Article QR

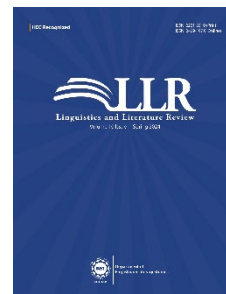|	
Title:	**Lexical Load in Class X English Textbooks: A Corpus-Based Comparative Analysis**
Author (s):	Abdul Farooq Khan
Affiliation (s):	Punjab School Education Department, Pakistan
DOI:	https://doi.org/10.32350/llr.101.07
History:	Received: June 01, 2023, Revised: December 03, 2024, Accepted: March 09, 2024, Published: March 31, 2024
Citation:	Khan, A. F. (2024). Lexical load in class X English textbooks: A corpus-based comparative analysis. *Linguistics and Literature Review*, *10*(1), 122–139. https://doi.org/10.32350/llr.101.07
Copyright:	© The Authors
Licensing:	This article is open access and is distributed under the terms of Creative Commons Attribution 4.0 International License
Conflict of Interest:	Author(s) declared no conflict of interest

Lexical Load in Class X English Textbooks: A Corpus-Based Comparative Analysis

Abdul Farooq Khan[*]

Punjab School Education Department, Pakistan

Abstract

English, in Pakistan, is taught since grade 1, yet the acquisition of receptive and productive skills in English language by Pakistani language learners remains questionable. The lack of vocabulary is a major factor as vocabulary is crucial for communication. In Pakistan, textbooks are the only source of vocabulary exposure for students. Keeping in view the importance of textbooks, this corpus-based research was conducted to know and compare the lexical load of Class X textbooks that are developed, published and distributed in public schools by the four textbook boards. These textbooks were evaluated on the basis of three parameters, namely the frequency of types and tokens present in textbooks and type-token ratio, the percentage of GSL/AWL covered, and the levels of BNC-COCA 25 present. The coverage of GSL/AWL lists for GSL_2nd_1000 and AWL750 was alarming. Most of the vocabulary present in these textbooks habituated first three levels of BNC-COCA 25 but rest of the levels had no or minor presentation. The results showed that these textbooks do not fulfill the needs of students as the amount of vocabulary included does not meet the internationally defined standards for vocabulary. Hence, vocabulary should be extended and distributed as per the standards.

Keywords: BNC-COCA 25, corpus linguistics, lexical load, GSL/AWL lists, textbooks, type-token ratio, vocabulary development

Introduction

In Pakistan, English is taught as per the National Education Policy, 2017 (Government of Pakistan, 2017), as a compulsory subject since grade 1 and it has been given the status of second language in the Single National Curriculum, 2022. Despite its obligation from early years and status of ESL, the majority of Pakistani students lack both productive and receptive skills in English even after the completion of their education with an English major. National Achievement Test resulted in alarming figures for English

---

[*]Corresponding Author: farooqsawansi@gmail.com

language learning and emphasized the necessity of measures to overcome the problem. Exclusive research to find out the causes of the lack in language skills is limited and attributed only to a few factors that is, motivation and its complexity (Shahbaz et al., 2012), ELT practices (Shamim, 2011) and language policies (Mansoor, 2005) and so on.

Multiple factors including teachers' training and qualifications, teachers' attitudes towards language teaching, available facilities, language resources, textbooks, language social status, and many more are responsible for the successful teaching-learning process (Azim et al., 2020). Hsueh-Chao and Nation (2000) placed vocabulary acquisition at the center of language learning and it has been considered synonymous with language acquisition (Lessard-Clouston, 2013). Textbooks, especially in public sector institutions, are the only resource in Pakistan for vocabulary development that are provided free of cost to the learners. These textbooks are produced by the textbook boards at provincial levels in accordance with the defined National Education Policy (NEP). Development of Lexical competence of L2 learners is essentially dependent on textbooks. Textbooks vocabulary and vocabulary activities dimensions have impact on the acquisition and development of language so, the lexical load and its distribution becomes the most important factor in language material development. Many textbooks contain vocabulary that has already been archived or outworn and has no or very little contemporary use. There is no benefit of including such vocabulary in language learning textbooks. Keeping this in view, Azim and Garcia (2020) conducted the research to evaluate the lexical load of Punjab Curriculum and Textbook Board's *English 1* and *English 2*. Kauser et al. (2016) evaluated the *Intermediate English Book-1 (Short Stories)* of the same board. No such research is conducted that evaluates the textbooks published by different textbook boards for vocabulary load. The research is focused on the evaluation of the textbooks published for and recommended to Class X students by provincial boards. The aim of the study was to know the lexical load and its distribution in these books and to compare the present vocabulary to GSL/AWL lists and BNC-COCA 20 lists in order to know whether it meets the contemporary linguistic needs of the learners or not.

