AsiaCALL Online Journal
ISSN 1936-9859; https://asiacall-acoj.org
Vol. 15, No. 1, 2024
pp. 117-128

Received: 09/03/2024
Revision: 30/05/2024
Accepted: 04/06/2024
Online: 11/06/2024

What Types of Vocabulary Do Students Learn? A Computer-assisted Study of EFL Textbooks

Nasreen Bibi[1], Samina Hayat[1], Abdul Aleem Yahya[2*]

[1]English Language Institute, Jazan University, Saudi Arabia
[2]Department of English, University of Education, Pakistan
[*]Corresponding author's email: aleemyahya@gmail.com
[*]https://orcid.org/0000-0001-6001-2466
https://doi.org/10.54855/acoj.241518

Abstract

There has been an increase in research based on corpus analysis, especially in the field of teaching English to foreign language students. Such research covers aspects like grammar, syntax, morphology, semantics, and pragmatics in a broader sense. However, corpus studies on vocabulary are limited, and the need to work on them has increased in the last decade. Vocabulary is essential in learning and teaching any second language to English as a foreign language (EFL) classes. This study aims to thoroughly analyze vocabulary in three textbooks through a corpus analysis tool that is used to teach English as a foreign language to level 1, level 2, and diploma-taking students at a college in Saudi Arabia. Vocabulary is scrutinized through AntConC (1.4.0), a freeware program that uses statistical algorithms to provide visualizations of word frequencies, lexical clusters, and keywords. The results can be seen in the form of lists, tables, or charts to help you understand and draw conclusions. This selected corpus contains 163,716 word tokens. After statistical analysis of these words, English language teachers (ELTs) will understand what type of vocabulary is important for students for a particular level. It will also tell whether the vocabulary used in these textbooks is the correct representation of academic words when compared to baby BNC. Moreover, the study will provide insight to academicians, language teachers, and curriculum designers on what type of vocabulary is suitable for developing reading, writing, listening, or speaking skills according to the needs and levels of students.

Keywords: Corpus Studies, ESL, Lexical Clusters, BNC

Introduction

Every textbook provides a basis for the foundation of learning in institutions. In the same way, second language textbooks also provide a basis for learning languages. The vocabulary used in those textbooks is crucial for the development of language skills in students, so it is important that the vocabulary is selected carefully while keeping in mind the needs of the learners. Unfortunately, the amount of research on textbooks is few that may assess their efficiency in second language learning (Chan, 2000), and there also seems to be a lack of standardized vocabulary in textbooks that, in turn, may affect student's learning (Yang, 2006). Therefore, the aim of this research is to look at the issue of vocabulary used in the selected textbooks.

Till now, the linguistic research on the construction of L2 textbooks and the content of

CITATION | Bibi, N., Hayat, S., & Yahya, A. A. (2024) What Types of Vocabulary Do Students Learn? A Computer-assisted Study of EFL Textbooks. *AsiaCALL Online Journal*, 15(1), 117-128. DOI: https://doi.org/10.54855/acoj.241518

vocabulary provided in them is limited, keeping in mind the pedagogical point of view. Some researchers have employed corpus-based approaches, but their main focus is on teaching material with regard to written academic materials (Hoey, 2000; Thompson, 2000). In the same way, studies regarding the needs of college learners have been ignored (Foster & Mackie, 2013; Keck, 2004). Keeping in view the importance of second language learning for college learners, it seems shocking that significant research is lacking (Konstantakis & Alexiou, 2012; Skolverket, 2006).

The current paper aims to examine the vocabulary of three textbooks used for level 1, level 2, and diploma-taking students at the college level in Saudi Arabia. The researcher intends to describe the type and frequency of the vocabulary used in the textbooks with a view to making comparisons with Baby BNC (British National Corpus).

