of Lodhran District, Punjab. These teachers were given their perceptions about the language needs of grade 7 students, vocabulary load, collocational patterns, the quality of the prescribed textbook, and difficulties in teaching the prescribed textbook. For the second phase the population consisted of the prescribed English book for Grade 7th from PECTA, Punjab. The whole book, with its reading passages, lessons, poems and pertinent textual content, was explored and analyzed as a digital corpus to find out how it can be described in terms of language and lexis. For the third phase, selected male and female grade 7 English Teachers of public schools in the Lodhran district was selected.

3.3 Sample of the Study

The first phase sample comprised 120-Lodhran District grade 7 English teachers from public schools. A sample of 60 male and 60 female teachers, who were teaching the prescribed Grade 7 English textbook, will be included in the sample. The teachers were given information about their perceptions of the textbook, the amount of vocabulary in it, the collocational patterns, linguistic needs, and the challenges they face with the textbook. The book of grade 7 English prescribed by PECTA (Punjab) was used as a sample for the second phase. The entire textbook was incorporated into the corpus-based analysis.

Table 3. Description of the Book

No	Chapter	Page	Themes	Genre
1	The last sermon of Rasoolullah	1	Ethics and values	Religious
2	Martyrs of Pakistan	12	Role models	Biographical\ Informational
3	A Nation's strength(poem)	22	Patriotism\ National pride	Poetry
4	Quaid-e-Azam	32	Self-people and places	Biographical\ Informational
5	Smog	41	Environmental education	Expository\ informational
6	Pakistani Female Mountaineer	50	Gender equality and equity	Biographical \ Narrative
7	Eid-ul-Adha	62	Festivals and cultural events	Descriptive\informational
8	The wise Old Woman	72	Peaceful co-existence\ peace education	Narrative story
9	Think not all is	84	Personality development	poetry

	over(poem)			
10	The young Boy's Adventure	94	Adventure	Narrative story
11	Pakistani culture	103	National pride	Descriptive\cultural essays
12	Role of Media	118	Media	Descriptive

A total of 40 selected Grade 7 English teachers from public schools in Lodhran District were selected as the sample for 3rd phase. The sample was included 20 male and 20 female teachers. These teachers participated in consultative workshops to get feedback and suggestions about the pedagogical module based on the corpus and its relevance, usefulness and practicability for teaching and learning of vocabulary.

Table 4. Consultative workshop

Phase	Males	Females	Total
One	60	60	120
Third	20	20	40

3.4 Data Collection tools

In the present study, various data collection instruments were used in the three phases of the study. The data was gathered and analyzed using these tools to investigate the language features of the English textbook for Grade 7 and the language learners' teachers' needs. In the first phase, a structured questioning was used to the Grade 7 English teachers to investigate their perceptions on the prescribed textbook, vocabulary load, collocational patterns, language needs, teaching practices, and challenges in the classroom. During the second phase, the prescribed Grade 7 English textbook by the Punjab Education, Curriculum, Training and Assessment Authority (PECTA), Punjab has been digitized and subsequently analyzed linguistically. The word frequency, vocabulary distribution, collocational patterns, vocabulary levels and other language features were examined using corpus-based computational tools: AntConc, Voyant Tools, and Text Inspector. The third phase involved the development of a corpus-based pedagogical module based on the results that have been obtained from the questionnaire and corpus analysis. Selected teachers participated in the consultative workshops with the module and provided their feedback by using a structured checklist. These tools were in totality used to link the language analysis of the textbook to the teachers' needs and classroom vocabulary teaching.

Table 5. Questionnaire section

Section	Questionnaire Component	Response Scale
A	General Assessment of English Textbook	5 Point Likert Scale (Agreement) (SA=Strongly Agree, A=Agree, N= Neutral, D= Disagree, SD= Strongly Disagree)
B	Teacher's Perceptions towards Textbook Chapters	Usefulness scale: Useful (U), Very Useful (VU), Moderately Useful (MU), Slightly Useful (SU), Not Useful (NU)
C	Teaching Prospective	5 Point Likert Scale (Agreement) (SA, A, N, D, DA)
D	Awareness regarding Corpus-based Pedagogy	Awareness Scale (Highly Aware, Aware, Moderately Aware, Slightly Aware, Not Aware)

E	Teachers prospective and Corpus-based pedagogy	5 Point Likert Scale (Agreement (SA=Strongly Agree, A=Agree, N= Neutral, D= Disagree, SD= Strongly Disagree)
F	Difficulties and Challenges	Difficulty Scale VD=very difficult, d= difficult, MD= Moderately Difficult, SD= Slightly Difficult, ND= Not Difficult)

