



A Corpus Analysis of SSC English Textbook Phrases of Khyber Pakhtunkhwa: A Lexicographical Study

Zarhat Ali¹, Umme Habiba² & Usama Subhan³

¹ (Corresponding Author) PhD Scholar, Kohat University of Science and Technology, Kohat
Email: zarhatali11@gmail.com

² M.Phil. Scholar, Department of English, Northern University Nowshera Email: hibaa5548@gmail.com

³ M.Phil. Scholar, Department of English, Northern University Nowshera Email: usamasubhan928@gmail.com

Abstract

The planned study aims to do a corpus analysis of words from the secondary School certificates (SSC) English textbooks in Khyber Pakhtunkhwa with a Lexicographic approach. This study aims to provide insight into prevalent wording and usage patterns by evaluating the linguistic setting of these textbooks. The study will look into how phrase usage varies in a variety of textbook scenarios and subjects using a corpus-based methodology. Through a thorough analysis of the phrases covered in the SSC British educational program, this work aims to further lexicographic research. The study will make use of qualitative methods, such as term collection and evaluation from the corpus using specialized programs. The results have detected in light on the linguistic characteristics of the textbooks, which is beneficial to lexicographers, instructional designers, and instructors.

Keywords

Corpus Analysis, Lexicographic Approach, SSC English Textbooks

Introduction

The language and educational setting of Pakistan's Khyber Pakhtunkhwa (KP) province are distinct, as seen by the expressions used in the English textbooks for the SSC (Secondary School Certificate). To improve our comprehension of the lexical characteristics unique to this educational setting, the purpose of this work is to analyse these phrases using a corpus. This study uses lexicographical approaches to identify the grammatical structures and social factors that are commonly seen in KP's English language syllabus (Aarts & Meijs, 2022). The process of creating dictionaries dates back to 1600, when Robert Cawdrey added terms from other languages that he thought were challenging to his own translation of the dictionary. The dictionary definitions, synonyms, and fixed forms were applied to words that were obtained from Latin-English dictionaries as well as from books that were readily available at the time. Samuel Johnson was the one who first published the procedures or methods used to build his lexicon in the 1700s. The committee tasked with creating "A New Dictionary," or the Oxford English Dictionary, in the 1800s used some of the same procedures. A corpus comprises genuine spoken and written text samples that are analyzed for word meanings, grammar, and use. The text is comparable to parole in Saussurian terminology, whereas the corpus offers the proof of language. When a corpus is explicitly used to study a language, the term corpus linguistics is employed (Gul et al, 2022(c))

It is anticipated that vocabulary levels ranging from three thousand to five thousand would allow the English language as a foreign language learner (EFL) to fully understand the language to a certain extent in an EFL circumstance. Given that vocabulary is seen as a crucial aspect of language learning, exposure to the target language and a variety of resources help them expand their list of words and enhance their comprehension. Even though there are many readily available vocabulary sources these days, students of EFL frequently only have access to what is provided and discussed in

classes. Since English is not the primary tongue of the nation, there is a dearth of exposure to it in the KP SSC English Course. In other words, most KP interacts with acquaintances, coworkers, and other strangers in their mother tongue, Pashto. Even though English is widely used in the nation's economy—millions of expats, in example, work in the business community and rely on it for communication—they usually acquire the native tongue of the KP to be able to interact with them. Since KP doesn't use English in every aspect of his life, academics are by far the most important source of language purchase (Ha, 2020).

Nation (2013), while developing instructional and educational resources for languages, the problem of organising and displaying items from the dictionary has received little attention. According to Nation and Beglar 2007 and Ishtiaq et al., 2021 research, text saturation in a text is directly correlated with the emergence of high-frequency words. Nation (2006), when acquiring terms, 1,500–2,000 high-frequency phrases should be learned fast and with greater attention. Nonetheless, there is insufficient research done on the high frequency of English words that native speakers of the language speak and write. This is especially true with English textbooks released in Pakistani circumstances. This prompted the researchers to look at the vocabulary items in the SSC English books and how well they were covered in the most widely used English dictionary. The purpose of this study is to examine the vocabulary and grammatical integrity employed in Khyber Pakhtunkhwa English basic texts. The links and interconnections that exist between paragraphs and words within a text are referred to as lexical collaboration (Habib et al., 2024). It is the technique of using conjunctions like hyponyms, metonyms, and comparable terms to give a written or spoken discourse consistency and unity. Grammatical integration is the use of language structures and methods to provide a document with clarity and unity.

