



## RESEARCH ARTICLE

# An empirical investigation of Linguistic Errors in a corpus of sixteen doctoral theses submitted to CHARUSAT to improve lexical repertoire and quality of Academic Writing

Prince Grover<sup>1\*</sup>, Bhaskar Kanaiyalal Pandya<sup>2</sup>

## Abstract

The academic writing is grounded in skillful and proficient incorporation of language-based systems which reinforce the expression of argumentative discourse, conceptual knowledge and discipline-oriented scholarly communication. The earlier scholarly investigations mainly focused on discourse-related patterns and grammatical systems and perceived lexical repertoire as ancillary constituents. Despite this, the systematic and structured utilization of lexical resources contributes substantially in developing semantic meaning and supporting logical coherence in research writing.

The current empirical investigation probes vocabulary-related linguistic resources through the empirical evidence by evaluating language-oriented systematic patterns in the dataset of sixteen doctoral dissertations submitted at Charotar University of Science and Technology (CHARUSAT), Changa between 2016 and 2020. It concentrates on sentence-structural arrangement, discourse-structural lexical terms, grammatical correctness, and lexical appropriateness. By investigating the sixteen dissertations using error-based analyses and lenses, the inquiry evaluates lexical-semantic selections facilitating argumentative discourse, writer's position and academic discourse logical coherence in research-oriented exchange.

The research study in the next stage explores cognitive, conceptual, knowledge-centered, and discipline-focused components pertaining to lexis-related deployment coupled with the acknowledgement of empirical results and implications leading to influence the pedagogical interventions in designing a layout of the training program and strategizing refinement and enhancement.

The present scholarly investigation points out the prominence of vocabulary-based expertise and capability in disciplinary textual discourse. It also suggests pedagogical implications in building and strengthening lexical resources understanding, discipline-specific linguistic systems and lexis-based formulaic sequences in higher-level research writing.

**Keywords:** Academic writing; Linguistic errors; Doctoral theses; Lexical repertoire; Empirical investigation.

<sup>1</sup>Researcher Scholar at Charotar University of Science and Technology,

<sup>2</sup>Research Guide and Professor of English, Principal at Charotar Institute of Languages, Arts and Social Studies (CLASS), Dean at Faculty of Humanities (FOH), Charotar University of Science and Technology (CHARUSAT), Changa, Anand, Gujarat, India

**\*Corresponding Author:** Prince Grover, Researcher Scholar at Charotar University of Science and Technology, E-Mail: 23drfoh003@charusat.edu.in

**How to cite this article:** Grover, P., Pandya, B.K. (2026). An empirical investigation of Linguistic Errors in a corpus of sixteen doctoral theses submitted to CHARUSAT to improve lexical repertoire and quality of Academic Writing. *The Scientific Temper*, 17(3):5878-5889.

Doi: 10.58414/SCIENTIFICTEMPER.2026.17.3.19

**Source of support:** Nil

**Conflict of interest:** None.

## Introduction

Academic writing emphasizes on linguistic-discursive meticulousness and comprehensibility in the articulation of scholarly as well as academic understanding. Conceptual coherence and linguistic cohesion in written discourse is typically established through the appropriate and proficient deployment of lexical bundles, conventional word combinations and subject-specific lexical repertoire. In some studies, the mere emphasis is given on the grammar development and advancement that leads in the enhancement of the grammatical accuracy, but it lacks in connectedness, structuring arguments, and taking authorial stance. The English for Academic Purposes (EAP) insists on the principles of keeping it at the central point by linking it with applied studies in linguistic domains. It is essential to have the integration of language-related practices, patterns and conventions to meet the requirements and expectations while shaping academic writing. The competence to

articulate in an academic style necessitates not limited to grammar-based exactness and formal cohesion, but it also incorporates the capability to depict sophisticated notions with the navigation of domain-specific and subject-oriented understandings as well as authorial positioning in the intellectual dialogues.

The background of language teaching is indeed an ancient concept. As early as the civilizations came into existence, language learning appeared. The communicative learning and teaching approaches are interpreted by means of integrated frameworks and mechanisms that are commonly labelled in the role of lexis, phonological structure, grammatical structure and discourse application. The mentioned frameworks as a whole underpin communicative mastery as well as meaning-construction. On the opposite side, the integration of grammar, vocabulary and discourse have gained more prominence and visibility due to rising demands of their applications in academic writing. The appearance of words in the specific contexts become closely involved with taking authorial positioning, structuring an argument and enhancing transparency levels by using the most appropriate vocabulary words in the specific domains and working on the enhancement of quality in articulation. The core status of vocabulary is always enshrined in educational expression including scholarly writing as well as academic writing. Notwithstanding the presence of wide-ranging and comprehensive conceptual and theoretical scholarly discussions on vocabulary-based inventory in EAP, the modest and inadequate corpus-based scholarly attention has been devoted to the exploration and assessment of vocabulary resources patterns in doctoral-stage academic composition.

The academic discourse is typically distinguished by theoretical and cognitive compactness with an emphasis on argumentative clarity with domain-oriented precision. All it requires is a balance of the effective and the most appropriate vocabulary. The lexical repertoire is determined by the way in which learning is articulated and the suitable and the appropriate ways in which the lines of arguments are structured. The effectiveness of vocabulary makes it possible to write concepts more precisely with judgement-based and interpretive arguments.

Considering the practices of language systems particularly grammar is more prevailing where adequate attention is unpaid for discourse, lexis and phonology. The rearticulation of word stock in the capacity of language system calls for transitioning by not looking into the quantity of words instead of quality of words. The ordered and interconnected structure of lexico-semantic tools that encompass habitual co-occurrences of lexical chunks or collocations, expression-based configurations with noun-based constructions along with subject-oriented specialized lexis. Quality negotiation can easily be avoided with effective use of vocabulary.

On the side of phonology, it is explicit that it has a very limited role in academic writing. The speech sound patterns hold a proportionally narrow function measured with lexical resources. Even though the speech production along with suprasegmental characteristics and aspects are crucial within speech-based or spoken exchange. While writing, the academic discourse is chiefly based on vocabulary-related and sentence structure tools. Moreover, the role of lexis in the specific domain of the language system shapes academic writing extensively.