Google forms were used to create the questionnaire from which data were gathered from grade 7 English teachers online. Data gathered were analysed using descriptive statistics such as frequencies, percentages, means and standard deviations (where appropriate).
A link to the questionnaire is provided below.
https://docs.google.com/forms/d/e/1FAIpQLScdcBq4Woc6JvDTzl_aiDaV7gA9xPRcqoV77EaudRYynfZEPw/viewform

4 Data Analysis and Discussion

4.1 Findings from questionnaire

The overall analysis of the questionnaire revealed that teachers held varied perceptions regarding the linguistic suitability, vocabulary content, and pedagogical effectiveness of the Grade 7 English textbook. Their responses suggested that the textbook provided useful vocabulary and supported reading and writing skills to some extent; however, it did not fully address the linguistic needs and proficiency levels of Grade 7 learners. Teachers appreciated the vocabulary presented in several chapters, particularly its usefulness for vocabulary development, while some chapters were considered less effective in meeting learners' vocabulary requirements. This indicates that the selection and distribution of vocabulary across the textbook need further consideration. The findings also highlighted several teaching-related challenges. Teachers reported difficulties in explaining unfamiliar vocabulary, teaching collocations, and facilitating the understanding of reading passages. Similarly, students faced problems in understanding, acquiring, and retaining new vocabulary, particularly when the language used in textbook lessons was beyond their level of comprehension. These difficulties suggest a gap between the linguistic content of the textbook and the practical needs of both teachers and learners. The findings further indicate that the textbook language is not always sufficiently accessible for effective classroom instruction. Overall, the questionnaire findings suggest that although the Grade 7 English textbook contains useful linguistic material, certain aspects require improvement to ensure better alignment with learners' vocabulary needs and proficiency levels. The difficulties reported by teachers and students emphasize the need for more effective vocabulary instruction, clearer language input, and practical teaching strategies. Furthermore, teachers' positive attitudes towards corpus-based pedagogy provide a useful foundation for developing a corpus-based pedagogical module. Such a module could support teachers in selecting appropriate vocabulary, explaining words through contextual examples, teaching collocations, and improving vocabulary instruction and reading and writing practices.

4.2 Findings from corpus-based analysis of the book

The overall corpus-based analysis of the 7th grade English textbook reveals a considerable range of linguistic features, including grammatical categories, vocabulary distribution, lexical complexity, word frequency, contextual usage, and lexical density. The POS tagging analysis indicates that nouns and verbs are the most dominant grammatical categories, followed by prepositions, whereas pronouns, adjectives, adverbs, and conjunctions occur comparatively less frequently. This distribution suggests that the textbook primarily emphasizes content-based vocabulary, concepts, actions, and narrative descriptions. The variety of lexical items also reflects the inclusion of religious, cultural, social, and moral themes relevant to the Pakistani

context. The CEFR analysis indicates that the overall linguistic level of the textbook corresponds to C1, representing an advanced level of English proficiency. This finding suggests a considerable gap between the linguistic complexity of the textbook and the expected proficiency level of Grade 7 learners. The presence of lengthy and complex sentences, advanced vocabulary, and academically oriented lexical items may create comprehension difficulties, particularly for learners with limited English proficiency. Although such linguistic features can expose students to a wide range of vocabulary and reading structures, they require appropriate instructional support. Furthermore, the word cloud and frequency analysis demonstrate that the textbook contains a substantial number of frequently occurring functional and content words. Words such as Pakistan, people, village, Muhammad, mother, Eid, and Media reflect the cultural, religious, and social orientation of the textbook. The frequency and trend analysis also reveals variations in vocabulary distribution across different sections of the book, with greater lexical concentration in the middle sections. This variation highlights the importance of chapter-wise vocabulary preparation and contextual reading activities.