Statement of the Problem

This work aims to conduct a comprehensive lexicographical analysis by analyzing phrases extracted from the SSC English textbooks of Khyber Pakhtunkhwa. Gaining a comprehensive understanding of the usage, context, and linguistic writing qualities of these phrases will contribute to developing language learning tools and enhancing language education methods in the given region. The present inquiry serves as the focal point of investigation. This study uses corpus analysis to lighten the underlying patterns, textual analysis, embedded within these linguistic phrases, enriching the lexicographic tools accessible to educator.

Research Objectives

To identify prevalent phrases in the SSC English textbook used in Khyber Pakhtunkhwa.

To develop a lexicographic model to highlight the corpus variations of the identified phrases.

Research Questions

What are the most frequently used phrases in the English textbooks of Khyber Pakhtunkhwa SSC curriculum?

How lexicographical presentations of these phrases enhance students writing skill?

Significance of the Study

The importance of this study lies in its capacity to unveil lesser-known language patterns and cultural nuances evident in the SSC English textbooks of Khyber Pakhtunkhwa. This study aims to get a deeper understanding of the decision-making process for vocabulary selection among curriculum planners and educators. This is done by employing lexicographic tools to analyze and evaluate various phrases. These judgments demonstrate the cultural and communicative interests of the region, as well as its preferences in terms of language. Curriculum designers and educators need to possess knowledge of these idiosyncrasies in language, as it enables them to create learning materials that are more impactful and suitable for the given environment.

Limitations and Delimitations

The limitations of this research include the possibility of resource constraints preventing access to and analysis of a sizable corpus of SSC English literature. The boundaries guarantee the study's breadth and viability by concentrating on particular grades, media outlets, and language characteristics within KP's learning environment (Misnawati et al., 2024). In summary, this corpus analysis study's qualitative methodology provides a thorough way to examine Khyber Pakhtunkhwa's SSC English text themes. The study intends to deepen the comprehension of language learning techniques and social factors in KP's academic landscape by combining textual examination, content assessment, and corpus-based approaches.

Literature Review

A comprehensive review of language pedagogical and socioeconomic variables is included in the investigation of the English language in the setting of the province of Khyber Pakhtunkhwa (KP). The main factors influencing phrase choice and usage in KP's SSC (Secondary School Certification) English texts are examined in this study of the literature (Wang, 2013). Numerous definitions are possible for a corpus. However, a corpus of literature is now generally acknowledged to be a grouping of (1) machine-readable, authentic texts, (2) representative samples of a particular language or linguistic variant, and (3) authentic texts (Aarts & Meijs, 2022). The same holds for corpora, which are computer-stored collections of textual or audio content. A corpus is an organized collection of writings that can be utilized for both qualitative and quantitative research, Baker (2010). Academics' perspective on the word studied. While databases are frequently utilised to offer more thorough descriptions of linguistic use, few scholars have directly used their information to critically assess contemporary textbooks and tools for teaching English as a foreign language (TEFL) (McEnery et al., 2006). In language training, one simple but important role of corpora is to provide more authentic examples of word usage that capture the nuances and complexity of real conversation. Nevertheless, corpora may also provide information that modifies the input provided in class, especially with relation to frequency. This is predicated on the notion that the most useful terms for pupils are also those that are most often used in a language. Different kinds of corpora may exist. The following corpora are listed by Sinclair (1995) and Ahmed et al., 2024.