In order to respond to the research gap, this investigation scrutinizes language-related recurring patterns in sixteen examined PhD-level dissertations to study how lexis-based repertoire serves within the context of scholarly communication. The assessment of lexical-semantic specificity, grammatical frameworks as well as discourse-related lexical terms application, this inquiry delivers data-driven and evidence-based observations and interpretations into the contribution of vocabulary frameworks in scientific writing.

The ongoing inquiry advances with the view that the vocabulary-related resources function as essential and crucial positions in the language system of academic scientific or academic writing. By means of corpus-based and data-driven of sixteen doctoral dissertations, it studies how lexical repertoire performs in shaping argumentative structures, preserving textual discourse, logical coherence, and fostering academic knowledge dissemination in academic and disciplinary settings.

### Literature Review

The inquiry concerning lexical repertoire has close relationship with EAP as it leads to positioning language approximately and more specifically. A number of theories established in the language domain strengthen the foundations of EAP. The early phases of the language system were more centered on formal and structural orientations as well as approaches, but nowadays an inclination toward communication-orientation, discourse-focused and research-based scholarly literacy conceptual frameworks. The linguist Halliday reconceived the language system in the capacity of a framework of alternatives, possibilities and selections whereby the word stock or terminologies together with grammatical structure perform collaboratively to develop cognitive social, interactional, and discourse-based interpretations (Halliday, 1978, 1985; Halliday & Matthiessen, 2014).

The corresponding progressions within lexico-semantic inquiries additionally brought to the fore of lexis by keeping it at a vital facet of linguistic systems capability. Ian Stephen Paul Nation (I.S.P. Nation) accentuated the relevance of lexical repertoire, scope of thoroughness, and context-based application within language acquisition

resulting in language competency (Nation, 1990, 2001, 2013). Correspondingly, the linguist Norbert Schmitt developed the comprehension and awareness of vocabulary-related mastery in the form of underscoring and foregrounding co-occurrence-related/collocations sensitivity, vocabulary-related interconnected structures along with purposeful and systematic lexical resources acquisition and development (Schmitt, 2000, 2010). Various inputs and efforts as a whole reoriented and reframed lexical resources from the initial and minor pedagogical focus to a core and essential component of expressive proficiency (Nation & Webb, 2019; Webb & Nation, 2017).

The data-driven and text-based linguistic studies established a notable procedural and model-based transformation within lexical inquiry. The linguist John Sinclair proved that the linguistic system is organized systematically and also in expression-based sense and interpretation through word groupings and associations (Sinclair, 1991). The ensuing data-driven examinations by the linguist Douglas Biber and co-researchers recognized widespread presence of vocabulary-related sequences and conventional patterns and arrangements within research-based and academic textual interaction demonstrating how those configurations sustain connectedness, authorial stance or positioning and argument structuring (Biber et al., 1999; Biber et al., 2004). The linguist Susan Huntson foregrounded the significance of the linguistic dataset in raising awareness of the functionality of the vocabulary embedded in uncontrived and original discourse settings. The aforementioned examinations supported the standpoint of vocabulary in a formulaic and system-based resource (Wray, 2002). The standalone application of vocabulary is not advised and the assemblage of such detached words have limited scope of enhancing quality in the academic writing context by fostering the effective use of rhetorical tools and discourse management in raising coherence and cohesion (Hyland, 2006; Hyland, 2018).

The intellectual investigation and formal academic English essentially require an extensive use of communicative practice, lexical repertoire, grammatical structures, and discipline-oriented approaches to connect all systems altogether for effective academic writing production. The linguist John Swales paid more importance to rhetorical genre analysis and emphasized on keeping genre study as a core element in indicating the extent to which lexical preferences can express academic norms, discourse goals, rhetorical functions and discipline-oriented practices (Swales, 1990; Swales & Feak, 2012). One of the linguists known as Key Hyland who laid a strong foundation of a mediation point where vocabulary, rhetorical devices, authorial stance, argument structuring and discourse management connect the discipline with meaning-making and knowledge development (Hyland, 2000, 2006).

The university-level and disciplinary academic literary practice in the research delivered an additional key component to the study of vocabulary repertoire. Another linguists Brian Street and Mary Lea modelled scholarly writing and academic writing in the form of societal and interactional engagement (Street, 1984; Lea & Street, 1998). In fact, the application of scientific writing necessitates the constituents of vocabulary related to specific discipline, intellectual content, and academic networks. Following this, other linguists Mary Scott and Theresa Lillis strengthened and validated the socio-knowledge-based alignment relating to academic literacy frameworks and such structures reveal the manner in which language use and discourse preferences indicate university-based, educational as well as domain-related conventions (Lillis & Scott, 2007).

Another linguist Averil Coxhead also presented a special emphasis on AWL (Academic Word List) and placed it at a primary position for effective research communication (Coxhead, 2000). Ensuing analyses further elaborated and developed the current study through domain-based vocabulary evaluations and dataset frameworks (Coxhead, 2018; Coxhead & Webb, 2021). This way, the studies after another boosted the domains of EAP by placing more emphasis on academic writing and laying strong foundations of EAP (Hyland & Shaw, 2021). Linguists such as Magali Paquot worked wonderfully in the domain by finding the disparities between early-stage learners and the learners at an advanced level using disciplinary terminology and academic lexicon (Paquot, 2010). On the same grounds, Nick Ellis and Rita-Simpson favored the same approach by considering prefabricated chunks, lexical bundles, and fixed expressions as structural and core elements in academic communication and research discourse (Simpson-Vlach & Ellis, 2010; Ellis, 2012).