Statement of Problem

In English language teaching, there is notable emphasis on how vocabulary is chosen, distributed, and presented in the textbooks. Despite

the proven importance of managing lexical load in language teaching, it has not been given due attention to understand the efficacy of recommended textbooks in effectively managing and distributing the lexical load. This study investigates the lexical load present in Class X English textbooks.

Objectives

The objectives of the research are:

- To compute and compare the vocabulary items present in Class X textbooks

- To find the percentage of GSL/AWL lists coverage in Class X textbooks

- To find out the level of Class X textbooks vocabulary against the levels of BNC-COCA 25

Research Questions

- What percentage of GSL/AWL lists is covered in Class X textbooks?

- How many vocabulary items are present in Class X textbooks?

- Which level of BNC-COCA vocabulary is present in Class X textbooks?

Significance of the Research

This research will inform and improve the lexical aspect of English language teaching in Pakistan providing insights for designing the required vocabulary-oriented curriculum. It will help the textbook boards to reconsider the recommended textbooks from a vocabulary development perspective since these textbooks are the primary source of language teaching and learning.

Literature Review

English, being an international language, has an influence on the education, commerce, culture and literature of Pakistan (Azim et al., 2017). In all the provinces of Pakistan, English is being taught as a second language since grade 1. Every province is independent to develop National Education Policy based material for English language teaching. For this purpose, provinces have official bodies that are known as textbook boards. These boards produce no other language resource but textbooks only. That is why textbooks are very crucial and the only resource for Pakistani learners in the language teaching-learning process.

Nunan (1989) stated that a textbook is the fundamental part of the instructional method, so it serves as the standard model for classroom instruction. A textbook, for certain subjects, is a reference chart and key guide (Gurmani et al., 2021). According to Tomlinson (2012), textbook guides students for exam preparation, saves teachers time in lesson planning, and guides administration in designing contents of course (Tomlinson, 2012). Teachers, most often, rely on textbooks instead of syllabus, while planning lessons (Nelson, 2012). Knight (2015) concludes that textbooks are considered reliable and healthy support for the teachers and did not consider books as an outlying resource but rather as a chief resource.

Textbooks perform many functions and responsibilities but in Pakistan, these have additional responsibility of being the only source for vocabulary learning. No other means of vocabulary learning-teaching is available for the students of public schools.

Chegeni (2016) states that the choice of books and their evaluation is very important. Ahmadi and Derakhshan (2016) also emphasized on the evaluation of textbooks and considered it "vital and necessary" in order to look into its weaknesses and strengths. Vocabulary should be placed as the first target while selecting and developing language teaching material. Qian (2002) identified the vocabulary breath as one of the reading adroitness and language abilities indications. Substantial vocabulary makes the learners' skills better and insufficient vocabulary hinders the teaching-learning process of a language. Nation (2006) stated that vocabulary plays an important role and makes a great difference in learners' L2 efficiency and functionality. Cameron (2001) stated that vocabulary is one of the knowledge areas of language and it takes great part in language learning. Linse (2005) emphasized that vocabulary development is crucial part of language proficiency of L2 learners. Successful communication in the second language is developed on the basis of vocabulary. Limited vocabulary hinders successful second language communication. Schmitt (2000) declared that "lexical knowledge is central to communicative competence and to the acquisition of a second language". Nation (2001) listed reciprocal connection between knowledge of vocabulary and use of language as vocabulary knowledge facilitated the use of language. Furthermore, the use of language strengthens the vocabulary knowledge. Nunan (1991) asserted that sufficient vocabulary acquisition is