1. References The corpora: These collections of data, such as BNC, are intended to offer extensive linguistic knowledge.
2. Monitor the Corpus: Constantly growing in size, it is continuously updated with fresh content, with older content being relocated to archive storage, for example. BOE (Baker, 2010 Afaq et al., 2022).
3. Parallel corpus: A group of texts, such as the CRATER Corpus, that have been individually translated into many different languages.
4. Comparative Corpus: Texts that is similar across multiple language varieties, such as ICE Clusters and Brown the Family Corpora. Unlike classical lexicography, which aims to situate word meanings upon a model of a linguistically independent ontological idea (e.g., by species proximal and differentia specific), the corpus approach attempts to record different practices in their link to distinct circumstances. The text supplied in most textbooks as a resource material could be more helpful. Instead, it is based on the authors' or teachers' assumptions, and the language chunks are not contextualized. When a language is taught in its actual usage context, learners learn it more successfully. The gap between what students learn in school and their outside-of-school experiences may also be narrowed with the use of corpora if textbook and instructional material designers are educated via corpus-oriented research and are instructed to make the necessary revisions (Cobb and Boulton, 2015). Additionally, with the aid of corpora, useful teaching tools for the instruction of the English language can be developed, including lesson plans, course books, vocabulary books, a glossary of standard terms, grammar books, and many more. To learn a language, students can obtain corpus-based resources such as worksheets, exercises, and other assignments (McEnery & Xiao, 2011 and Khan et al., 2023). These exercises are accessible in print and online and are called data-driven learning (DDL) activities. The investigation of lexical linguistics across several academic disciplines has been the subject of scholarly inquiry among university students (Durrant, 2017). A corpus has been compiled, including several subjects, including science and technology, humanities and social sciences, life sciences, and business. Substantial disparities were identified among these domains. Furthermore, researchers have identified variations among writers within specific disciplines (Esfandiari & Barbary, 2017). Furthermore, it has been shown that a significant proportion of the disciplines exhibit a high degree of internal homogeneity. The investigation conducted by Yurievna, Antonovna, and Vasilievna (2022) focused on examining lexical bundles found in textbooks and lectures within university settings. The researchers integrated textbooks derived from the field of natural sciences into their dataset. Observable changes in lexical bundles between textbooks and lectures were identified. A study examining lexical bundles in physics research articles (Hussain, Zahra, & Abbas, 2021). The researchers methodically

identified lexical bundles, but they needed to give more consideration to their discourse roles. A study has investigated the presence of anticipatory it lexical bundles in research articles within the fields of linguistics and chemistry (Birhan, 2021). Chambers (2019) conducted a corpus-based study in order to investigate the vocabulary employed in literature targeted toward children. In order to ascertain if there are discernible linguistic characteristics that distinguish writing meant for children from writing intended for adults, scholars conducted a comparative analysis of the vocabulary employed in the CLLIP corpus and a corpus of adult literature and newspaper articles. The researchers determined that, although there are distinct dissimilarities between the vocabulary profiles found in news text and children's literatures, there are indeed several shared characteristics with the language employed in adult fiction. A notable degree of intersection existed among the highest-ranking lexical verbs, adjectives, and nouns in both corpora. The study conducted by Alexiou, Mattheoudakis, Saratsli, and Vagenas 2019 and Gul et al., 2022(b)) investigate the lexical content of five English as a Foreign Language (EFL) textbooks utilized in Greece throughout the initial two years of primary education. This research contributes to the existing body of literature that examines vocabulary acquisition within a classroom context. The argument is that the Greek texts lack an adequate number of mid- and low-frequency terms, which stands in striking contrast to the viewpoints expressed by researcher (Norberg & Nordlund, 2018). Kulas (2012) and (Gul et al., 2023) conducted a study investigating educational resources lexical composition. The researcher examined seven introductory English textbooks designed for first-year learners to ascertain whether there existed any consensus or divergence regarding the extent and substance of vocabulary included. Comparable to the results of Alexiou et al. (2019), the researcher observed notable disparities in the perceived optimal amount of words to expose children to during their initial year of English language acquisition. The research conducted by the author also revealed a need for more consensus about the lexical selections employed in the analyzed texts. A researcher looked at how vocabulary is learned in four English as a Foreign Language (EFL) course that are often used in Spanish primary and secondary schools (Catalán & Francisco, 2008; Nasir et al., 2023; Afaq et al., 2023). The primary argument of the authors is that using corpora and corpus-based analyses greatly improves the creation of efficient language teaching resources and methods. The authors argue that by incorporating corpus-based insights into language instruction, the gap between classroom learning and real-world language usage can be bridged, resulting in more contextually relevant and comprehensive language learning experiences (Baker, 2010; Ishtiaq et al., 2022). Concerning the teaching of languages, the English language literature is significant, so for the secondary school pupils. Inasmuch as the significance of carrying out corpus assessment and the possible consequences of such assessment on the English textbooks meant for SSC students in the educational system of Khyber Pakhtunkhwa cannot be overemphasized, it is the aim of this review of literature to explore the research questions that ensue from the proposed study. Teaching focus is on the lexicogrammatical features, with the purpose of revelation of morphological and phrasal patterns shared in a number of teaching aids.