The contemporary scholarly inquiry incorporates lexicon study alongside disciplinary and academic writing considering its-in-depth engagement in meaning specificity and precision. A number of linguists including Sheena Gardner and Hilary Nesi also presented their research outcomes by emphasizing the role of word stock in a close alignment with genre, discourse category, rhetorical form and cognitive applications (Nesi & Gardner, 2012; Gardner & Nesi, 2013). James Paul Gee stressed by stating that language use, and discourse application in forming cognition and self-concept and social identity affirming a close association linking lexicon with discipline-based involvement (Gee, 2014, 2015). The linguist Christopher Tribble underlined the correlation connecting linguistic dataset inquiry as well as scientific writing educational methodology (Tribble, 2015).

This investigation also explored by broadening the given observations and understandings by means of investigating the rationale of vocabulary by functioning on communicative tools, rhetorical, and cognitive in

educational and scholarly communication. The linguist Sylviane Granger studied language learner datasets for the purpose of comprehending word-based growth and progression during educational settings (Granger, 2015, 2021). The linguist I.S.P. Nation also studied vocabulary mastery mechanisms, stages and operations as well as instructional applications (Nation, 2016; Nation, 2022).

Throughout the given theory-based along with evidence-based paradigms, it is absolutely explicit that vocabulary does not work as supporting and auxiliary elements, but it functions as a fundamental mechanism influencing academic writing (Hyland, 2019; Webb, 2023). It has a significant role to play in EAP. A special emphasis on lexicon in EAP evidences the fact that vocabulary-based and lexico-semantic tools regulate intellectual content expression and articulation, stylistic alignment as well as discipline-based participation (Hyland, 2020; Coxhead, 2024).

Even though the previously mentioned academic inquiries foreground the centrality of lexical resources in academic and scholarly discourse, a significant amount of available research literature continues to be conceptual and theory-based at a broad scale. The corpus-informed analytical assessments of doctoral-level thesiss generate meaningful and insightful observations into how lexical inventory functions in real-world academic writing settings (Charles, 2020). As a result, this inquiry explores linguistic-systematic configurations or patterns in sixteen PhD-level dissertations to investigate the importance of lexis-based frameworks in academic writing.

**Research Methodology**

**Aim**

To investigate linguistic errors in PhD theses submitted to CHARUSAT in the faculties of Engineering, Sciences, Medical Sciences, Computer Science and Application, and Management Studies to improve lexical repertoire and quality of academic writing.

**Objectives**

- To cite existing literature related to needs evaluation of academic writing pertaining to PhD-level scholars.
- To carry out an empirical observation to investigate the theses submitted to CHARUSAT in order to understand constraints encountered by doctoral researchers in academic writing
- To categorize errors committed by scholars in writing PhD theses

**Scope and Limitation**

The scope of the current inquiry is to detect, recognize and evaluate language-related errors notably those related to lexical repertoire as well as vocabulary-grammar elements in PhD-level academic writing. It concentrates on linguistic

system in the role of an independent and autonomous framework examining the manner in which similar errors affect precision, logical flow as well as overall academic communication in research-oriented writing.

However, the study is confined within its scope given that it is founded on specific dataset consisting of 16 PhD theses that is aimed as being a pilot study. It does not incorporate extensive cross-disciplinary examination throughout a broader scope of scholarly domains, and consequently the results are unlikely to be entirely applicable to every discipline of PhD research.

**Research Design**

The current investigation employs descriptive-analytical and data-driven scholarly research methodological framework to study vocabulary resources in academic writing. The investigation makes use of language-related deviations’ analytical assessment to evaluate recurring patterns of lexis-based and grammar-based application in PhD-level manuscripts.

**Source of Data**

The compiled body of data is composed of a sample of sixteen selected and examined PhD dissertations drafted in English-language. Such theses embody higher-level academic written discourse composed in the PhD stages and deliver appropriate and relevant textual corpus for assessing vocabulary-related and grammar-based configurations in academic and disciplinary scholarly communication.

The current study is based on a corpus of sixteen theses submitted to CHARUSAT between 2016 and 2020, representing diverse academic disciplines including Engineering, Medical Sciences, Computer Science and Application, Science and Management Studies. For the purpose of systematic analysis and to maintain privacy of doctoral manuscripts, the academic documents are labelled as D1 (document 1), D2 (document 2) and so on.

**Methodological Framework**

The analytical assessment carried out in this study employed well-defined and structured methodological design

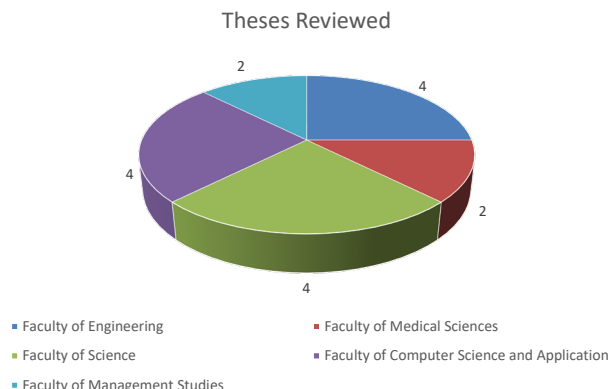


Figure 1: Theses Reviewed

encompassing a number of language-related classifications such as grammar-based correctness, sentence-structural pattern, verb-related deployment, determiners, lexical specificity, and stylistic register. The mentioned analytical indicators were applied to detect and recognize lexico-grammatical configurations and measure the importance of lexis-based repertoire in academic communication. The grammarian model followed in this investigation is Lindley Murray’s grammatical framework mentioned in *English Grammar*.

**Procedure and methodology of data analysis**

Every single doctoral dissertation is scrutinized in an organized manner to observe and locate lexico-grammatical recurring patterns pertaining to vocabulary resources and grammatical structures. The marked errors were sorted and classified in accordance with the grammatical structures presented and published by Linley Murray in his book *English Grammar*. Lindley Murray is considered to be the father of English grammar. The deviations were previously determined on methodological and analytical dimensions. The occurrence rate and distribution pattern were examined to comprehend and explore how lexical repertoire serves in academic communication.