The role that vocabulary plays in language learning is crucial. The choice of words affects the way students write sentences and produce utterances in spoken language. Language shapes the students' discourses, making their communication more or less effective. Words are containers of meaning, and the need to use the right words in the right context is important for communicative competence. Hence, the given study is conducted to evaluate the vocabulary in terms of its representativeness when compared to Baby BNC.

Baby BNC Edition

Baby BNC is a sub-part of BNC (British National Corpus). It contains four million word samples representing four genres: fiction, newspapers, academic writings, and spoken conversations. The texts in these corpora are annotated similarly to the original full corpora. Its first edition was released in March 2007.

The four genres included in baby BNC contain four million words, and each genre represents one million words, hence creating a pool of equality in terms of representativeness. Texts are copied directly from BNC XML, 2008 edition, without adding or subtracting. The selection of texts is made based on the information provided in the original corpus and in accordance with the text classification given by David Lee (Lee, 2021).

Vocabulary

Building sufficient vocabulary is crucial for learning any second language. Without learning a substantial amount of second language words, learners may be unable to produce correct sentences according to the context. So, vocabulary knowledge plays an important role in the development of reading skills (Cameron, 2001) and is also closely related to the overall success of the learners in learning a second language. Memorizing and studying words are not enough, but understanding them requires knowledge of meaning, form, collocation, register, and association (Nation, 2013). To grasp the knowledge of any word, it must recur many times in multiple contexts, which may, in turn, automatize word knowledge (Tyler, 2012). Words are not stored in our brains as isolated units but rather in the form of semantic networks. These semantic networks are built when the same words appear before learners in multiple contexts, forming connections among them. These connections are internalized and remembered afterward (Cameron, 2001).

Researchers are not sure how many times a word should occur before learners completely comprehend it. Some say the number of occurrences is between 5-6 (Cameron, 2001; Nation, 1990) and twenty (Waring & Takaki, 2003). As such, repetition is more important for beginner learners than for advanced learners (Zahar, Cobb & Spada, 2001). Another important aspect is to understand the size of vocabulary that may be included in textbooks necessary for learning a second language. Moreover, the number of words that a learner may know during the course

of second language learning is also crucial. Such data is found in curricula of Greece, Spain, and South Korea (Jiménez Catalán & Mancebo Francisco, 2008) but not in Saudi Arabia.

As said, it is difficult to know the number of words a learner may know to understand a passage of an average difficulty level. Some researchers show that high-frequency words are closely related to successfully learning a second language (Nation, 2006). In the same way, 250 words occurring most frequently in children's speech account for 75 to 80 percent of language production (Roessingh & Cobb, n.d.). These works suggest that a small number of words may help considerably in comprehending a second language. Learning 1000 to 2000 most frequent words in any discourse may help learners develop, comprehend, and produce the vocabulary of a second language.

Research Objectives

This research has the following objectives: first, to understand the extent to which the selected books may represent the required vocabulary suitable for the students; second, to describe the frequencies of functional and content words and their importance in the corpora of English as a foreign language learning textbooks.

Literature Review

Research focusing on structuring textbooks, especially vocabulary, by considering learners' needs is scarce. However, Shin and Chon (2011) are some examples of such research. They worked on vocabulary profiles of Elementary and Secondary textbooks taught in South Korea. By comparing the vocabulary of these textbooks with West's (1953) 2,000 General Service List (GSL), they conclude that 68% of the vocabulary is not found in the GSL list. This suggests that most words used in these textbooks are used infrequently in language production.

Thompson and Sealey (2007) also carried out a similar study for the vocabulary used in children's literature. When they compared children's literature corpus (CLLIP corpus) to that of adults' literary vocabulary profiles, they found that much of their vocabulary is common and shares most of the linguistic properties. Rixon (1999) also examined vocabulary in teaching materials. She examined seven textbooks taught to first-year English learners and concluded that most of the vocabulary in these seven textbooks was unsuitable for first-year English learners. Moreover, half of the vocabulary used in those seven books was different and unique, which may create problems for first-year learners. The present study also aims to understand the type of vocabulary used in the selected textbooks and their differences when compared to baby BNC.