Lexicographical Examination of Course Expressions

This study's lexicographical portion deals with gathering and evaluating lexical materials from SSC English courses. The goal of lexicography is to build and enhance dictionaries and language tools through the study of terminology and phrases. Lexicographical research will concentrate on cataloguing and evaluating the terms and expressions used in the SSC English texts of KP, taking into account their prevalence, meaning connections, and learner-friendliness (Birhan, 2021).

Relevance in Contextual and Its Effects on Training

For language instruction, it is important to comprehend a collection of terms found in Khyber Pakhtunkhwa-specific SSC English course. First of all, it offers perceptions of the language needs and inclinations of pupils in this area, enabling the creation of curriculum and teaching methodologies adapted to the local circumstances. Second, lexicographical studies help teachers create focused language resources and lesson plans that match students' language acquisition goals and competence levels.

1) Word Choice.

Because many words have been borrowed from other languages, Johnson had to choose which terms to include to the dictionary and categorize each one as either foreign or English. Additionally, he had to determine whether words from particular occupations belonged in the dictionary.

2) The Practice of Orthography.

Johnson said that word spellings should never be changed without a good reason, as doing so would only annoy other people and be seen as a sign of weakness or inconsistent thinking (Baker, 2010).

3) Enunciation.

According to Johnson's, pronunciation should remain consistent in addition to orthography since language stability is essential to a language's longevity and any alterations would result in virtually entirely new speech, which would contaminate spoken English at the moment.

4) Variation and Terminology.

With so many words borrowed from other languages, it can be challenging to identify which terms are original to English, therefore understanding word's etymology is crucial.

5) Comparison.

Included are the guidelines that controlled the usage of the terms.

6) Grammar.

The way words are constructed in English is displayed because the language's composition is so irregular that it would be challenging to distil it down to a set of principles.

7) The Grammar of Words.

The word's examples of usage are provided to show the various contexts in which it can be employed.

8) Analysis.

Johnson believes that the most challenging aspect of producing a dictionary is interpreting a word, as it requires careful consideration of the various applications of each word and the best possible explanation.

9) Dispersal.

Following the completion of the aforementioned procedures, Johnson had assigned each word to its appropriate class. Johnson's dictionary, which served as the primary source of reference for over 150 years and underwent multiple modifications, was found to fall short of the norms of contemporary scholarship (McEnery & Banerjee, 2019).

Methodology

Based on certain methodological constraints, the current study exclusively uses the preface sections of English textbooks from two major domains: SSC and bachelor's level. Because there are so few textbooks on sale, many disciplines may not be included in the data. Finally, there is a discrepancy in the quantity of prefaces included in the textbooks. For instance, the English textbooks are sourced from the Khyber Pakhtunkhwa Textbook Board. The maximum number of subjects that can be included in an English textbook preface is three. The study's findings are evaluated in the context of the information at hand. Software called AntConc was used to capture frequencies and analyses concordance lines (Habib & Umar, 2017).

Research Design

The SSC (The Secondary School Certification) English textbook keywords within KP are the subject of a thorough corpus analysis conducted using a qualitative research design. The study examines the lexical elements and cultural nuance encoded in the instructional materials using a variety of qualitative techniques, such as textual analysis, content evaluation, and corpus-linguistic tools. Mixed Methodology Research (MMR) is selected for the study because it intends to conduct a thorough investigation into the distribution and features of lexical elements presented in the English language program used in KP. MMR is "the broad types of research in which elements or methods from both qualitative and quantitative studies are combined or blended in a research study," according to Briggs, Coleman, Morrison (2012) and Gul et al., (2024). The study's produced qualitative and quantitative findings provide to a deeper understanding of the connection between KP English textbooks' lexical coverage and the degree to which they give students opportunities to

learn vocabulary. As a result, the current study used an explanatory sequential Mixed Method design, which starts with the gathering and interpreting of quantitative data before going on to the gathering and analysis of data that is qualitative. This made it possible for the researcher to gather more specific data about how many lexical elements were covered. The purpose of this study will analyze the linguistic content of SSC English textbooks in Khyber Pakhtunkhwa (KPK), Pakistan, using a corpus analysis methodology. A qualitative research design will be chosen in order to provide a comprehensive understanding of the linguistic and cultural context (Misnawati et al.2024; Gul et al., 2023).