indispensable for using second language successfully because without sufficient vocabulary, learners will not be able to use language structures and functions they have learned for the understandable communication. It has been observed that L2 readers depend on the knowledge of vocabulary and insufficient L2 vocabulary knowledge is the chief and huge obstacle for L2 learners to deal with (Huckin et al., 1993). Wikins (1972, as cited in Alqahtani, 2005), stated that, "there is not much value in being able to produce grammatical sentences if one has got the vocabulary that is needed to convey what one wishes to say… while without grammar very little can be conveyed, without vocabulary nothing can be conveyed". Krashen (1989, as cited in Alqahtani, 2005), disclosed many grounds for giving attention to vocabulary. "First, an extensive vocabulary is of course essential for the mastery over language. Second, language acquirers know it that they carry dictionaries with them, not grammar books and regularly report that the lack of vocabulary is a major problem". Vocabulary is difficult to master because of its open-endedness besides, not having any defined rules like syntax and phonology. "Despite these difficulties that language learners face in L2 vocabulary, they still have to deal with it and their examination (Schmitt, 1997)". The importance of vocabulary development has been recognized by the researchers and language teachers but it is still a less explored area of research in Pakistan.

Gries (2009) stated that the corpus linguistics is not a branch of linguistics but an approach to language study. Practice and involvement of corpus linguistics remained less active in syllabus designing, reference works, and teaching activities and the use of corpora by both teachers and student was very rare and they were not familiar with the use of corpus tools. However, in the past two decades it has been extensively used for linguistic research and language learning-teaching. Now, the teachers and learners benefit from a large range of corpus-generated and corpus-based available works. Many researches have proven the significance of the use of corpora in language teaching. Currently, corpora have direct and indirect pedagogical implications in language teaching and learning. Where direct application makes it possible for the teachers and learners to interact with the corpora, the indirect application as adopted in this study, provides opportunities for the researchers and material writers to evaluate, investigate, and develop syllabus, reference works, and materials. Despite proven significance of corpora in language teaching-learning process, few researches have been carried out in Pakistan to investigate the textbooks in

general and the vocabulary load of English textbooks in specific. Jahan et al. (2019) did a lexical analysis of English course books developed by Punjab Curriculum and Textbook Board for the intermediate class. They analyzed English Book I, English Book II and English Book III based on English literature including short stories, one-act plays, poetry and prose. Khalil et al. (2022) investigated the frequency distribution of selected adjectives in English textbook used for intermediate classes among the four provinces of Pakistan. Gurmani et al. (2021) evaluated the vocabulary of Pakistani and Saudi English textbooks for secondary level classes.

Conceptual Framework

In this research, corpus linguistics was employed as the theoretical framework. From the offered concepts and methodologies of corpus linguistics, the frequency-based analysis and the paradigms of pedagogical corpus linguistics were integrated. The frequency-based analysis quantified the occurrences to identify the appearance and dominance of lexical items in the textbooks. Whereas, the pedagogical corpus linguistics analyzed the suitability of textbooks' lexical load, for language learning considering the following criteria on GSL/AWL lists and BNC-COCA 25.

General Service List (GSL)

GSL is the list of 2000 words published by Michel West in 1953. The list contains the words that are most frequent in English and were taken from the written English corpus. The target audience of this list was ESL teachers and learners. A learner can understand 90-95 percent of colloquial and 80-85 percent of the everyday written English if he or she knows the words present in GSL.

Academic Word List (AWL)

AWL contains 570 most frequently used academic words and was published by Averil Coxhead. A learner will be able to understand 90 percent of the academic texts if he or she is familiar with AWL. These words are not specific to any subject or discipline so, they are useful for every subject.

BNC-COCA 25

BNC-COCA 25 are the lists compiled by Paul Nation and his colleagues in 2012. These are the harmonized product of British National Corpus (BNC) and Corpus of Contemporary American English that contain 25000

most frequently used words of both British and American varieties of English.