The KWIC analysis further demonstrates that vocabulary is presented within meaningful linguistic and thematic contexts. Keywords associated with religious practices, cultural traditions, moral lessons, and social awareness provide opportunities for contextual vocabulary learning. Similarly, the lexical density analysis indicates a moderate level of vocabulary richness, suggesting a balance between the introduction of new lexical items and the repetition of familiar vocabulary. This repetition can support vocabulary retention, reading fluency, and comprehension among learners. From a pedagogical perspective, the combined findings suggest that the textbook offers considerable lexical variety and meaningful thematic content; however, its advanced linguistic complexity may not fully correspond to the proficiency needs of Grade 7 learners. The dominance of nouns and verbs, variations in word frequency, contextual use of vocabulary, and moderate lexical density provide useful foundations for developing vocabulary-focused teaching activities. Nevertheless, teachers need to employ scaffolding strategies, pre-teach unfamiliar vocabulary, explain complex sentence structures, and use contextual examples to facilitate comprehension. Corpus-based activities, including word identification, collocation practice, KWIC-based exercises, and chapter-wise vocabulary lists, can further support reading and writing development. Overall, the corpus-based investigation indicates that the 7th grade English textbook is linguistically rich and thematically diverse, but its advanced CEFR classification raises questions about its suitability for the expected proficiency level of middle-school learners. The findings emphasize the importance of aligning textbook vocabulary and linguistic complexity with learners' actual needs and proficiency levels. They also demonstrate the potential of corpus-based analysis as a valuable resource for identifying linguistic challenges and developing more appropriate, evidence-based pedagogical practices for English language teaching.

4.3 Findings from corpus-based analysis of the book:

The consultative workshop was held to gather teachers' comments on the proposed pedagogical module on the use of a corpus in vocabulary instruction in grade 7 English. In this regard, a checklist was filled out by 40 selected public-school grade 7 English teachers among which 20 teachers were males and 20 teachers were females. The checklist emphasized how useful the module was, how relevant, practical, and how applicable it was to classroom instruction. Teachers' responses and suggestions were collectively analyzed to determine the strengths and areas for improvement in the module. The positive feedback from the teachers' checklist indicated that the teachers had positive perceptions of the corpus-based pedagogical module and how it may be able to contribute to vocabulary instruction and teaching practices for the Grade 7 English textbook. The feedback revealed that the module

could offer a more structured and evidence-based vocabulary teaching strategy, assisting teachers to identify high frequency words, to explore collocational structures, and to determine words relevant to learners' linguistic needs and proficiency levels.

Findings also indicated that corpus-based results can help teachers make decisions to select the best words to teach, prioritize valuable words, and offer learners the contexts to understand the meaning of the word rather than just the individual word or textbook explanation. The answers also emphasized the need to form a close relationship between goals of the English textbook for grade 7, learners' language needs and teacher's methods of teaching. The module provides an opportunity for teachers to review the vocabulary that is embedded in the lessons in a textbook and ask themselves whether the vocabulary is adequate to meet the learning outcomes, and whether classroom activities offer opportunities to build vocabulary. Information on word frequency, vocabulary distribution and collocation can be used to determine the significance of key vocabulary items, which can then guide instructional practices more closely to the learning goals. This alignment can lead to more meaningful vocabulary instruction and support teachers to address learners' challenges in vocabulary comprehension and usage of English vocabulary. In addition, the module will help teachers to facilitate the learning of vocabulary to improve students' reading and writing skills. Using corpus evidence, teachers will be able to show students how words are used in context, show them that words have a habit of being combined together in certain ways, and offer practice using them appropriately. These high frequency words can be covered during reading sessions and collocational patterns can be integrated in writing tasks to give students an idea of how words are used in a natural way. These activities can enhance students' chances of reading from textbooks, increasing their vocabulary, and using new words in appropriate contexts. This means that the module can be adapted to link vocabulary learning to the overall language requirements of the textbook, while keeping the focus on vocabulary, reading and writing.

The suggestions provided by the teachers are another source of information used for improving the module and its application in the class. They can provide input on the need for further clarification, more appropriate and specific vocabulary examples, potentially restructured activities, and further integration with the text learning objectives. The suggestions may also be used to help adapt the module to the varying levels of learner ability and to the different needs of the classroom. Teachers are not only users of the module, but their own experiences in teaching can help make the content of the module more relevant, accessible, and practical. Taking their suggestions into account can facilitate the module to improve its adaptability to the linguistic features of the textbook and the difficulties that teachers may meet in teaching vocabulary. The module's successful use is also an aspect of teacher training and professional guidance. While the corpus can be a useful resource for vocabulary work, teachers might require assistance in knowing how to interpret word-frequency data, understand collocational patterns, use the corpus tool, and translate the linguistic results into useful classroom activities. The training can assist teachers to build up their confidence and practice in using corpus-based evidence for vocabulary teaching and lesson planning. It can also enable them to know how to apply the module to assist in the objectives of the textbook and meet students' language needs. If then, teachers provide feedback on training and guidance, the feedback can be used in developing practical support for effective classroom implementation, as long as the latter is represented in the teachers' actual checklist. Practicality is also related to the module's applicability to daily teaching needs.