Research Methodology

Research Design

This study employs quantitative research methodology to investigate numerical data and qualitative research methodology to interpret the results. According to Creswell (2012), the quantitative method is the process of collecting numerical data from people using an instrument such as questions or interviews. After collecting the data, numeric results are interpreted, and the researcher gets the optical values for the frequencies of the vocabulary.

The given study explores the vocabulary of English textbooks used for level 1, level 2, and diploma-taking colleges in Saudi Arabia. The text is analyzed at word, phrase, and sentence levels. One of the basic procedures adopted for corpus analysis is to find the frequencies of

different words or word units (morphemes). Using the frequency tool, the researchers may assess the importance of high-frequency words in a text. Such words may also be compared to other compiled corpora (e.g., baby BNC) to assess their usefulness in a certain textbook. It is a high frequency and may be high in significance, making it a characteristic feature of the text. Therefore, a mixed methodological research design is adopted for this particular study.

Subject of the Study

The following study investigates three English textbooks used for teaching English as a foreign language. These books are used to teach level 1 and level 2 and diploma-taking students at the higher secondary school level in Saudi Arabia.

These textbooks are analyzed for their vocabulary only. The corpus prepared for analysis does not contain figures, charts, tables, footnotes, endnotes, references, or pictures. Only plain text is to be considered in order to develop a suitable and representative corpus.

Collection of Data

The data collected for this study is taken from English textbooks specially designed to cater to the language needs of second-language learners in Saudi Arabia.

Foreign English authors write all these textbooks. The corpus that will be used for analysis is prepared from these textbooks, and it contains 163,716 word tokens. The researcher's purpose is to look at the number and types of vocabulary used in these textbooks and compare them with baby BNC to see whether the vocabulary represents a fair number of suitable words.

Research Instrument

A research instrument is a tool used to process data for scientific analysis. A corpus is a large amount of data extracted from various digital sources. In the case of this study, AntConC (1.4.0) is used to analyze corpora extracted from selected academic textbooks. AntConC (1.4.0) is multi-task software that is primarily developed to perform vocabulary profiling in corpus linguistics. Now, it is being used in many fields of knowledge to perform language-related tasks. It is developed by Anthony Laurence (2014) from Waseda University to perform various automated analyses on digital language. This software performs two functions: vocabulary profiling and file viewing and editing. Through the profiling tool, lists of the vocabulary are generated based on statistics and frequencies. The statistics and frequencies are then compared to other standardized lists, such as BNC. The file viewing and editing option allows one to color-code and highlight different levels of vocabulary, while the editing option allows making changes in the file and then seeing the results.

Rationale

The study uses a computer-based freeware program known as AntConC (1.4.0), which was developed by Anthony Laurence in 2014. This tool provides various types of language processes, but the researcher intends to use it to analyze vocabulary in EFL textbooks. Many other researchers have used this program previously, so it caters to everyone's needs. In this regard, the process being taken in this research is not new, but the research objectives are unique, so there still needs to be a lot of work to be done in this field

Analysis

Target and Reference Corpus

Target corpora in the given research are collected and prepared from three textbooks taught in Saudi Arabia to English as a second language learners. On the other hand, baby BNC is used

as a reference corpus to find differences in both corpora and understand the use of vocabulary. Representativeness of the target corpus is maintained by selecting textbooks that may balance the use of academic vocabulary in various contexts. Textbooks selected provide contextual usage of words in multiple contexts ranging from fictional stories to non-fictional texts.

Word lists

Following are the word lists of target and reference corpora generated through computer-assisted software. The word list is a list of all the words that occur in a particular text and are arranged in order of its frequency. To explore the type of words (content/functional) used in these corpora, lists are generated to find out their type and frequencies. For convenience, only excerpts from the original source are provided here.