Qualitative Investigation Advantages and Disadvantages

One of the many advantages of qualitative data collection is its capacity to offer in-depth, nuanced insights into intricate phenomena, such as language use throughout learning environments. It permits researchers to investigate new patterns and meanings by providing flexibility in data collecting and analysis. However qualitative exploration can also be subjective and costly, requiring careful evaluation of information and sometimes prejudiced researcher viewpoints. Qualitative research gives a unique view of the complexity of the human experience by looking into the reasons and processes that drive actions and events. It has the ability to give us insights that go beyond just numbers and that cannot be reached by using only quantitative methods (Choy, 2014; Gul et al., 2022(a)). Methods like in-depth interviews, participant observation, and text analysis are used in qualitative research to help people understand complex social and cultural events. The fact that there are biased points of view adds to the difficulty and possibility for bias in this situation. To make sure a research method is complete, it is important to find a balance between how precise quantitative analysis is and how detailed the insights are that come from qualitative ways. Limitations in extrapolating from qualitative data may be due to their narrative structure and smaller sample numbers.

Data Analysis

The data analysis in this study delves into a systematic examination of the SSC English textbooks used in Khyber Pakhtunkhwa (KP), aiming to unveil the linguistic and educational patterns embedded within these materials. The analysis adopts a dual approach, combining qualitative and quantitative methodologies to comprehensively understand the frequency, context, and implications of common phrases and vocabulary items within the textbooks. Central to this analysis is the utilization of AntConc, a software application specifically designed for corpus linguistics research. AntConc serves as the primary tool for data analysis, offering features that facilitate the extraction of concordance lines, frequency lists, and collocational patterns. These features enable a meticulous investigation of the lexical elements present in the textbooks, shedding light on the nuances of language use and educational content. The analysis process is structured and meticulous, encompassing several key steps. Firstly, the textbooks are systematically examined to identify common phrases and vocabulary items. These are then entered into AntConc for further analysis. Through AntConc, the frequency of each word and phrase is determined, providing insights into their prominence within the texts. Additionally, the context in which these words and phrases appear is closely examined. By analyzing concordance lines, researchers can understand how words are used in different contexts, uncovering the underlying themes and messages conveyed through language. Furthermore, the implications of these linguistic patterns are explored. Researchers seek to understand how the choice of words and phrases influences the overall tone and message of the textbooks, as well as how they reflect broader societal and educational trends. Overall, the data analysis in this study offers a comprehensive and detailed examination of the SSC English textbooks used in KP. Through the use of AntConc and a meticulous approach to analysis, researchers aim to uncover the intricacies of language use and educational content within these materials, providing valuable insights for educators, curriculum developers, and policymakers.

Figure 4.1

Rank Phrases Frequency

1. Hazrat Muhammad 120.
2. Allama iqbal 90.
3. Quaid e 80.
4. Holy Quran 76.
5. Prophet Muhammad 72.
6. Muslim community 68.
7. Islamic teachings 66.
8. According to 64.
9. Young generation 60.
10. Great leader 58.
11. The Quran 54.

Textual Analysis

The analysis of the SSC English textbooks used in Khyber Pakhtunkhwa (KP) reveals a rich tapestry of themes and values embedded within the educational content. Key phrases extracted from the

textbooks highlight significant cultural, moral, and educational messages that are repeatedly emphasized to shape students' character and worldview.

Contextual Analysis

In Islamic texts, *Hazrat Muhammad* is referenced extensively as the last prophet. Concordance analysis shows his name often appearing in sentences emphasizing his teachings and virtues, such as 'Hazrat Muhammad emphasized honesty,' reflecting his role in guiding moral and ethical conduct. His life and sayings serve as a model for Muslims worldwide. Allama Iqbal frequently appears in concordance lines in the context of his poetry and philosophical thoughts. Sentences like 'Allama Iqbal's vision for youth empowerment' highlight his influence on cultural and intellectual spheres, inspiring socio-political movements. His works focus on self-awareness, spirituality, and awakening. The term Quaid-e appears predominantly in references to Quaid-e-Azam Muhammad Ali Jinnah. Concordance analysis indicates its use in discussions about leadership, as in 'Quaid-e-Azam led Pakistan to independence,' underscoring his pivotal role in the nation's history. He is celebrated for his vision and determination in founding Pakistan.