**Findings**

This chapter outlines data-driven and research-based observations drawn from the discourse-related and language-related investigation of sixteen doctoral theses submitted to the CHARUSAT between 2016 and 2020. It concentrates on diverse features encompassing grammatical precision, syntactic arrangement, verbal constructions, use of articles and determiners, conventions of punctuation and capitalization, lexical appropriateness, and language register. The statistical description techniques compute occurrence rate and pattern in linguistic problems in the sampled doctoral manuscripts. The tables below illustrate the framework of analysis detected in recurring patterns.

For the purpose of investigation, the impact of vocabulary range in scientific writing, diverse discourse-related, and language-based classifications and groups were selected and established for evaluation which constitute fundamental and core elements of the aspects of the language that determine comprehensibility and communicative effectiveness in academic writing. The table given below displays language-based dimensions and features adopted in the evaluative standards in the current research inquiry.

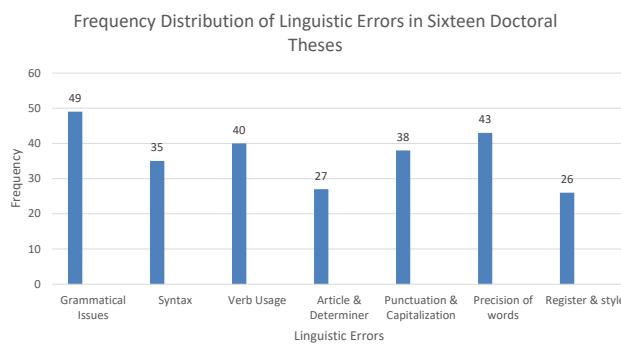
The linguistic constituents specified above present a well-defined conceptual framework for analyzing interplay between vocabulary repertoire, discourse features, and grammatical structures in the scholarly writing at the doctoral level. Such classifications facilitate methodical determination of linguistic tendencies throughout sampled doctoral dissertations.

**Table 1: Linguistic Dimensions of Analysis**

<i>Variable</i>	<i>Description</i>
Syntax	Sentence structure and word order
Verb Usage	Tense consistency and verb form accuracy
Article Usage	Use of definite and indefinite articles
Lexical Specificity	Appropriateness of vocabulary choice
Punctuation	Commas, capitalization, and punctuation
Register	Academic tone and stylistic appropriateness

Upon the completion of designing and framing examination-based framework, the observed and detected language-oriented shortcomings were divided into categories and enumerated across the range of the total of selected sixteen doctoral-level manuscripts. The examination and analytical review of statistical frequency was executed to ascertain the manifestation and incidence of particular error segment and class. The second table summarizes collective and aggregate prevalence patterns of the recognized and detected grammatical and lexical errors.

Figure 2 exhibits a distribution pattern of language-related errors identified and recorded spanning selected sixteen-dissertations at the doctoral level. The analytical reviews demonstrate that the grammar-related deficiencies account for most frequent occurrences closely followed by lexical appropriateness and verbal constructions. Comparatively, linguistic register, grammatical article use are relatively less dominant and frequent. The empirical



**Figure 2: Frequency Distribution of Linguistic Errors in Sixteen Doctoral Theses**

**Table 2: Frequency Distribution of Linguistic Errors**

<i>Linguistic category</i>	<i>Frequency</i>	<i>Percentage</i>
Grammatical Errors	49	19%
Syntax	35	13.6%
Verb Usage	40	15.5%
Article & Determiner	27	10.5%
Punctuation & Capitalization	38	14.7%
Precision of words	43	16.7%
Register & style	26	10.1 %

findings point to grammatical proficiency and vocabulary control continue to represent most prevalent problem areas in PhD-level academic writing.

The third Figure 3 depicts proportional breakdown of language-based errors recognized and evaluated in selected PhD theses. The diagrammatic illustration establishes grammar-related deficiencies and word-choice precision account for majority share in language-based discrepancies.

The recurrence dispersion uncovers language-structural limitations which compose the most significant statistical composition of language-oriented obstacles (19%), occurred subsequently by lexicological exactitude complications (16.7%), and action word deployment (15.5%). Such research results signify that both language accuracy and terminology selection serve meaningful and notable functions in structuring coherence, intelligibility, and comprehensibility in academic communication at the doctoral level. The research outputs suggest that the doctor-stage candidates must exhibit well-developed and substantial cognitive awareness. It is possible that candidates at such stages intermittently and sporadically encounter difficulties in productively implementing grammar-related and lexical language resources.

Supplementary to general recurrence trends and configurations, it is noteworthy to inspect the method by which language-oriented and discourse-based concerns categorized and organized across the range of respective doctoral manuscripts. The third statistical table outlines the statistical distribution of language classifications and segments recorded individually in the sixteen selected doctoral level dissertations examined in the scholarly study.

Figure 4 displays occurrence patterns throughout the corpus of sixteen dissertations reviewed and examined in the study undertaken. The chart-based representation brings out considerable disparity among individual

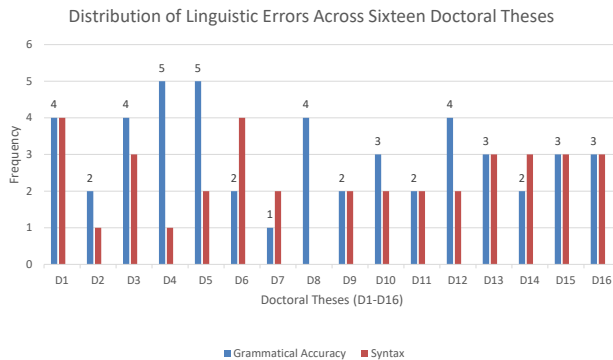


Figure 4: Distribution of Linguistic Errors Across Sixteen Doctoral Theses

dissertations implying inconsistencies in language use continue to exist even in advanced academic contexts and scholarly writing.

The visual 4 demonstrates the occurrence pattern of grammatical and structural errors across selected research theses. The study marked the considerable fluctuation of linguistic structures in individual dissertations. For the sake of clarity, the effective graphical depiction exclusively incorporated key indicators and wider language phenomena identified in the findings. Despite the fact that supplementary categories and further classifications reviewed including article use, punctuation and capitalization, lexical appropriacy, and register were also investigated, such classifications revealed less prevalent instances and it was the reason that they were disregarded in chart-based illustration in order to ensure readability and simplify visual presentation. Notwithstanding, the appearance of such errors in the dataset substantiates key results and central observation of language-related discrepancies that remain and continue to exist with diverse dimensions in doctoral-level academic writing.