Table 1

Type, rank & frequency in the reference corpus

Type	Rank	Freq	Range	NormFreq	NormRange
the	1	70362	30	69913.366	1
of	2	44256	30	43973.82	1
and	3	26841	30	26669.86	1
to	4	26348	30	26180.003	1
in	5	25265	30	25103.908	1
a	6	23252	30	23103.743	1
is	7	17432	30	17320.852	1
that	8	12476	30	12396.452	1
for	9	9438	30	9377.823	1
be	10	9401	30	9341.058	1
as	11	9252	30	9193.008	1
it	12	8268	30	8215.283	1
are	13	7551	30	7502.854	1
by	14	7340	30	7293.2	1
this	15	7084	30	7038.832	1
with	16	6991	30	6946.425	1
which	17	6026	30	5987.578	1
on	18	5843	30	5805.745	1
or	19	5547	30	5511.632	1
not	20	5229	30	5195.659	1
was	21	5078	30	5045.622	1
have	22	4712	30	4681.956	1
from	23	4634	30	4604.453	1

Table 2

Type, rank & frequency in the target corpus

Type	Rank	Freq	Range	NormFreq	NormRange
the	1	9918	1	60580.517	1
a	2	5070	1	30968.262	1
in	3	3867	1	23620.172	1
to	4	3743	1	22862.762	1
and	5	3401	1	20773.779	1
you	6	2878	1	17579.223	1
i	7	2739	1	16730.191	1
of	8	2700	1	16491.974	1
is	9	1806	1	11031.298	1
s	10	1439	1	8789.611	1
it	11	1395	1	8520.853	1
with	11	1395	1	8520.853	1
t	13	1383	1	8447.556	1
for	14	1337	1	8166.581	1
are	15	1286	1	7855.066	1
we	16	1278	1	7806.201	1
your	17	1274	1	7781.768	1
do	18	1122	1	6853.331	1
at	19	1083	1	6615.114	1
have	20	1025	1	6260.842	1
about	21	984	1	6010.408	1
what	22	942	1	5753.866	1
b	23	918	1	5607.271	1

As can be seen, the highest frequencies are found to be of functional words with ‘the’ word at the top. ‘The’ shows a frequency of 70362 (69913.366 norm frequency) in the reference corpus and 9918 (60580.517 norm frequency) in the target corpus. Similarly, 'of' comes second with a frequency of 44256 (43973.82 norm frequency) and 'a' with 5070 (30968.262 norm frequency), and the list goes on. The differences in the use of lexical items suggest that students are given opportunities to use words in a variety of contexts that may enhance their understanding of the second language. The target corpus provides a balanced source of content to functional words ratio, implying that the selected textbooks are written with great care, keeping in mind the linguistic needs of second language learners. To further understand the phenomenon stated above, cluster charts of both the reference and target corpus are generated and presented below.

Lexical Clusters

Clusters are the repeated sequences of words that occur in texts. It may help in understanding the importance of sequences of words (lexical clusters) that repeatedly occur for a significant number of times. The following are the two-word clusters found in target and reference corpora, and they enable the researcher to hypothesize about the word sequences that are more useful for the learning of a second language for students.

Table 3

Cluster, rank, and frequency in reference corpus

Cluster	Rank	Freq	Range	NormFreq	NormRange
the same	1	875	30	0.012	1
the first	2	624	29	0.009	0.967
the other	3	526	29	0.007	0.967
the most	4	499	29	0.007	0.967
the two	5	356	28	0.005	0.933
the case	6	331	27	0.005	0.9
the second	7	300	29	0.004	0.967
the time	8	269	27	0.004	0.9
the number	9	262	23	0.004	0.767
the party	10	259	8	0.004	0.267
the problem	11	256	28	0.004	0.933
the use	12	241	27	0.003	0.9
the end	13	239	27	0.003	0.9
the last	13	239	29	0.003	0.967
the new	15	231	26	0.003	0.867
the way	16	227	27	0.003	0.9
the fact	17	224	28	0.003	0.933
the law	17	224	14	0.003	0.467
the following	19	220	28	0.003	0.933
the form	20	217	26	0.003	0.867
the whole	21	213	27	0.003	0.9
the more	22	212	29	0.003	0.967