Conclusion/ Findings

This study was designed in such a way so that it could comprehensively carry out the corpus based attribution style of SSE English textbooks used in KPK. Under this them, the area covering linguistics and education was examined in-depth. With the help of corpus based techniques, the research has disclosed the complex features of the textbooks, which simultaneously promote a child beyond merely a language attention to an immersion into the morals and cultures.

Research Objective Achieved

The main task that must be done in the course of achieving this purpose is to carefully select and analyze the corpus of textbooks in use. The choice was these textbooks based on their widespread usefulness and the degree of their approval in the education system of KP. An analysis of different linguistic elements, such as phrase, word, and thematic frequencies, was implicated in the process of the investigation. This methodological glance granted a detailed overview of the linguistic patterns and cultural components that the educational literary works on read. The research demonstrated the frequent use of certain words and the frequent mention of the Prophet Muhammad, Allama Iqbal, Sid Field as well as the first statesman Quaid-e-Azam. These vermin are not only the wonderful historical and religious individuals but also have a great moral. Families and communities in the used textbooks apply these images in the moral and cultural raising of their children. Thus the learning process is not only

English but also it protects the culture base of students.

Research Question Answered

The central research question posed by this study was: What is the relationship between the way SSC English textbooks in KP express patterns and the cultural as well as moral courses of education? Through investigating both the content of linguistic and cultural knowledge in these textbooks, this research made it possible to answer this question in the depth. The analysis, however, shows that the textbooks usually mention Majid-e-Pakistan, Allama Iqbal and Muhammad amongst others. These bodily pictures have a sacred and/or a religious personality in them, but they mainly pass down moral and ethical principles. The curriculums get augmented with such characters who are then used throughout the books as a medium by which students get exposed to important cultural myths and values. This approach serves a dual purpose: mainly teaching English but also instilling other cultural and normative values. The similar expressions used here e.g. 'Hazrat Muhammad said' and 'the holy Qur'an,' act to impart both the language and morals in the narration. Integrated learning. On this account CLIL facilitates, concurrently, acquiring language skills and subject-specific knowledge directly, which, in turn, promotes holistic educational method. Through the magenta of cultural education within the framework of language learning, students are able to accumulate more knowledge about English than that but, most importantly, they are able to absorb cultural values and stories.

Recommendations

Based on the results of this study, numerous recommendations could be suggested not only to improve academic productivity, but also the level of students' learning experiences in Khyber Pakhtunkhwa province. The recommendations seek to improve educational resources like curriculum

and implementation of proper points to ensure that the students can excel academically and also in a personal manner.

Expanding Corpus-Based Studies

Research should go the full length and exhaust corpus based strategies for educational purposes. Examining a more prevalent and more wide-ranging corpus; researchers will be able to get an even better perspective about language and teaching. This method holds also the potential of rolled out to other regions and educational environments while presenting the complete picture of how language is used across the different settings.