The indexing framework supports the scholarly research to illustrate the observation-based and data-driven verbal and lexical arrangements and distributions while upholding concealment of identity of the assessed university-level manuscripts.

The statistical dispersion of communicative problems and weaknesses among the set of sixteen selected doctoral dissertations highlights that the largest number of manuscripts manifest a range of segments and divisions of complications and obstacles. The count of errors is exactly not the same across the doctoral manuscripts. The heterogeneity and inconsistency points toward that despite this sophisticated and upper-level of academic writing stages, the complexity related to precision in vocabulary use, sentence framework and grammatical structures persist widespread. The observations from the research validate and substantiate the prominence as well as relevance of lexical knowledge and mastery as a principal aspect of language framework in academic writing discourse.

Percentage Distribution of Linguistic Errors

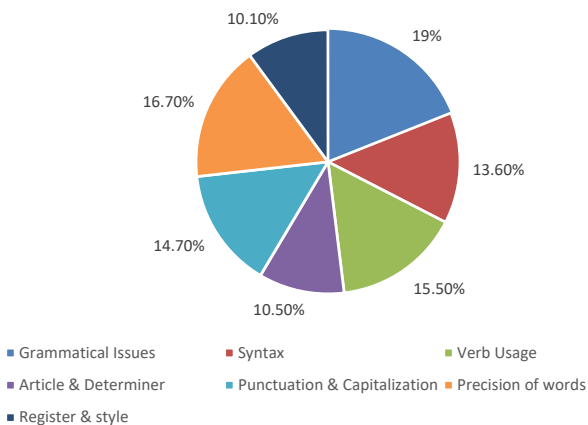


Figure 3: Percentage Distribution of Linguistic Errors in Sixteen Doctoral Theses

**Table 3:** Distribution of Linguistic Errors across Sixteen Doctoral Theses

<i>Document code</i>	<i>Grammatical Accuracy</i>	<i>Syntax</i>	<i>Verb Usage</i>	<i>Article &amp; Determiner</i>	<i>Punctuation and capitalization</i>	<i>Precision of words</i>	<i>Register &amp; style</i>
D1	4	4	3	2	2	3	2
D2	2	1	2	1	1	2	1
D3	4	3	3	2	3	4	1
D4	5	1	0	1	6	6	1
D5	5	2	3	3	3	2	2
D6	2	4	3	0	0	0	2
D7	1	2	1	1	0	0	1
D8	4	0	1	4	5	6	0
D9	2	2	2	0	0	0	0
D10	3	2	3	2	3	3	3
D11	2	2	2	2	2	0	0
D12	4	2	2	2	2	5	3
D13	3	3	4	2	3	3	2
D14	2	3	4	2	3	4	2
D15	3	3	3	3	2	3	3
D16	3	3	4	2	3	2	3

In summary, the research-based, factual and data-driven analytical review brings attention to the pivotal function of word-based and lexico-semantic structures, and grammar-related language resources in strengthening coherence and cohesion across academic documents and doctoral-level manuscripts.

### ***Methodological framework for scrutinizing vocabulary-based range***

The lexical resources function as primary and pivotal positions in determining and constructing semantic meaning, logical coherence and textual unity in scholarly discourse. The advanced-level academic discourse, the lexical repertoire promotes markedly expression of argumentative structures, theoretical transparency and intelligibility, and positioning within the discipline. With the purpose of evaluating how vocabulary range performs in academic communication, the current inquiry examines lexico-grammatical recurring patterns in a sample of sixteen doctoral dissertations, applying structured and methodological framework.

The analytical assessment of doctoral dissertations demonstrates the constraints typically occur due to the non-existence of conceptual and intellectual knowledge which is not the truth. It happens due to the inadequacies and shortcomings of specificity and correctness in lexis and grammar-based construction. In numerous occurrences, the incorrect and ineffective word form selections weaken collocational forms. As a result, inexact and

imprecise vocabulary resources hinder comprehensibility in argumentative structures.

The analytical assessment of sampled doctoral dissertations implies that the isolated and individual lexis is found to be deficient and ineffective in accomplishing sophisticated and advanced academic written discourse. Above all, native speakers and English experts admit that the effective use of collocations and lexical chunks is strongly recommended to produce more quality and precision in the formal writing tasks, especially manuscripts. The accurate use of lexis depends on sensitivity and consciousness of collocational words, expression-based configurations, and domain-specific precision and particularity. The deployment of words tactically and purposefully toward expressive expectations and requirements necessitates to have an alignment with scholarly standards and practices. The requirements of words are not exactly the same in all the contexts. It seems to have a good command of academic vocabulary to produce academic writing tasks including research papers, articles and doctoral dissertations. Several doctoral dissertations displayed and illustrated inadequate employment of lexical bundles and phraseological expressions, collocational configurations, that are extensively acknowledged and regarded as indispensable constituents of effective academic discourse.

The scholarly and academic lexis encompasses a widely used set of academic lexemes, conventional and patterned constructions and utterances along with professional and technical vocabulary. The appropriate use of such lexis

enables scholars, authors and researchers to formulate conceptual and theoretical notions to build critical and evaluative claims in order to engage with field-specific understanding. The phrase sequences constitute phrasal strings to enhance textual connectedness and consistency in the flow of ideas in academic writing that are essential elements and strong pillars while determining the quality in expressions. The application of such lexical chunks and collocation forms support the writers to write abstracts and summaries at an ease. The authorial stance and argument constructions become absolutely clear that explicitly indicates that the vocabulary has great significance in conveying the overall meaning with utmost clarity. The disciplinary and technical lexis should not be replaced with simple words, assuming the readers are uncomfortable and unable to understand. The investigation of dissertations submitted at the PhD-level further points out that discipline-oriented and technical vocabulary undertake significant and vital contributions in building academic authority and trustworthiness, corresponding to discipline-oriented guidelines, practices and conventions.