A module that provides explicit vocabulary foci, examples of these words in context, and activities that align with the module goals is likely to be more easily integrated into teachers' existing lessons. It can be made more effective when teachers know how to use the information provided by it and do not regard it as an extra

and separate piece of work. Better implementation on the other hand, could rely on teachers' knowledge of corpus tools, resources, time to prepare, and the appropriateness of the suggested activities for their students. These practical considerations can be identified through teachers' observations and suggestions, which can be used to further improve the module. In general, the results of the mostly positive checklist responses indicated that teachers saw the corpus-based pedagogical module as a valuable tool to help improve vocabulary teaching and to aid teachers' instructional processes in the English textbook used by Grade 7 students. Its usefulness may be related to supporting the teacher to select vocabulary, high-frequency words and to consider collocations and classroom activities in evidence-informed decision making. Its application and usefulness demonstrates in how it can be used to support class planning, tie in the book learning objectives to learners' linguistic needs, and serve as a yardstick for vocabulary instruction. Teachers' suggestions also help to highlight how the module can be developed further and what forms of professional support can help implement the module.

5. Justification of the Research Questions:

5.1 What are Grade 7 English teachers' perceptions of the linguistic needs, vocabulary load, and collocational patterns in the Grade 7 English textbook?

The question was answered in four relevant parts of the questionnaire: General Assessment of the Textbook, Teachers' Perceptions towards Textbook Chapters, Teaching Perspective and Difficulties and Challenges. The results of the General Assessment of the Textbook show that teachers' perceptions of the overall suitability of the Grade 7 English textbook were mixed. While certain aspects of vocabulary and skills development were seen favorably, concerns were raised about the textbook's role in overall English language development and language level of the textbook in the classroom. The section on Teachers' Perceptions towards Textbook Chapters also showed differences in teachers' perception of the usefulness of the vocabulary items in the various chapters of the text. There were some chapters that were found to be more useful in relation to the vocabulary and others that were found to be relatively less useful, indicating that the vocabulary load and coverage was not found to be equally suitable across the chapters. The Teaching Perspective revealed that teachers had problems when teaching various aspects of textbook language especially when introducing new words, reading texts, teaching collocations, and dealing with difficult language. The results show that the linguistic material in the textbook is not always sufficient to support the needs of teachers in the classroom.

The Difficulties and Challenges revealed similar difficulties for teachers and students with book vocabulary and language. Students were also reported to need more assistance with some of the language in some lessons, and with learning and retaining new words, and teachers experienced some difficulties in explaining difficult words. Results from the four sections indicate that while the vocabulary and language-learning tasks in the Grade 7 English textbook are helpful, the appropriateness of the textbook does not hold true for all the chapters and some vocabulary, linguistic and collocational features are problematic for English classroom learning and teaching. Hence the results of the questionnaire have shown the need to pay more attention to the linguistic needs of the learners and the use of suitable load of words and the effective treatment of collocations in the English textbook for Grade 7, which are the empirical foundations for solving the problem of this study and for developing the pedagogical module to develop the English textbook for Grade 7. Gurmani et al. (2021) made a comparison of the vocabulary used in secondary school English books of both countries, Pakistan and Saudi Arabia. They reported differences in the vocabulary of the textbooks and concerned about the vocabulary appropriateness and coverage for the students at the relevant education level. This helps to validate the current study's emphasis on teachers' perceptions of the suitability of vocabulary, the vocabulary load and the need for vocabulary in the textbook for Grade 7 English.

Chegeni (2016) stated that the choice and evaluation of the textbook is significant for discovering the strengths and weaknesses of the language teaching textbook. The study also pointed to vocabulary as an important factor to be taken into account in designing and evaluating language-teaching materials. It helps support the present study's research on teachers' perspectives on the appropriateness of vocabulary and language in Grade 7 English textbook

5.2 To what extent do the vocabulary and collocational patterns in the Grade 7 English textbook align with the linguistic needs and proficiency levels of Grade 7 leaners?

This study aimed to find out whether the vocabulary and collocational patterns of the textbook matched the linguistic requirements and level of proficiency of the Grade 7 students. The corpus analysis thus enabled a more objective analysis of the content of the textbook, whereas the teachers' perceptions gained from the questionnaire gave information about the problems and needs the teachers had noticed in the teaching and learning process in the classroom. However, the analysis of the total set of words in the entire textbook revealed that the words were not only available at a single level of proficiency and/or frequency, but at several levels. The analysis of the lexical content showed variation in frequency, complexity, and level of vocabulary used in the textbook. The analysis of the proficiency level was of great value in finding out the level of difficulty of the textbook in terms of lexical level. The results indicated that the text wasn't totally consistent in terms of its proficiency level throughout the material. Rather, there were words that were found at varying levels which was a reflection of the linguistic complexity of the text. The expected variation is significant because Grade 7 learners are expected to work with language that is developmentally and proficiently appropriate. The chapter-wise Corpus analysis also revealed that not all the chapters in the textbook had the same lexical characteristics.