Research Methodology

The nature of this research is quantitative and evaluative. The reason for keeping this research quantitative is that, as suggested by Azim and Garcia (2020), it suits best for the textbook's vocabulary profiling and is suitable for the vocabulary load calculation as well.

The textbooks for Class X published by Azad Jammu and Kashmir Textbook Board (AJKTBB), Baluchistan Textbook Board (BTBB), Khyber Pakhtunkhwa Textbook Board (KPKTBB), and Punjab Curriculum and Textbook Board (PCTB) were selected for the analysis. The strong motivation behind selecting these public domain textbooks for this research was that they are accessible by a large number of students who solely depend on these textbooks. These textbooks were converted into .txt format to make them usable as corpora. These corpora were then compiled and stored individually and named as AJKTBB.txt, BTBB.txt, KPKTBB.txt, and PCTB.txt for Azad Jammu and Kashmir Textbook Board, Baluchistan Textbook Board, Khyber Pakhtunkhwa Textbook Board and Punjab Curriculum and Textbook Board respectively.

After compilation, these corpora were analyzed and compared with each other by using the Sketch Engine (a web-based corpus manager) and text analysis software developed by Lexical Computing founded by Adam Kilgarriff in 2003 on the parameters of wordlist and N-grams. GSL/AWL coverage was found with the help of AntWordProfiler, another freeware tool used for profiling the vocabulary level and text's complexity. Some manual statistics was also done at the profiling stage.

'Compleat Web Vp' with BNC-COCA 25, a tool in Compleat Lexical Tutor's VP-Compleat, a free online software for vocabulary profiling available at lextutor.ca, was used to find out the distribution of textbooks vocabulary at different BNC-COCA levels. After that, obtained results were compared manually.

BNC-COCA 25 is a combination of British National Corpus (BNC) and Corpus of Contemporary American English lists. Initially, BNC lists were up to 2k levels with having most frequently used 2000 British English words. These words were taken from various sub-categories including spoken and written English. Later, by adding COCA, the lists went up to

25k levels comprising of the most frequently used 25000 words from both British and American English. These 25000 words are divided into 25 levels each having 1000 words in total ordered from 1 to 25. These lists were generated to help learners of English as a Foreign Language (EFL) but they are also used for the evaluation of existing language teaching material and the development of new one.

The researcher analyzed the selected textbooks individually and then compared them to know the most loaded one and up to the standard textbook. AntConc, a freeware by Anthony Lawrence for concordance, was utilized for GSL/AWL lists and Compleat Lexical Tutor's VocabProfilers, was used for BNC-COCA levels. After that, the results of these textbooks were compared manually in terms of the lists percentage they are covering.

Data Analysis

Vocabulary Items Present in Class X Textbooks

Type-Token Frequency and Ratio (TTR)

Table 1

Frequency of Word Types, Tokens, and their Ratio

Corpus	Types	Tokens	TTR (In Percent)
AJKTBB	2309	7924	29.14
BTBB	1937	7349	26.36
KPKTBB	2840	10845	26.18
PCTB	1778	6373	27.90

KPKTBB had the highest frequency of types of words and tokens that were 2840 and 10845 respectively. PCTB had the lowest text types that is 17748 and tokens 6373 as compared to others. 1937 word-types and 7349 tokens were listed in the corpus of BTBB whereas, AJKTBB corpus was comprised of 2309 word-types and 7924 tokens.

TTR of 29.14% was the highest for AJKTBB and 26.18% was the lowest for KPKTBB. Similarly, TTR for BTBB was 26.36% and 27.90% for PCTB as being the second highest. The highest TTR of AJKTBB makes it lexically the most variant of all and lowest TTR of KPKTBB makes it as the least lexical variant.