This is all about the research papers, dissertations and manuscripts where the target audience of each field belongs to the same domain. The usage of technical and professional terminologies cannot be ignored, rather the connection of the text with rhetorical tools and discourse markers can help readers to relate and understand the message more in-depth and with absolute clarity. The awareness of application of words in pairs offers framework-based scaffolding justification with connectors' explanation enabling unification in discourse that is an essential element in academic writing. The domain-related terminology denotes domain-specific word stock connected with distinct and specific subject areas. Using vocabulary systematically embodies theoretical structures and models to align with scholarly conventions and philosophical approaches.

The university-level, scholarly and disciplinary competency incorporate conventions and procedures obligatory for the engagement and contribution in academic writing contexts. It entails the capability and competence to examine and construct discourses in alignment with subject-specific protocols and it engages evaluatively and reflectively in understanding of academic writing. It can be concluded that lexis performs core function in disciplinary and academic proficiency and literacy especially in bridging and supporting availability to subject-specific interpretations and facilitating research-based exchange. Thus, vocabulary functions as a structural and organized means forming disciplinary and academic writing.

### ***Importance of lexis in English***

The lexical stock functions as a pivotal feature of the language-based system in EAP. The vocabulary precision in academic communication fosters comprehensibility,

specificity, logical coherence and textual connectedness that are indispensable. The analytical assessment of sixteen selected PhD theses highlights vocabulary selections markedly impact communicative effectiveness in academic communication.

- The data-driven investigation reveals that lexis-based inventory functions at multiple dimensions in academic discourse:
- At the word level, the lexical repertoire promotes specific articulation of conceptual ideas and viewpoints.
- At the phrasal level, the formulaic sequences, phraseological units and collocational forms enhance linguistic fluency, naturalness as well as textual connectivity.
- At the discourse level, lexical resources reinforce theme-related and topic-oriented progression with logical coherence and textual connectedness.
- At the disciplinary level, the domain-specific and technical vocabulary signify specialized and subject knowledge.

These particular vocabulary-related domains influence each other synergistically and interactively to affect intelligibility, comprehensibility, and semantic continuity in academic writing. The empirical findings evaluated in doctoral dissertations reveal that the vocabulary-related specificity and correctness promote significantly argumentative structure in addition to the progression in authorial positioning as well as domain-specific alignment.

Moreover, it adds to interpretive and understanding-based alignment and stance. The effective usage of lexis is expected to articulate definiteness, explanation and assessment to have an excellent alignment with disciplinary and academic writing conventions. The word-level tools namely qualifying or hedging devices, reporting verbal forms as well as assessment-related qualifying words which hold together the academic writing and help the academic writers to navigate claims and arguments as part of discourse management and scholarly writing. The intellectual writing and research-based writing heavily lean on the effective and appropriate application of lexis and its linkage with other sets of words to convey meaning.

### ***Lexical repertoire***

The academic writing performs by means of well-structured language-based selections that determines how conceptual knowledge is articulated in scholarly communication. Lexical chunks have strong potential to the construction of arguments. It is central to formal communication and essential to balance the meaning and authorial stance in academic writing. A particular determining attribute of scholarly writing is vocabulary-related accuracy. The analytical assessment of sixteen examined dissertations at the PhD level suggests that the lexical-semantic accuracy and linguistic precision are vital determinants contributing

to lucidity in academic argumentative structures. The research-based lines of arguments and thoughts call for the careful selection of nomenclatures and technical terms and the conceptual clarity is built accordingly. The authors are expected to decide on lexis-based components which represent and mirror conceptual perspectives, subject-related traditions, and absolute transparent research-based standpoints. The absence of clarity regularly causes indeterminacy that results in weakening the arguments.

The terminologies additionally facilitate textual connectedness and logical flow across intellectual, research-based, and academic writing. Excellent selection of vocabulary comprehends multifaceted lines of arguments that strengthens associations and correlations across concepts. The textual connectedness or cohesion is a major outcome in academic writing by using a high-range of vocabulary and it performs as a system by guaranteeing research-based compositions to ensure that the text is coherently and systematically arranged. Thus, the linguistic bundles assume a central contribution to the academic text production. The particular and habitual compound-word series act as structural elements in scholarly discourse and as a result, it empowers researchers to systematize line of reasoning while submitting ideas and concepts in different domains of academic writing. The scrutiny of dissertations shows that the occurrence of lexical bundles facilitates more effective discourse related linguistic cohesion and textual connectivity. On the other hand, the nonexistence of lexical chunks gives rise to incoherent and segmented scholarly argumentation.

The academic writing widely depends on noun formation and noun-dominant grammatical structures to condense multifaceted and sophisticated theoretical ideas into concise articulations. The nominalization facilitates systematic approach using compacting details and it also permits conceptualization. Using noun-based formats, the researchers are capable of articulating lines of reasoning in which arguments are compressed and well-structured. The linguistic range is a further attribute in the productive and impactful written discourse. The researchers are required to harmonize the reiteration of ideas alongside consistency together with rhetorical and compositional refinement.

The deployment of vocabulary in interpretively contexts associated with lexis-related elements that inhibits redundancy whilst upholding subject-based persistence. The linguistic adaptability and versatility demonstrate high-level and intellectual proficiency especially in enhancing comprehensibility and ease of reading among the readers. It also influences the author's stance, writing style and the tone through the expressions. The conventional linguistic decisions indicate research-oriented aim and objectives in addition to subject-based recognition and consciousness. The professional and technical terminologies evaluate the

language produced and it also builds reliability in alignment with academic writing practices and standards. A number of dissertations illustrated narrow lexical diversity that hindered stylistic-discursive linguistic sophistication and overall communicative effectiveness.