The vocabulary frequency, the level of vocabulary proficiency, the complexity of the vocabulary and the use of vocabulary items were found to be different in the various chapters. Several chapters offered vocabulary that seemed relatively easier and more manageable for learners while some chapters had relatively more difficult vocabulary items. The chapter-wise variation is important because it suggests that the difficulty level of the language of the textbook can only be understood when based on the overall profile. Each chapter may have multiple versions of the vocabulary for students to learn, and may demand a varying level of instructional support. The variation is also in line with the results of the questionnaire in which teachers indicated that the usefulness and suitability of the vocabulary varied from chapter to chapter in the textbooks. Collocational patterns analysis has been used to give another dimension to the evaluation of the textbook. The corpus analysis was not conducted at the level of the single word, but rather at the level of the word in context and also through frequency and KWIC analysis, which revealed frequent word combinations. Findings revealed words occurred in meaningful combinations and patterns within context, which allowed the learners the chance to come across the words in relation to other words. Meanwhile, the teachers' questionnaires showed that teachers faced challenges in the handling and teaching of collocations. This implies that having collocational patterns in the textbook does not guarantee the learner will be able to identify, comprehend and apply them well. Thus, collocations can be better highlighted in the classroom, especially when the collocations are less common or challenging for Grade 7 students. The chapter-level proficiency and lexical analysis also indicated that other measures may yield different levels of difficulty measures for individual chapters. For instance, when looking at A Nation's Strength as a whole, some lexical measures classified it as a C2, while the Flesch-Kincaid grade level classification was an A2, and the average words per sentence was an A2. Academic vocabulary proportion was 0.99% (A1) and B1 vocabulary was 8.11%. The findings in this study support the view that textual difficulty is a multidimensional phenomenon and that one measure of difficulty is insufficient.

The results of the questionnaire revealed that teachers reported problems with the vocabulary, language, reading passages and collocation of the textbooks and students were thought to need extra help in reading and remembering hard words. The corpus analysis also revealed differences between the frequencies of vocabulary, levels of proficiency, complexity of vocabulary and its association with other words across the textbook. The relation between teachers' reported classroom problems and corpus-based features of the textbook indicates that some of the linguistic problems teachers reported are also supported by the lexical profile of the textbook. In general, it can be concluded that the vocabulary and collocational elements in the Grade 7 English textbook are useful and broaden students' vocabulary in terms of exposure to a variety of vocabulary and word combinations. But the language as a whole and the collocations are not consistently congruent with the language requirements and predicted competence of grade 7 students. Some words of the text seem to be suitable and easy to access, while others are more challenging and might need further instructional scaffolding. Thus, the results indicated some degree of alignment but not full alignment between the vocabulary/collocational patterns in the textbook with the needs and proficiency level of the targeted learners. The discovery of this will serve as a solid foundation for the next step in the corpus-based pedagogical module: the identification of words and collocations that need to be explained, contextualized, and put into use in the classroom. Likewise, Webb and Nation (2017) talk about the need to build vocabulary through suitable opportunities and repeated word exposures. The study of collocations, therefore, can be used to help answer the questions of whether learners are exposed to useful word combinations as part of meaningful texts and whether useful word combinations can facilitate vocabulary development and written expression in the Grade 7 textbook.

5.3 How can a corpus-based pedagogical module be designed based on the findings of the needs analysis and corpus analysis for the Grade 7 English textbook?

The needs analysis identified the problems that teachers encountered when they explained new words, taught reading passages, used collocation and when they had to cope with complicated textbook language. The results further showed that students needed more help with vocabulary understanding and retention, and that the usefulness of vocabulary was different for different chapters of the textbook. The findings guided the identification of the key aspects which were to be addressed in the pedagogical intervention of the module. The linguistic evidence used to select and cluster the module content was obtained by undertaking corpus analysis. The vocabulary, the proficiency levels, lexical complexity, collocational patterns, KWIC analysis and lexical density were used to analysed the complete book and the individual chapters. Analysis helped identify frequently used vocabulary items, collocations that were contextually important and/or comparatively difficult for language learners. These were then chosen to be included in the module and examples in KWIC form and context information were used to show where words and collocations were found in the real book. The module was created based on a combination of the results, including classroom activities focusing on vocabulary and collocation that could address the difficulties identified by the teachers and the linguistic patterns found in the corpus. The module was thus systematic in process, starting with the need's identification, followed by the analysis of the corpus, selection of appropriate vocabulary and collocation, designing pedagogical activities and teacher consultation. In general, the module was needs-based, corpus-informed and practically relevant, since the content was extracted from the actual Grade 7 English textbook and the activities were aimed at giving teachers practical help in the teaching of vocabulary and collocational problems. Corpus-based materials have been found to assist

vocabulary learning by providing suitably prepared examples (Liu & Jiang, 2009).