POS Comparison

Table 2

Frequency of Parts of Speech

Corpus	Adjective	Adverb	Conjunction	Noun	Preposition	Pronoun	Verb
AJKTBB	332	112	5	1106	56	27	409
BTBB	300	107	5	858	47	25	343
KPKTBB	430	153	6	1339	58	28	539
PCTB	279	114	4	768	49	23	342

Among all of the corpora, the most occurring part of speech is noun. KPKTBB is at the top in this respect and PCTB at the bottom. AJKTBB stands at the second position, while BTBB at the third. There are in total 7 coordination conjunctions in English language. None of them has all 7. However, KPKTBB has 6 of them that is the highest number as compared to others. PCTB has only 4 conjunctions that is the least number. Despite being the largest corpus, AJKTBB has only 5 conjunctions, equal to BTBB. AJKTBB has 332 adjectives, 112 adverbs, 56 prepositions, 27 pronouns, and 409 verbs. On the other hand, BTBB has 300 adjectives, 107 adverbs, 47 prepositions, 25 pronouns, and 343 verbs. Figures for adjectives, adverbs, prepositions, pronouns, and verbs in KPKTBB are 430, 153, 58, 28, and 539 respectively. 279 adjectives, 114 adverbs, 49 prepositions, 23 pronouns, and 342 verbs are present in PCTB. The highest adjectives are recorded in KPKTBB and the lowest in PCTB. Subsequently, KPKTBB has the highest number of adverbs, prepositions, pronouns, and verbs as well. PCTB has the lowest rate for all of them. A great variation is measured only for the noun class. The rest of the categories have less variations in all corpora.

Adjective, adverb, noun, and verb are the important parts of speech and verb is the most important but, in all the textbooks it has low occurrence than the noun.

N-Grams

Table 3

N-Grams up to Level 4

Corpus	2-Grams	3-Grams	4-Grams
AJKTBB	70	3	0
BTBB	83	12	2
KPKTBB	93	5	1
PCTB	61	3	0

There are 70, 83, 93, and 61 2-grams (also named as bigrams), 3,12,5, and 3 3-grams and 0,2,1, and 0 4-grams in AJKTBB, BTBB, KPKTBB and PCTB respectively. The highest numbers of 2-grams are present in KPKPTBB, while 3-Grams and 4-Grams in BTBB. PCTB has the lowest number of 2-grams that is 61 in total and 3-grams rate is only 3. Both PCTB and AJKTBB do not have any 4-grams.

GSL/AWL Lists Covered in Class X Textbooks

Table 4

GSL_1ˢᵗ_1000, GSL 2ⁿᵈ_1000, and AWL_570 Coverage in Textbooks

Corpus	GSL 1ˢᵗ 1000	GSL 2ⁿᵈ 1000	AWL 570	Off Lists	Total
AJKTBB	6031	538	332	1023	7924
BTBB	5807	524	524	764	7349
KPKTBB	8187	665	360	1630	10845
PCTB	5130	381	288	574	6373

GSL_1ˢᵗ_1000

6031 AJKTBB, 5807 BTBB, 8187 KPKPTBB, and 5130 PCTB word-tokens are recorded on GSL_1ˢᵗ_1000. KPKTBB scored the highest word-tokens on this list and PCTB has the lowest number. The sSecond highest is AJKTBB and the third is BTBB. These figures indicate that KPKTBB vocabulary is most promising for general purpose usage of English language on the other hand, least promising is of PCTB. AJKKTBB and BTBB have good weightage of vocabulary that makes them useful for general purpose English language use.

GSL_2ⁿᵈ_1000

Word-tokens on GSL_2ⁿᵈ_1000 for AJKTBB, BTBB, KPKPTBB, and PCTB are 538, 524, 665, and 381 respectively. Again, KPKTBB has the highest number of word-tokens on the list, PCTB has the lowest, AJKTBB the second highest, and BTBB ranked as the third. Usefulness of their vocabulary regarding usage of English for general purpose remains the same as of GSL_1ˢᵗ_1000 list.

AWL_570

In AWL_570, BTBB stands at the top with 524 word-tokens and PCTB at the bottom with only 288. KPKTBB is at the second position with 360 tokens and AJKTBB is at the third position with 332 word-tokens. These

results show that BTBB servers the best in all academic texts. PCTB is least at serving in academic writing as it has the lowest number of vocabulary items from this list. KPKTBB is the second best and AJKTBB is at the third level.

Off-List Words

1023 word-tokens of AJKTBB, 764 of BTBB, 1630 of KPKTBB, and 574 of PCTB were not on GSL/AWL lists.