### ***Knowledge development through lexis-based constructions***

Academic writing fulfills fundamental and essential contribution in the progression and articulation of academic knowledge within the context of academic and intellectual communities. The written composition is an integrated process of idea formation, rational thinking and expertise formation. The linguistic materials allow researchers to systematize ideas generation, cognition and the effective and structured organization of information scholarly viewpoints. The nomenclature operates as an intellectual device that reduces complexity in notions and ideas in verbal forms and types. The research-based domains draw technical lexis-based units in depicting abstract and conceptual frameworks and models, systematic techniques, and evaluative structures.

The empirical observations of the inquiry demonstrate that the deployment of higher-level and sophisticated academic lexis promotes the expression and clarity in multifaceted ideas and theory-based ideas in PhD level discourse writing. The concentration of high-quality words is a mental and intellectual facet of written discourse. The scholarly writing commonly holds considerable percentage and shares substantive and informational lexical items. The researcher needs to regulate vocabulary-related compactness and concentration deliberately and attentively by maintaining sophistication in precision of lexis. The lack of words and excess of words equally harm the meaning in the discourse and formal writing. They have a lot to deal with cognitive alignment. The authors are expected to demonstrate likelihood, assessment and assurance whilst addressing evaluatively in the current domains of information.

The vocabulary resources including hedging linguistic expressions, assessment-related descriptive adjectives and reporting verb forms support academic writers articulate authorial stance and situate argumentative claims within the context of scholarly discussions. By means of vocabulary preferences, the writers mediate to build the scholarly self-concept. The mental processing and vocabulary is essential and vitally important in evaluative composition. The capability of lexis to convey meaning explicitly lies in the appropriate placement in distinct sentence structures. Thus, the lexis has an immense role in fostering analytical and argument cognition.

The writers assimilate field-specific lexis that display consciousness with excellent theoretical models in

academic conventions. The vocabulary conformity with the text indicates affiliation with scholarly writing. On the grounds of intellectual standpoint, the terminology serves a system of arranging and structuring information in association with argument structuring particularly in field-specific material. The role of vocabulary goes beyond language-based articulation and it also embraces psychological mechanisms from the basis of discourse writing. The analysis of dissertations, therefore, confirms that command in vocabulary is strongly correlated to intellectual arrangement and scholarly argument construction in academic composition.

### ***English for academic purposes and pedagogical outcomes***

The empirical findings of the current investigation generate a number of pedagogical implications for language teaching pedagogy in EAP. The lexico-grammatical recurring configurations discovered in the analytical assessment of a sample of sixteen doctoral dissertations emphasize the relevance and necessity of strengthening vocabulary specificity and academic lexical repertoire sensitivity among higher-level academic researchers. The instructional interventions on vocabulary lay a strong foundation for formal compositions. Perceiving lexical repertoire in the capacity of one of the systems in English holds considerable outcomes with regard to academic English educational practice. The established strategies and techniques toward terminology training underscore committing to memory relating to lexical inventories and compilations as well as separate lexis-based components. These similar practices and strategies fall short in order to convey that specific structural and organized as well as utilitarian essence of the lexis put in place. The formal training in word stock is required to be incorporated alongside textual composition progression together with subject-based acquisition.

The evaluation of the selected sixteen doctoral dissertations suggests that a considerable number of lexis-based challenges develop insufficient and constrained vocabulary resources' expertise, but from deficient knowledge of collocational constructions, phraseological units, multiword expressions and specialized lexicon. The text-type based and genre focused strategies provide efficient plans concerning embedding word stock within academic writing training. Through the investigation of the scholarly discourse samples, the language users are able to perceive the manner in which terminology functions internally to distinct and defined discourse forms as well as fields. Therefore, this methodology of context-based study enriches lexis-based insights and it also facilitates textual production progress. The text database applications permit language users in order to investigate vocabulary formations, expression-related structures and collocational forms in scholarly communication.

The tutors undertake a responsibility of helping learners develop vocabulary. The scaffolding is necessary as the teacher stands at the central role in lexis advancement sessions. The classroom sessions are also to be followed by assessments. The evaluation is expected to be in a position to deal with lexical precision. The lexis assimilation necessitates awareness toward subject-based settings. The pedagogy in EAP is required to encourage and support the learners in cultivating technical terminologies pertinent for belonging to them across distinct disciplines. The subject-oriented vocabulary formal training strengthens research-based interaction and it also readies language users for the contribution within academic groups. The techniques to be applied in pedagogy are required to transition away from standalone vocabulary and using words in lexical chunks and collocational forms. Thus, the lexis plays a central role in academic writing.

The observations accordingly recommend that EAP language and pedagogical instruction introduce corpus-supported pedagogical approaches that provide exposure to EAP and English language learners to naturally occurring academic contexts. By means of examination of theses, the academic learners can acquire sensitivity and familiarity of phrase-based constructions, phrase-based structures, lexical recurring patterns as well as academic discourse conventions in academic writing.

### ***Academic program and scheme of assessment***

The empirical observations of the inquiry additionally reveal that the pedagogical implications of academic curriculum design and academic assessment of EAP instructional programs are closely associated with each other. The incorporation of the course content is required to be in alignment with program architecture where EAP is at the central point. The assessment is as important as course design. It is common to consider that pedagogy is an integral part of both the course curriculum and appraisal. In order to understand the requirements of evaluation, the assessment scheme can either be formative or summative and the test can either be controlled practice or freer practice where a number of factors can influence the whole process of assessment.

The lexical development ought to be integrated in producing writing tasks where the vocabulary is primary and can derive the quality in academic writing. The learners participate with such real and quality writing materials where formulaic sequences and collocational forms are illustrated in terms of meaning and forms in the specific contexts. The self-discovery sheets and scaffolded exercises permit students vocabulary rich resources. The self-directed compositions and assignments generate possibilities to embed lexis into more elaborative forms and add more quality in the scientific writing. The assignments have extensive potential in presenting communicative clarity

and enhancing quality in terms of tone, style and discourse management. The language-related trends discovered in the theses establish the importance of academic assessment parameters that determine lexical correctness and specificity, grammatical precision, and discourse-oriented lexical repertoire deployment in academic communication.