5.4 What is the effectiveness of the corpus-based pedagogical module in enhancing vocabulary instruction and teaching practices for the Grade 7 English textbook?

Selected Grade 7 English teachers participated in a consultative workshop and evaluation of the corpus-based pedagogical module using the checklist to address this question. The needs analysis, the corpus analysis, especially the needs of the students in using words, collocational problems, and differences in the vocabulary features of the textbook were used as sources to develop the module. At the consultive workshop, teachers reviewed the module based on its relevance to their curriculum, its usefulness to their students, and its ability to facilitate vocabulary instruction and classroom teaching practices. The feedback suggested that the module was helpful in relation to difficult word frequency items, in the presentation of words in meaningful contexts and also in the presentation of common collocational patterns. In addition, the corpus-based activities provided a systematic framework for teachers to identify and present vocabulary based on frequency, proficiency level and contextual usage of the vocabulary instead of the word-level approach. Besides, the module was also deemed relevant to the real teaching content of the grade 7 English textbook, leading to a higher practical value for classroom teaching. Therefore, the teachers' responses indicated that the module could support concentrating on the vocabulary and teaching in the module in a more contextualized and focused way, and also provide some practical help in relation to the language related issues that had been identified in the needs analysis. The overall findings of the consultative findings were that the corpus-based pedagogical module is effective, relevant, useful and practically applicable in supporting the vocabulary instruction and teaching practices of the Grade 7 English textbook. In this study, however, the effectiveness that have been reported cannot be interpreted as a measure of students' vocabulary achievement because there was no pre-test/ post-test or experimental comparison was done. Increased awareness of corpus-based approach to language pedagogy and an awareness of the possible contribution of corpus-based approach towards vocabulary instruction of teachers was found by Çalışkan & Kuru Gönen (2018) after training.

5.5 How do Grade 7 English teachers perceive the utilization and practicality of the corpus-based pedagogical module in teaching the Grade 7 English textbook?

The teachers' responses to the consultative workshop checklist indicated that the corpus-based pedagogical module was quite beneficial and practical to use in teaching vocabulary in the Grade 7 English textbook. The checklist was directed towards learning outcomes, clarity of learning outcomes, usefulness of activities, anticipated challenges and suggestions for the management of the module in the classroom. Feedback from teachers showed that there was a need to relate the corpus-based results to learners' vocabulary needs and teachers' day-to-day practices in class. The module can help teachers select key vocabulary, grasp word frequency patterns, and provide collocated words in context. The above-mentioned features can support the teaching syllabus for introducing vocabulary in a more structured and systematic way and link the vocabulary with the assigned textbook lessons. The module is also supported by previous research of its practical value. Jamal et al. (2021) explored the views of the teachers of English on the adoption of corpus-based approach in language teaching classrooms in Karachi city, Pakistan. They showed that teachers were aware of the potential of the corpus-based materials and activities in teaching vocabulary, reading and writing skills in English. The study is meaningful to the current study because it explores teachers' attitudes in the Pakistani educational context and advocates the use of corpus-based materials in the classroom for teaching. In the present research, the findings of their study however, offer supporting evidence from a different context to the practicality of the module developed in the present research.

In the same way, Çalışkan and Kuru Gönen (2018) investigated the teachers' attitudes towards the use of corpus-based materials and concordance lines in vocabulary teaching. They found that teachers saw these materials as useful for learner's awareness of vocabulary. The participants also pointed out issues with technology and the preparation of corpus-based materials, however. This research confirms that the implementation of a corpus-based module is not only based on the quality of the content of the module but also on the teachers' knowledge of the activities that can be carried out with the corpus and the guidance that is given to them. The module may be easier to fit into regular lessons in the Grade 7 English textbook, if the instructions are clear, examples are appropriate, and classroom activities are manageable. Suggestions from the teachers are also significant towards enhancing the applicability of the module in the classroom. Activities should be suitable to grade 7 students, linked to the appropriate unit of the text and suitable for the time available for teaching. The teachers might also require assistance in clarifying the meanings of challenging words, modifying tasks for varying learning styles, and incorporating the information from a corpus in lesson planning. Guidance like providing more examples and practice, flexibility in lesson timing, and some brief teacher guidance, can enhance the usefulness of the module. The recommendations are made as they highlighted teacher preparation in the use of corpus-based materials. Overall, the results indicate that the corpus-based pedagogical module can be used as a potential tool by teachers for incorporating vocabulary frequency, collocation and contextual language use in the teaching of English textbook for Grade 7. There is a need for flexibility in its use if it is to be integrated with the normal curriculum and to fit the classroom environment. Teachers' feedback can thus inform further development of the module and aid in making it more appropriate for teachers' teaching purposes. These findings do not, in themselves, prove the module was effective in enhancing students' vocabulary skills, but they do indicate what teachers thought of the module's usefulness and practicality