The first five frequent words that occur with the article 'the' are 'same, 'first', 'other', 'most' and 'two' in the target corpus, while 'sentences', 'questions', 'article', 'correct' and 'past' in the reference corpus. This suggests that the target corpus uses more adjectives, highlighting nouns' qualities. Adjectives create importance for nouns and, provide attraction for the readers and develop their interests. For second language learners, creating interest is important as it keeps them engaged in reading. On the other hand, nouns are used more frequently with the article ‘the’ suggesting that nouns present less interesting information for the readers. The readers have to put more effort into understanding the phenomenon given in the text due to the lack of descriptions. As such, the target corpus is more engaging and interesting than the reference corpus for second language learners. Moreover, the target corpus provides more variety of words, especially nouns and their descriptors (adjectives).

Table 4

Cluster, rank, frequency in target corpus

Cluster	Rank	Freq	Range	NormFreq	NormRange
the sentences	1	354	1	0.036	1
the questions	2	240	1	0.024	1
the article	3	215	1	0.022	1
the correct	3	215	1	0.022	1
the past	5	169	1	0.017	1
the words	5	169	1	0.017	1
the same	7	148	1	0.015	1
the present	8	135	1	0.014	1
the video	8	135	1	0.014	1
the world	10	126	1	0.013	1
the photo	11	113	1	0.011	1
the first	12	105	1	0.011	1
the city	13	96	1	0.01	1
the grammar	14	78	1	0.008	1
the verbs	15	76	1	0.008	1
the future	16	75	1	0.008	1
the text	17	73	1	0.007	1
the verb	18	66	1	0.007	1
the expressions	19	65	1	0.007	1
the other	20	64	1	0.006	1
the people	21	61	1	0.006	1
the information	22	58	1	0.006	1

Keywords (Keyness)

Keyness refers to the frequency of words that occur in the target corpus in comparison to the reference corpus. The idea of keyness points out the relative importance of words compared to a reference corpus, providing insight into their usage in context. The comparison chart is given below.

The above table presents relative frequencies and keyness in both target and reference corpora. As is evident from the data, tokens show that, by and large, their frequencies in the target corpus are greater in number than in the reference corpus. This suggests that the target corpus uses the given words relatively more than the reference corpus, and so they have more keyness effect.

Table 5

Type, keyness & keyness effect in both corpora

Type	Rank	Freq_Tar	Freq_Ref	Keyness (Likelihood)	Keyness (Effect)
you	1	2878	1660	5896.906	0.034
your	4	1274	579	2891.477	0.015
my	5	775	294	1882.3	0.009
sentences	6	659	151	1860.595	0.008
do	7	1122	1036	1742.254	0.014
listen	8	435	6	1650.377	0.005
pairs	9	425	21	1509.655	0.005
questions	10	600	277	1351.235	0.007
complete	12	494	163	1257.274	0.006
verbs	14	292	7	1084.732	0.004
go	15	475	214	1080.155	0.006
exercise	16	363	79	1037.303	0.004
work	17	855	1084	1031.397	0.01
watch	18	268	7	991.444	0.003
words	19	493	316	953.015	0.006
look	20	435	230	923.577	0.005
people	21	719	868	905.541	0.009
about	22	984	1629	902.801	0.012
read	23	378	174	851.923	0.005
what	24	942	1714	769.992	0.011
think	25	387	240	760.82	0.005
yes	26	245	42	737.699	0.003
past	27	373	229	737.043	0.005
video	28	194	4	725.363	0.002
grammar	29	234	35	723.268	0.003