AJKTBB proved to have good vocabulary in most parameters but figures show that it has the highest number of word-tokens that does serve neither the general purpose of English language use nor the academic purpose. In this regard, PCTB has the least word-tokens that are not on the lists. To make it clearer, percentages for both corpora were calculated. 9% PCTB and 12.9% AJKTBB word-tokens are not present on the GSL/AWL lists. Therefore, a large number of word-tokens of AJKTBB are not serving the general purpose of language use and academic texts of English language.

Levels of BNC-COCA Vocabulary Present in Class X Textbooks

Table 4

Levels of BNC and COCA Present in Textbooks

	AJKTBB	BTBB	KPKTBB	PCTB
K-1	6,351	5,934	8,563	5,133
K-2	706	726	900	629
K-3	390	272	534	302
K-4	147	129	267	85
K-5	95	89	165	51
K-6	49	38	86	46
K-7	40	22	51	18
K-8	17	20	48	21
K-9	13	8	26	9
K-10	10	13	13	20
K-11	3	11	15	22
K-12	1	12	13	3
K-13	6	5	14	12
K-14	2	0	17	1
K-15	6	1	6	1

	AJKTBB	BTBB	KPKTBB	PCTB
K-16	8	3	2	2
K-17	0	0	2	0
K-18	0	2	1	2
K-19	1	0	1	1
K-20	0	1	33	0
K-21	1	0	0	0
K-22	1	0	1	0
K-23	1	1	0	0
K-24	20	1	21	2
K-25	1	0	0	0

BNC-COCA has 25 levels, starting from K-1 and ending at K-25. Each level has 1000 words so, K-1 indicates the first 1000 most frequent words in BNC-COCA lists and K-2 presents the next 1000 most frequent words in English language. This goes up to K-25 which results in a total number of 25000 words. Highest number of words for every corpora are present at 1st level and gradually decrease with the progression towards next level. Up to K-8, there are reasonable word-tokens for all the corpora but after that, the numbers start to drop from there on.

Off-List Word

Table 5

Frequency of Off-List Words

	AJKTBB	BTBB	KPKTBB	PCTB
K-Off	56	54	24	17

There are 56 AJKTBB, 54 BTBB, 24 KPKTBB and 17 PCTB word-tokens that are not on BNC-COCA 25 lists. Apparently, the results show that AJKTBB has the highest number of off-list word-tokens but as per the percentage, BTBB has the highest. BTBB has 0.73% off-list word-tokens whereas, 0.71% of AJKTBB, 0.22% of KPKTBB, and 0.27% of PCTB word-tokens are not on the lists.

Off-list words can further be classified as English and non-English words. Most of the non-English off-list words are nouns and they are religion, culture and tradition oriented. Words like Allah, Muhammad, Nabi, Anbia, Rasool, Rasoolullah, Makkah, Madina, Umar, Abu Bakkar, Bukhari, Daood, and so on are purely religious. And names of the persons, for

example Allama Aurangzeb, Ayaz, Ayesha, Azam, Badshahi, and so on and names of places, namely Gujranwala, Gujrat, Karachi, Lahore are local adaptations to make the text relatable for the students.

English off-list words like website, talk shows, and Facebook are present in textbooks but not on BNC-COCA 25 lists. This is the limitation of the aforementioned lists that they are required to be updated in order to include new words for daily use.

Discussion

The following three parameters were set to analyze the vocabulary load of Class X English textbooks for Pakistani students, namely types-tokens frequency and type-token ratio, their coverage of GSL/AWL lists, and BNC-COCA 25 levels.

AJK Textbook Board's textbook introduced 2309 word-types with 70 bigram and 3 3-gram collocations. 1937 word-types with 83 bigram, 12 3-gram and 2 4-gram collocations were introduced in the textbook published by Baluchistan Textbook Board. Khyber Pakhtunkhwa Textbook Board's textbook for Class X was on the top in frequency list of tokens and collocations. It introduced the learners with 2840 word-types and 93 2-gram, 5 3-gram and 1 4-gram collocations. Punjab Curriculum and Textbook Board's textbook had the least amount of vocabulary that was of only 1778 word-types with 61 2-gram and 3 3-gram collocations. The parts of speech analysis revealed that while all textbooks contained notable number of nouns and verbs, they lacked in adjectives that are considered as the largest component of a language.