6. Conclusion

Based on the current research, it is concluded that corpus-based analysis can contribute to the English language teaching field because it can provide valuable information about the linguistic patterns of the Grade 7 English textbooks and its relevance. The study provided a systematic analysis of the language used in textbook by analyzing the vocabulary distribution, word frequency, collocational patterns, lexical profiles, and lexical complexity. Teachers' perspectives further facilitated the connection of these linguistic features with classroom needs and vocabulary-related classroom practices. The study also shows the possibility of embedding the corpus results in a pedagogical module to aid in vocabulary teaching using context-based examples and meaningful classroom activities. The teachers' suggestions gave rise to the insights into the usefulness and practicability of the module in teaching the Grade 7 English textbook. In conclusion, the study establishes that with the help of teachers' pedagogical reflections, corpus based textbook analysis could contribute to design of teaching materials appropriate for the Pakistani education context. However, additional classroom-based research is needed to prove the effectiveness of the module to increase students' learning outcomes.

References:

Allen, H. W. (2008). Textbook materials and foreign language teaching: Perspectives from the classroom. *The NECTFL Review, 62*, 5–28.

Aston, G. (Ed.). (2001). *Learning with corpora*. Athelstan.

Bennett, G. R. (2010). *Part I: An introduction to corpus linguistics*. University of Michigan.

Bergström, D., Norberg, C., & Nordlund, M. (2022). Do textbooks support incidental vocabulary learning? A corpus-based study of Swedish intermediate EFL materials. *Education Inquiry, 16*(1), 69–87. https://doi.org/10.1080/20004508.2022.2163050

Breyer, Y. (2009). Learning and teaching with corpora: Reflections by student teachers. *Computer Assisted Language Learning, 22*(2), 153–172.

Chambers, A. (2019). Towards the corpus revolution? Bridging the research–practice gap. *Language Teaching, 52*(4), 460–475.

Choi, I. (2008). The impact of EFL testing on EFL education in Korea. *Language Testing, 25*(1), 39–62.

Crosthwaite, P., et al. (2023). Exploring language teachers' lesson planning for corpus-based language teaching: A focus on developing TPACK for corpora and DDL. *Computer Assisted Language Learning, 36*(7), 1392–1420. https://doi.org/10.1080/09588221.2021.1995001

Farooq, M. S., Uzair-ul-Hassan, M., & Wahid, S. (2012). Opinion of second language learners about writing difficulties in English language. *South Asian Studies, 27*(1), 183–194.

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics, 35*(3), 305–327. https://doi.org/10.1093/applin/amt015

Gedik, T. A., & Kolsal, Y. S. (2022). A corpus-based analysis of high school English textbooks and English university entrance exams in Turkey. *Theory and Practice of Second Language Acquisition, 8*(1), 157–176. https://doi.org/10.31261/TAPSLA.9152

Haider, G. (2012). Teaching of writing in Pakistan: A review of major pedagogical trends and issues in teaching of writing. *Journal of Educational and Social Research, 2*(3), 215. https://doi.org/10.5901/jesr.2012.v2n3p215

Hendrigan, H. (2019). Mixing digital humanities and applied science librarianship: Using Voyant Tools to reveal word patterns in faculty research.

Hetenyi, G., Lengyel, A., & Szilasi, M. (2019). Quantitative analysis of qualitative data: Using Voyant Tools to investigate the sales-marketing interface. *Journal of Industrial Engineering and Management, 12*(3).

Jamal, J., Shafqat, A., & Afzal, E. (2021). Teachers' perceptions of incorporation of corpus-based approach in English language teaching classrooms in Karachi, Pakistan. *Liberal Arts and Social Sciences International Journal, 5*(1), 611–629. https://doi.org/10.47264/idea.lassij/5.1.40

Joubert, E. (2021). Distant reading in French music criticism. *Nineteenth-Century Music Review, 19*, 291–315.