Discussion and Conclusion

The present study looks at the words used both in target and reference corpora. The target corpus is prepared from three textbooks taught at the higher secondary school level containing 163,716 tokens. The main intent of the research is to look at the type of vocabulary used in the target corpus and its effectiveness in learning a second language. This is done by analyzing the corpus with the help of computer-assisted software AntConc (1.4.0), which generates results. Word lists and clusters are generated to draw results and make generalizations. The reference corpus used is baby BNC, which comprises four million words related to fiction, newspapers, academic writing, and spoken conversations. Baby BNC is used as a standard against which the target corpus is compared. The lists of the results given in this research are taken from original reports and contain limited tokens to avoid length. Following would be the final results that may be concluded from the analysis:

i. The textbooks used at the higher secondary school level for teaching English as a foreign language contain a balanced number of content to functional word ratios in a

variety of contexts when compared to a reference corpus. This may enhance students' understanding of the use of words and enable them to see words in various contexts.

ii. The frequency of functional words is the highest in the target corpus, which also seems to be the case when the reference corpus is examined individually.

iii. The high-frequency content words in target corpus are the adjectives that provide opportunities for students to learn them along with nouns, enriching their vocabulary.

iv. The third most high-frequency content words are found to be verbs, and again, verbs are important to learn and present a rich repertoire of words.

v. The examination also reveals the recurrences of prepositions like 'the' with certain nouns, hence forming lexical clusters. The list of these lexical clusters may help students to learn along with individual vocabulary words.

vi. Keyness and its effect is major indicator in highlighting the importance of high frequency words that appear at the top of table 5. For example, the keyness (likelihood) of the token ‘you’ is 5896.906 with an effect of 0.034 in the given data. This indicates that ‘you’ is significantly important either in isolation or in relation to other words and the list goes on with decreasing keyness effect.

Apart from the results, it is well understood that students need more than just a few hundred recurring high-frequency words to fully comprehend the language. It is not about memorizing and understanding 1000 to 2000 most frequent words to communicate in the English language, but it requires both high and low-frequency words to cover situation-based contexts. There is no way that a text should contain only specific words because it will render the text boring and unrealistic. There must be a balance between the uses of different word classes to produce reasonable native-like texts. Learners should be provided maximum types of words in varying contexts to enhance their comprehension level.

The books that are analyzed in this regard contain a good balance of high and low-frequency words. Compared to a reference corpus, almost the same types of words are found with a 2:1 ratio. The researcher approves that these books are written with great care and caution and in accordance with the lexical needs of the second language learners. This research may help academicians select and sort ELT books for better teaching practice for foreign language practical teaching.

This study limits itself to specific aspects of English as a foreign language, potentially overlooking other broader issues of context, and so the findings may not be generalized to students of other cultures. Moreover, English language trends and usage can change rapidly, making it challenging for research to keep pace with evolving language dynamics. The use of modern technologies like interactive online chatbots may also pose a threat to students' learning of English vocabulary (Bin-Hady et al., 2023 & Ho, 2024). At the same time, online chatbots help students get linguistic inputs and provide real-time assistance (Kohnke et al., 2023), but since these chatbots are automated machines, they have drawbacks as well.

References

Bin-Hady, W. R. A., Al-Kadi, A., Hazaea, A., & Ali, J. K. M. (2023). Exploring the dimensions of ChatGPT in English language learning: A global perspective. *Library Hi Tech*. Vol. ahead-of-print No. ahead-of-print. DOI: https://doi.org/10.1108/LHT-05-2023-0200

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

Chan, Y. J. (2000). A study of the development and trend of children's English education and issues on teachers, materials, and methodology. *Journal of National Taipei Teachers College*, 13, 203-238.

Creswell, J. W. (2012). *Educational Research*. Pearson.

Foster, J., & Mackie, C. (2013). Lexical analysis of the Dr. Seuss Corpus. *Concordia Working Papers in Applied Linguistics*, 4, 1–21.