References

- Aarts, J., & Meijs, W. (2022). Corpus linguistics: Recent developments in the use of computer corpora in English language research. *Journal of Language Teaching and Research*, 9(3), 27. Retrieved from [Academypublication.com](https://www.academypublication.com)
- Ahmed, S., Waqar, A., Habib, S., & Gul, N. (2024). The Reflection of Mystical Elements in the Poetry of William Butler Yeats. *International Journal of Social Science Archives (IJSSA)*, 7(3).
- Afaq, M., Shah, M., & Gul, N. (2023). The Role of Henchard's Character in Loss and Profit of His Business in the Novel the Mayor of Casterbridge. *Pakistan Languages and Humanities Review*, 7(2), 318-327.
- Afaq, M., Sabih-ul-Hassan, S., & Gul, N. (2022). A Linguistic Critique on Naturalism in Thomas Hardy's Novel 'The Mayor of Casterbridge'. *Journal of Education and Social Studies*, 3(2), 143-150.
- Alexiou, T., Mattheoudakis, M., Saratsli, D., & Vagenas, A. (2019). Words don't come easy: Linguistic analysis of vocabulary in Magic Books. *Journal of Applied Linguistics*, 54(32), 25–41. <https://doi.org/10.9779/pauefd.798049>
- Aqeel, M., Rubbani, A., & Hamid, A. (2022). Error analysis in learners' essay writing: A corpus-based study. *International Journal of Special Education*, 37(3), 18331–18341. Retrieved from ResearchGate
- Ayu, M., & Inderawati, R. (2019). EFL textbook evaluation: The analysis of tasks presented in English textbooks. *Teknosastik*, 16(1), 21–25. <https://doi.org/10.33365/ts.v16i1.87>
- Baker, P. (2010). *Corpus methods in linguistics*. (93). Cambridge University Press.
- Balouch, A., Ahmed, R., & Channa, S. A. (2022). A multidimensional analysis of secondary level textbooks used in Sindh Board and Federal Board, Pakistan. *Kashmir Journal of Language Research*, 25(2), 73–94.
- Birhan, A. T. (2021). Effects of teaching lexical bundles on EFL students' abstract genre academic writing skills improvement: Corpus-based research design. *International Journal of Language Education*, 5(1), 585–597.
- Briggs, A. R. J., Coleman, M., & Morrison, M. (2012). *Research methods in educational leadership and management*. (3rd ed., pp. 1–400). SAGE Publications. <https://doi.org/10.13134/rmel.2012.123997>
- Carter, R. (1998). *Vocabulary: Applied linguistic perspectives*. (2nd ed., pp. 1–350). Routledge. <https://doi.org/10.vr36743/valp.1998.1thrt544>
- Catalán, R. M. J., & Francisco, R. M. (2008). Vocabulary input in EFL textbooks. *Spanish Journal of Applied Linguistics*, 21(21), 147–166. Retrieved from Dialnet
- Chambers, A. (2019). Towards the corpus revolution? Bridging the research–practice gap. *Language Teaching*, 52(4), 460–475. <https://doi.org/10.1017/S0261444819000089>
- Channa, L. A. (2017). English in Pakistani public education: Past, present, and future. *Language Problems and Language Planning*, 41(1), 1–25.
- Choy, L. T. (2014). The strengths and weaknesses of research methodology: Comparison and complimentary between qualitative and quantitative approaches. *IOSR Journal of Humanities and Social Science*, 19(4), 99–104.
- Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), *Cambridge Handbook of Corpus Linguistics* (pp. 478–497). Cambridge University Press. <https://doi.org/10.1017/CBO9781139764377.027>
- Cowie, A. P. (1999). *Phraseology: Theory, analysis, and applications*. (pp. 1–250). Oxford University Press. <https://doi.org/10.fa224/ptaa.1999.16432>

- Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing.
- Esfandiari, R., & Barbary, F. (2017). A contrastive corpus-driven study of lexical bundles between English writers and Persian writers in psychology research articles. *Journal of English for Academic Purposes*, 29, 21–42. <https://doi.org/10.1016/j.jeap.2017.09.002>
- Gul, N., Farooq, A., & Ali, A. (2022(a)). Saussure's Concept of Binary Oppositions: Looking through the Lens of the Holy Qur'an. *Al-Azhār*, 8(01), 111-119.
- Gul, N., Sabih-Ul-Hassan, S., & Imran, S. (2022(b)). A linguistic analysis of the impact of Covid-19 on English language diction in Pakistan. *Pakistan Journal of Social Research*, 4(2), 1010-1017.
- Gul, N., Ali, A., Hassan, S. S. U., & Rasheed, A. An Investigation into the Politics of English Language Assessment in Pakistan.
- Gul, N., Hassan, S. S. U., & Wasti, A. T. (2024). A Study of Linguaging in English Language Classroom Discourse (ELCD) in Pakistani Context: From Theory to Practice. *Shnakhat*, 3(2), 236-246.
- Gul, N., Imran, S., & Wasti, A. T. (2023). A Study of the Bidirectional Causality Relationship Between English Language and Environment: Looking Through the Lens of Ecological Discourse Analysis (EDA). *Journal of Development and Social Sciences*, 4(4), 114-125.
- Gul, N., Khatoon, S., & Hassan, S. S. U. (2022(c)). An Investigation into the Role of Language in Regional Connectivity in Afghanistan and Pakistan. *University of Chitral Journal of Linguistics and Literature*, 6(I), 341-348.
- Ha, M. (2020). Exploring academic vocabulary in advertising English corpus. (57), 223–243. Retrieved from DBpia
- Habib, A., Ali, M., & Gul, N. (2024). An Analysis of Politeness Strategies in the Speeches of King Abdullah II of Jordan at the United Nations General Assembly. *International Journal of Social Science Archives (IJSSA)*, 7(3).
- Habib, A., & Umar, H. (2017). A qualitative inquiry into the alignment of English textbooks with the national curriculum at secondary level. *NUML Journal of Critical Inquiry*, 15(2), 171–XIV. Retrieved from ProQuest
- Hawtin, A. (2019). The written British National Corpus 2014: Design, compilation and analysis. Lancaster University (United Kingdom).
- Hinkel, E. (2002). Second language writers' text: Linguistic and rhetorical features. (pp. 1–300). Lawrence Erlbaum Associates. <https://doi.org/10.refrgf/sltw.2002.i99r43>
- Hornby, A. S. (1954). A guide to patterns and usage in English. (pp. 1–200). Oxford University Press. <https://doi.org/10.34345g/gpue.1954.12984>
- Hunston, S. (2002). Corpora in applied linguistics. (pp. 1–275). Cambridge University Press. <https://doi.org/10.576dft/cial.2002.12hjy55>
- Hunston, S., & Francis, G. (2000). Pattern grammar: A corpus-driven approach to the lexical grammar of English. (pp. 1–240). John Benjamins Publishing Company. <https://doi.org/10.1424434/pgcd.2000.ju754>
- Hussain, G., Zahra, T., & Abbas, A. (2021). Discourse functions of lexical bundles in Pakistani chemistry and physics textbooks. *GEMA Online Journal of Language Studies*, 21(1), 221–238. <https://doi.org/10.17576/gema-2021-2101-13>
- Ishtiaq, M., Gul, N., & Khan, Y. (2022). English to Urdu Transliteration AS A Major Cause of Pronunciation Error in L1 & L2 Urdu Speakers of English: A Pedagogical Perspective. *Global Language Review*, VII, 282-298.
- Ishtiaq, M., Gul, N., & Hayat, Q. (2021). Linguistic Analysis of the Gender Representation in Jane Austen's Novel, Persuasion, Using Systemic Functional Grammar. *Global Sociological Review*.
- Johnson, S. (1755). A dictionary of the English language. (Vol. 1, pp. 1–500). W. Strahan. <https://doi.org/10.3r3ur3/del.1755.1o904ef>
- Kövecses, Z., & Csábi, S. (2014). Lexicography and cognitive linguistics. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, 27(1), 118–139.
- Kulas, K. (2012). The selection of vocabulary for EFL lower-primary school textbooks. Paper presented at the 10th Teaching and Language Corpora Conference, The Institute of Applied Linguistics, University of Warsaw, Warsaw, Poland.

- Khan, A. A., Khan, I. A., & Gul, N. Exploring Personal and Ethnic Identity in Ernest Hemingway's *The Sun Also Rises*.
- Lew, R. (2024). Dictionaries and lexicography in the AI era. *Humanities and Social Sciences Communications*, 11(1), 1–8.
- Li, Q., & Ng, Y. (2024). Effective Chinese-to-English biotic interpretation in ecotourism destinations: A corpus-based interdisciplinary study. *Journal of Sustainable Tourism*, 1–31.
- Lipinski, P. (2010). Language learning strategies in the classroom. *Journal of Applied Linguistics*, 12(3), 45–67. <https://doi.org/10.19764/jal.2010.1456>
- Mahlberg, M. (2006). Lexical cohesion: Corpus linguistic theory and its application in English language teaching. *International Journal of Corpus Linguistics*, 11(3), 363–383.
- McEnery, T., & Banerjee, J. (2019). Corpus linguistics, learner corpora, and SLA: Employing technology to analyze language use. *Annual Review of Applied Linguistics*, 39, 74–92. <https://doi.org/10.1017/S0267190519000096>
- McEnery, T., & Xiao, R. (2011). What corpora can offer in language teaching and learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 364–380). Routledge.
- Nasir, S., Gul, N., & Ullah, A. (2023). Existential Elements in the Character of the Underground Man. *Pakistan Languages and Humanities Review*, 7(2), 185-195.