The assessment methods need to indicate the value of vocabulary. It is required to be based on criteria and benchmarks that are widely accepted and centered on well-established conventions, standards and practice. The conventional forms of tests emphasize subject-specific concentrations and transparency, level of formality, and textual connectedness.

In the technology-driven era, it has immense competence to facilitate program application. The rising number of technological platforms present enormous options for the investigation and exploration arrangements of lexis in systematic forms. The terminologies encourage autonomy and self-governance among the learners and support them extensively in producing scholarly works. A holistic program in addition to an evaluation framework secures professional support in academic writing.

## Conclusion

- The investigation of sixteen doctoral theses establishes that lexical repertoire serves as core constituents of language systems in the context of academic writing.
- The empirical findings confirm that the preciseness in the lexical items, grammar along with purposeful application of lexical components perform a pivotal role in the process of constructing well-structured lines of reasoning with clear authorial positioning.
- The scholarly study brings to light that weakness related to collocation-based knowledge, lexical bundles and appropriateness in lexis undermine comprehensibility and intelligibility in the text and resultantly, the quality of academic writing is affected adversely.
- The research outcomes reinforce that academic writing proves to be fundamentally built upon vocabulary-related expertise in which lexical repertoire performs as a significant tool for conceptual meaning structuring and research-based articulation.
- The scholarly research brings attention to lexis-based resources that operate extending across diverse domains encompassing vocabulary-level distinctiveness, lexical-pattern-based sequences and area-specific technical terms.
- The conclusions support research in academic English by offering empirically grounded substantiation drawn from academic discourse pointing to key role of lexis-based range in the context of academic communication.
- The inquiry puts forward that more substantial priority upon vocabulary-related enhancement, corpus-

guided instructional approach and specific teaching within academic vocabulary range has the potential to profoundly elevate the excellence in academic discourse.

## Acknowledgement

We extend our deepest gratitude to the Provost, the Registrar, the Dean of Research Cell, and the concerned research supervisors and scholars at Charotar University of Science and Technology, Changa for their extensive and unwavering support, which has enabled us to further excel in our research endeavors.

## References

- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.
- Carter, R. (1998). *Vocabulary: Applied linguistic perspectives*. Routledge.
- Charles, M. (2020). *Corpora and discourse in EAP*. Routledge.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238.
- Coxhead, A. (2018). Vocabulary and English for specific purposes research. In *The Routledge handbook of vocabulary studies*. Routledge.
- Coxhead, A. (2024). *Academic vocabulary in disciplinary communication*. Routledge.
- Coxhead, A., & Webb, S. (2021). *Vocabulary and English for academic purposes*. Routledge.
- Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The academic spoken word list. *Language Learning*.
- Ellis, N. C. (2012). Formulaic language and second language acquisition. *Annual Review of Applied Linguistics*, 32, 17–44.
- Flowerdew, J. (2008). *Discourse in English language education*. Routledge.
- Flowerdew, J., & Peacock, M. (2019). *Research perspectives on English for academic purposes*. Cambridge University Press.
- Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*.
- Gardner, S., & Nesi, H. (2013). *Genres across the disciplines: Student writing in higher education*. Cambridge University Press.
- Gee, J. P. (2014). *An introduction to discourse analysis*. Routledge.
- Gee, J. P. (2015). *Social linguistics and literacies: Ideology in discourses* (4th ed.). Routledge.
- Granger, S. (2015). Learner corpora in language teaching and learning. In *Handbook on corpus linguistics*. Routledge.
- Granger, S. (2021). *Learner corpus research meets EAP*. Cambridge University Press.
- Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Edward Arnold.
- Halliday, M. A. K. (1985). *An introduction to functional grammar*. Edward Arnold.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
- Hyland, K. (2000). *Disciplinary discourses: Social interactions in*

- academic writing*. Longman.
- Hyland, K. (2003). *Second language writing*. Cambridge University Press.
- Hyland, K. (2006). *Academic discourse: English in a global context*. Continuum.
- Hyland, K. (2015). *Academic publishing: Issues and challenges*. Oxford University Press.
- Hyland, K. (2018). *Metadiscourse: Exploring interaction in writing*. Bloomsbury.
- Hyland, K. (2019). *Second language writing* (2nd ed.). Cambridge University Press.
- Hyland, K. (2020). *Academic writing and disciplinary knowledge*. Routledge.
- Hyland, K. (2022). *Teaching and researching writing*. Routledge.
- Hyland, K. (2024). *Writing in the disciplines: Research-based practices*. Routledge.
- Hyland, K., & Shaw, P. (2021). *The Routledge handbook of English for academic purposes*. Routledge.
- Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary"? *TESOL Quarterly*.
- Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, 23(2), 157–172.
- Leong, P., & Hyland, K. (2020). Lexical bundles in academic writing. *Journal of English for Academic Purposes*.
- Lillis, T., & Scott, M. (2007). Defining academic literacies research: Issues of epistemology, ideology and strategy. *Journal of Applied Linguistics*, 4(1), 5–32.
- McCarthy, M., & O'Dell, F. (2016). *Academic vocabulary in use*. Cambridge University Press.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Heinle & Heinle.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
- Nation, I. S. P. (2016). Making and using word lists for language learning. *Language Teaching*.
- Nation, I. S. P. (2022). *Learning vocabulary in another language: Updated perspectives*. Cambridge University Press.
- Nation, I. S. P., & Webb, S. (2019). *The Routledge handbook of vocabulary studies*. Routledge.
- Nesi, H., & Gardner, S. (2012). *Genres across the disciplines*. Cambridge University Press.
- Paquot, M. (2010). *Academic vocabulary in learner writing: From extraction to analysis*. Continuum.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.
- Schmitt, N., & Schmitt, D. (2023). *Vocabulary in language teaching: New directions*. Cambridge University Press.
- Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, 31(4), 487–512.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Street, B. V. (1984). *Literacy in theory and practice*. Cambridge University Press.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students*. University of Michigan Press.
- Tribble, C. (2015). *Writing academic English*. Routledge.
- Webb, S. (2023). *Vocabulary learning and teaching in academic contexts*. Oxford University Press.
- Webb, S., & Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford University Press.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.