Ho, P. X. P. (2024). Using ChatGPT in English Language Learning: A Study on I.T. Students' Attitudes, Habits, and Perceptions. *International Journal of TESOL & Education*, 4(1), 55-68. https://doi.org/10.54855/ijte.24414.

Hoey, M. (2000). The hidden lexical clues of textual organization: A preliminary investigation into an unusual text from a corpus perspective. In L. Burnard & T. McEnery (Eds.), *Rethinking Language Pedagogy from a Corpus Perspective* (pp. 31–42). New York: Peter Lang.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for Language Teaching and Learning. *RELC Journal*, 54(2), 537-550. https://doi.org/10.1177/00336882231162868.

Jiménez Catalán, R.M., & Mancebo Francisco, R. (2008). Vocabulary input in EFL textbooks. *RESLA*, 21, 147–165.

Keck, C.M. (2004). Corpus linguistics and language teaching research: Bridging the gap. *Language Teaching Research*, 8(1), 83–109.

Konstantakis, N., & Alexiou, T. (2012). Vocabulary in Greek college learners' English as a foreign language course books. *The Language Learning Journal*, 40(1), 35–45.

Lee, D. (2001) 'Genres, registers, text types and styles: clarifying the concepts and navigating a path through the BNC Jungle' in *Language Learning and Technology*, vol 5 no 3, September 2001; available online at http://llt.msu.edu/vol5num3/lee/default.html.

Nation, I.S.P. (2013). *Learning Vocabulary in Another Language* (2nd ed.). Cambridge: Cambridge University Press.

Nation, I.S.P. (1990). *Teaching and learning vocabulary*. New York, NY: Newbury House.

Nation, I.S.P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59–82.

Rixon, S. (1999). Where do the words in EYL textbooks come from? In S. Rixon (Ed.), *College learners of English: Some research perspectives* (pp. 55–71). Harlow: Longman.

Roessingh, H., & Cobb, T. (n.d.). Profiling the vocabulary of K-2 learners. Retrieved on April 19, 2016 from http://www.uclagary.ca/langugaelearning/profilingvocabulary.

Shin, D., & Chon, Y.V. (2011). A corpus-based analysis of curriculum-based elementary and secondary English textbooks. *Multimedia-Assisted Language Learning*, 14(1), 149–175.

Skolverket. (2006). Läromedlens roll i undervisningen. Grundskollärares val, användning och bedömning av läromedel i bild, engelska och samhällskunskap [The role of teaching materials in teaching: Secondary school teachers' choice, use and evaluation of teaching materials in Arts, English and Social Sciences]. Stockholm: Skolverket.

Thompson, P. (2000). Citation practices in PhD theses. In L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective* (pp. 91–102). New York, NY: Peter Lang.

Thompson, P., & Sealey, A. (2007). Through children's eyes? Corpus evidence of the features of children's literature. *International Journal of Corpus Linguistics*, 12(1), 1–23.

Tyler, A. (2012). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. London: Routledge.

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15(2), 130–163.

West, M. (1953). *A general service list of English words.* London: Longman.

Yang, I. L. (2006). On the issue of vocabulary size in English Teaching in Taiwan. *Journal of the National Institute for Compilation and Translation*,34(3), 35-44.

Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review,* 57(4), 541–572.

Biodata

Abdul Aleem Yahya is working as a lecturer at the University of Education, Multan Campus, Pakistan, since 2021. He has worked with various institutions and students of diverse backgrounds. His research interests include topics related to cognitive linguistics, language psychology, and teaching of English as a second language.

Nasreen BiBi, English Language Institute, Jazan University, Al Maarefah Road, Jazan, Saudi Arabia. Tel: +966 53 324 9728 E-mail: nfayyaz@jazanu.edu.sa; nqurashi91@gmail.com

Samina Hayat, English Language Institute, Jazan University, Al Maarefah Road, Jazan, Saudi, Arabia. Tel: +966 59 173 1939 E-mail: shmuhammad@jazanu.edu.sa; samina.jan16@gmail.com