The only parameter where PCTB's textbook had higher levels of lexical occurrences was GSL/AWL lists. It covered approximately 91.00% of these lists. 80.50% of GSL_1ˢᵗ_1000, 6.00% of GSL_2ⁿᵈ_1000, and 4.50% of AWL was covered in it. 4.5% of AWL coverage makes the PCTB worthier of all for academic texts and writing. KPK Textbook Board's textbook despite having the largest number word-types, covered the least portion of GSL/AWL lists. It covered about the 85% of the lists in total and 75.50% of GSL_1ˢᵗ_1000, 6.20% of GSL_2ⁿᵈ_1000, and 3.30% of AWL. AJK Textbook Board's textbooks covered the 87.10%, while that of Baluchistan Textbook Board covered the 89.60% thus, making it second high.

The first three levels of BNC-COCA 25 were targeted in the textbooks by four major publishing houses of Pakistan. 80.40% of level 1, 9.0% of

level 2, and 5.0% of level 3 were targeted in AJK Textbook Board's textbook. Baluchistan Textbook Board targeted the 81.2% of level 1, 9.9% of level 2, and 3.7% of level 3. Similarly, KPK Textbook Board's textbook coverage was 79.2%, 8.4%, and 4.9% for level 1, 2, and 3 whereas, for PCTB's textbook this was 80.5%, 9.9%, and 4.8%.

Limitations

Neither dedicated corpus nor electronic version of some textbooks was accessible. The data was obtained from the .JPEG formatted images of the textbooks by using google-lens tool. This conversion of date from .JPEG to .txt files might have affected the corpus. Secondly, it was not in time constraint to address the lexical representation of the textbooks on other international standards, that is CEFR.

Recommendations and Suggestions

Vocabulary is classified as receptive and expressive vocabulary. Receptive vocabulary is the collection of words that a child can understand and expressive vocabulary is the repertoire of words that a child can use. By the age of 12, a child has receptive vocabulary of 50,000 words (Stahl, 1999) and a high school educated adult knows the vocabulary size of 15,000 words. Research shows that none of the textbooks meets the overall vocabulary need of the learners. Vocabulary load should be increased according to the need and age of the learners. Curriculum designers should consult contemporary research based on language teaching in general and vocabulary development in specific. International standards like GSL/AWL lists, BNC-COCA, and CEFR should be followed when developing language teaching materials.

Conclusion

This investigation of the lexical load of Class X English textbooks by using corpus tools has given way to valuable insights into the coverage of textbook vocabulary and the alignment with BNC-COCA and GSL/AWL. The quantitative assessment uncovered the extent to which the vocabulary resonates with the lexical resources in these lists. Some textbooks exhibit a higher lexical congruence with these lists while, others deviate from these lists which may cause challenges for students thus, hindering their language performance.

Moreover, this study shows the importance of corpus-based approaches in lexical evaluations of language textbooks. Curriculum developers can gain insights into the frequency and distribution of vocabulary by utilizing the lists such as BNC-COCA 25. It will help them make informed decisions regarding the selection of vocabulary syllabus design. It is recommended to language material developers to incorporate diverse vocabulary in accordance with well-known word lists that will foster comprehensive language proficiency among the learners. Additionally, a qualitative analysis of vocabulary usage within Class X textbooks is endeavored to delve deeper for knowing the difference of meaning, collocations, and contextual relevance.

In essence, this study contributes to the contemporary discussion on the importance of lexical load in language teaching material and focuses on the importance of evidence-based curriculum development approaches and language pedagogy.

Conflict of Interest

The author of the manuscript has no financial or non-financial conflict of interest in the subject matter or materials discussed in this manuscript.

Data Availability Statement

The data associated with this study will be provided by the corresponding author upon request.

References

Ahmadi, A., & Derakhshan, A. (2016). EFL teachers' perceptions towards textbook evaluation. *Theory and Practice in Language Studies*, *6*(2), 260–267. https://doi.org/10.17507/tpls.0602.06

Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be   taught. *International Journal of Teaching and Education*, *3*(3), 21–34. https://doi.org/10.20472/TE.2015.3.3.002

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.