Khojasteh, L., & Mukundan, J. (2025). A systematic review of corpus-based methodologies in textbook analyses and evaluation. *Language Teaching Research Quarterly, 47*, 175–195. https://doi.org/10.32038/ltrq.2025.47.10

Kyle, K. (2022). *Measuring syntactic development in L2 writing: Fine-grained indices of syntactic complexity and usage-based indices of syntactic sophistication*. Digital Archive @ GSU. https://doi.org/10.57709/8501051

Kyle, K., & Crossley, S. A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *The Modern Language Journal, 102*(2), 333–349.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English.

Le Foll, E. (2024). Research on the language of school EFL textbooks. In *Textbook English: A multi-dimensional approach* (Studies in Corpus Linguistics, Vol. 116, pp. 25–63). John Benjamins. https://doi.org/10.1075/scl.116.c3

Mahmood, M. A. (2009). *A corpus-based analysis of Pakistani English* Doctoral dissertation, Bahauddin Zakariya University.

Maufiroh, Y. L., Hussen, B. T. E., & Femilia, P. S. (2024). Lexical density of reading materials in an English textbook: A content analysis. *Journal of Language Intelligence and Culture, 6*(2), 165–174. https://doi.org/10.35719/jlic.v6i2.104

Ming, X. (2018). Smart learning models of certified legal translators and interpreters in China. *Comparative Legilinguistics, 36*(1), 47–64.

Mirshojaee, S. B., & Sahragard, R. (2015). Reading comprehension passages of Iranian general English books and MA reading comprehension tests: A corpus analysis. *Journal of Modern Research in English Language Studies, 2*(2), 77–98.

Murphy, B. (2015). A corpus-based investigation of critical reflective practice and context in early career teacher settings. *Classroom Discourse, 6*(2), 107–123. https://doi.org/10.1080/19463014.2015.1030762

Nguyet, H. T. T., & Long, N. V. (2020). Applying corpus linguistics to English textbook evaluation: A case in Viet Nam. *VNU Journal of Foreign Studies, 36*(6), 109–121. https://doi.org/10.25073/2525-2445/vnufs.4632

Nur, S., & Islam, M. (2018). The (dis)connection between secondary English education assessment policy and practice: Insights from Bangladesh.

Reppen, R. (2010). *Using corpora in the language classroom*. Cambridge University Press.

Rosyida, E. (2016). Teachers' perceptions toward the use of English textbook. *English Education: Jurnal Tadris Bahasa Inggris, 9*(1), 43–54. https://doi.org/10.24042/ee-jtbi.v9i1.418

Saif, S., Faisal, U., Tahir, A. S., & Farid, S. (2025). Evaluating Pakistani English textbooks through corpus-based frequency and collocation analysis. *Journal of Applied Linguistics and TESOL, 8*(1). https://doi.org/10.63878/jalt2840

Schmidt, N. (2023). Unpacking second language writing teacher knowledge through corpus-based pedagogy training. *ReCALL, 35*(1), 40–57. https://doi.org/10.1017/S0958344022000106

Sert, O., et al. (2025). Corpus-based reflective practice for professional development: A collaborative micro auto-ethnography. *Education Sciences, 15*(1), Article 79. https://doi.org/10.3390/educsci15010079

Sheldon, L. E. (1988). Evaluating ELT textbooks and materials. *ELT Journal, 42*(4), 237–246.

Sinclair, J. (Ed.). (2004). *How to use corpora in language teaching*. John Benjamins.

Siddique, A., et al. (2025). Reevaluating textbooks organization: A corpus-based analysis of content complexity in PTB English textbooks for intermediate learners. *Journal of Applied Linguistics and TESOL, 8*(2). https://www.researchgate.net/publication/392693889

Tai, S., & Chen, H.-J. (2015). Are teachers test-oriented? A comparative corpus-based analysis of the English entrance exam and junior high school English textbooks. In F. Helm et al. (Eds.), *Critical CALL: Proceedings of the 2015 EUROCALL Conference* (pp. 518–522). Research-publishing.net. https://doi.org/10.14705/rpnet.2015.000386

Underwood, P. (2010). A comparative analysis of MEXT English reading textbooks and Japan's National Center Test. *RELC Journal, 41*(2), 165–182.

Yu, M., & Renandya, W. A. (2021). A corpus-based study of the vocabulary profile of high school English textbooks in China. *LEARN Journal: Language Education and Acquisition Research Network, 14*(1).