Azim, M. U., Hussain, Z., Bhatti, A. M., Iqbal, M., & Chohan, M. (2017). Caught between the Extremes: A comparative study of state owned news

channel and a private news channel. *Hamdard Islamicus*, *40*(2), 301–314.

Azim, U. A., & Garcia, M. I. M. (2020). Lexical load of Punjab curriculum and textbook board's English1and English 2. *Bulletin of Education and Research*, *42*(2), 15–32.

Cameron, L. (2001). *Teaching languages to young learners*. Cambridge University Press.

Chegeni, N. (2016). General English textbook evaluation: A closer look at "four corners." *Theory and Practice in Language Studies*, *6*(12), 2325–2330. https://doi.org/10.17507/tpls.0612.13

Government of Pakistan. (2017). *National education policy 2017*. https://pbit.punjab.gov.pk/system/files/National%20Educaton%20Policy%202017.pdf

Gries, S. T. (2009), What is corpus linguistics? *Language and Linguistics Compass*, *3*, 1225–1241. https://doi.org/10.1111/j.1749-818X.2009.00149.x

Gurmani, M. T. A. A., Yosuff, Z. S., & Shahid, C. (2021). An evaluation of vocabulary of English textbooks of Pakistan and Saudi Arabia through CEFR. *Palarch's Journal of Archaeology of Egypt/Egyptology*, *18*(10), 3375–3387.

Hsueh-Chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, *13*(1), 403–430.

Huckin, T., Haynes, M., & Coady, J. (Eds.) (1993). *Second language reading and vocabulary learning*. Ablex Publishing Corporation.

Jahan, K., Mahmood, M. A., & Azhar, W. (2019). Lexical analysis of intermediate English course books: A corpus based study. *International Journal of Education Science*, *24*(1-3), 13–22,

Kauser, G., Mushtaq M., & Badshah, I. (2016). The evaluation of English language textbook taught at intermediate level. *Gomal University Journal of Research*, *32*(1), 32–43.

Khalil, S., Muhammad, S., Hamid, N., & Khatoon, S. (2023). Frequency and collocational meanings of near-synonymous adjectives in

secondary school level English textbooks in Pakistan: A corpus-based study. *Erevna: Journal of Linguistics and Literature*, *6*(2), 30–42.

Knight, B. (2015). Teachers' use of textbooks in the digital age. *Cogent Education*, *2*(1), Article e1015812. https://doi.org/10.1080/2331186X.2015.1015812

Lessard-Clouston, M. (2013). *Teaching vocabulary*. TESOL International Association.

Linse, C. (2005). *Practical English language teaching: Young learners*. McGraw-Hill.

Mansoor, S. (2005). *Language planning in higher education. A case study of Pakistan*. Oxford University Press.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, *63*(1), 59–82. https://doi.org/10.3138/cmlr.63.1.59

Nelson, J. (2012). How is the textbook used by teachers and students? *Nordina: Nordic Studies in Science Education*, *2*(2), 16–27. https://doi.org/10.5617/nordina.421

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge University Press.

Nunan, D. (1991). *Language teaching methodology: A textbook for teacher*. Prentice Hall.

Qian, D. D. (2002) Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, *52*, 513–536. https://doi.org/10.1111/1467-9922.00193

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Descriptive, acquisition and pedagogy* (pp. 199–227). Cambridge University Press.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.

Shahbaz, M., & Liu, Y. (2012). Complexity of L2 motivation in an Asian ESL setting. *Porta Linguarum*, (18), 115–132.

Shamim, F. (2011). English as the language for development in Pakistan: Issues, challenges and possible solutions. In H. Coleman (Ed.), *Dreams and realities: Developing English in Development Contexts* (pp. 1–21). British Council.

Stahl, S. A. (1999). *Vocabulary development: From reading research to practice*. Brookline Books.

Tomlinson, B. (2012). Materials development for language learning and teaching. *Language Teaching*, *45*, 143–179. https://doi.org/10.1017/S0261444811000528