

**A Corpus-Based Investigation of Linguistic Patterns in Academic Writing by Non-Native Speakers in Pakistan**

**Zainab Hameed Ullah**

[zainabhameedullah6@gmail.com](mailto:zainabhameedullah6@gmail.com)

Visiting Lecturer, Department of English, Ghazi University, Dera Ghazi Khan

**Abida Kausar Chaudhary**

[achuadary@gudgk.edu.pk](mailto:achuadary@gudgk.edu.pk)

Assistant Professor, Department of History, Ghazi University, Dera Ghazi Khan

**Syed Awon Muhammad Bukhari**

[awonbukhari060@gmail.com](mailto:awonbukhari060@gmail.com)

Lecturer, Department of English, Ghazi University, Dera Ghazi Khan

**Corresponding Author: \* Zainab Hameed Ullah** [zainabhameedullah6@gmail.com](mailto:zainabhameedullah6@gmail.com)

|                      |                     |                      |                       |
|----------------------|---------------------|----------------------|-----------------------|
| Received: 12-09-2025 | Revised: 10-10-2025 | Accepted: 17-10-2025 | Published: 04-11-2025 |
|----------------------|---------------------|----------------------|-----------------------|

**ABSTRACT**

*This study adopts a corpus-based mixed-methods approach to investigate the linguistic patterns in academic writing produced by non-native English-speaking students. Using purposive sampling, authentic academic texts specifically essays and short research papers written by postgraduate students from two Pakistani universities were collected to compile a clean, anonymized corpus of approximately 100,000-150,000 words. The texts were converted into plain-text format, checked for consistency, and organized into sub-corpora based on academic level and genre. Quantitative analysis using corpus tools such as AntConc was employed to examine frequency patterns, lexical diversity, collocational behaviour, grammatical structures, cohesive devices, and recurrent phrase formations. Complementing this, qualitative concordance analysis was conducted to interpret how these linguistic patterns reflect broader academic writing practices among non-native English speakers. The findings reveal notable deviations from native-speaker academic norms, including restricted lexical variation, overuse of general vocabulary, dependence on formulaic expressions, and challenges in producing complex syntactic constructions. These patterns highlight the influence of first-language transfer, academic socialization, and proficiency levels. The study contributes to applied linguistics and English for Academic Purposes (EAP) by offering corpus-informed insights into the linguistic difficulties faced by NNEs writers and suggesting pedagogical implications for enhancing academic writing instruction.*

**Keywords:** *Corpus-based analysis, academic writing, non-native English speakers, lexical and syntactic patterns and English for academic purposes (EAP).*

**BACKGROUND OF THE STUDY**

Academic writing is among the most multifaceted and advanced types of language usage that presupposes not only the command of grammar and vocabulary but also strong knowledge of disciplinary norms, coherence, and rhetoric. Academic writing is a complex task of non-native English speakers (NNEs) who struggle to address the requirements of the global academic community in terms of language and style. Since English is the language of research, publication and higher education, NNEs are usually subjected to linguistic and pragmatic limitations affecting their capacity to write according to the standard of native speakers. It is, then, essential to investigate the linguistic features of their academic writing to learn how the language proficiency, culture, and the educational setting influence written discourse (Haider and

Mehmood, 2022; Haider and Asim, 2022 and Ozturk and Tasci, 2023). With the advent of corpus linguistics, language studies have been transformed to offer empirical and quantitative tools of analyzing large amounts of authentic texts. Corpus based approach allows the researcher to discover common linguistic trends including lexical preferences, syntax, collocations, and discourse markers that would otherwise be unseen through a qualitative research process (Paquet, 2019; Jaafar, 2022 and Nawaz, Siddiqui and Hakim 2025). Corpus linguistics makes it possible to identify trends and deviations in non-native language writing in comparison with native corpora. These observations are priceless to applied linguistics, second language learning, and English to Academic Purposes (EAP) pedagogy. Several corpus-based studies have investigated academic writing over the last few decades, and it has shown that there are substantial differences between non-native and native writers of English in terms of lexical density, cohesion, grammatical accuracy and stance-taking. It has been found that non-native authors tend to use a limited number of lexical variations, excessive use of some cohesive devices, and they experience difficulties in building complex syntactic constructions (Gulzar and Javed; Tarzeen and Tahir 2024). These trends can be attributed to language transfer, lack of exposure to scholarly discourse or cultural differences in the style of rhetoric. Although this is increasing, further corpus based studies which focus on the linguistic behaviour of non-native writers of English are still required with respect to various educational and linguistic backgrounds of the writer- especially in developing countries like in Pakistan. Academic writing is also a key to academic success in higher education because it directly determines the capacity of the students to communicate on a scholarly level, to publish research and to be able to contribute to the global academic discourse. Nevertheless, the existing disparity in the native and non-native writing is usually accompanied by such issues as a decreased rate of publication and a scarce presence at the international academic conferences (Shen et al, 2023; Song et al, 2024). Thus, a systematic corpus-based study of non-native academic writing in English can offer some useful ideas about the linguistic patterns, which determine their discourse.

This kind of findings can be utilized in the curriculum development, language teaching and teacher training programs, in order to develop academic writing competence. Finally, the research aims at exploring grammatical, lexical and discourse levels characteristics of academic writing that are composed by non-native learners of the English language using a corpus-based methodology. Through the discovery of recurring linguistic tendencies and underlying patterns, the study will play a role in the creation of data-driven pedagogical practices and support systems to the specific linguistic needs of NNES in an academic setting.

Therefore, the study aims are;

1. To determine and describe the repetitive linguistic patterns e.g. lexico-grammatical patterns, cohesion markers and syntactic features employed by non-native English writers in academic texts based on a corpus methodology.
2. To investigate the effects of these patterns in linguistics on the overall clarity, coherence and the quality of academic writings of non-native speakers of English by comparing it to the existing standards of academic English.

## **LITERATURE REVIEW**

The article by Gulzar and Javed (2024) considered a collection of writing samples provided by Pakistani undergraduate students to investigate the morphosyntactic characteristics of their academic writing. The authors collected a corpus of undergraduate writing samples and processed the samples through computerised corpus analysis using the AntConc tool. They aimed at determining the language-competency of the learners through determining morphosyntactic constructions, patterns, and frequency of errors in their writing. The theoretical framework that led them to their analysis was the Distributed Morphology according to which morphology and syntax as parts of linguistic production can be

investigated. The results showed that the undergraduate writers had high frequency of linguistically incorrect patterns and structures with a high frequency of errors in the tense, sentence structure and morphology and syntax in general. The authors concluded that the overall competency of the learners is in the middle level with a majority of them having weak mastery of morphology, syntax and grammar with only a tenth of the students demonstrating native competencies in language production.

In the study, Tarzeen and Tahir (2024) used a corpus of academic writing texts by Pakistani authors in which the authors performed an analysis on the instances of discourse markers using the software AntConc (Version 4.3.1). In the article, the authors performed a corpus-based analysis of the texts by examining the occurrences of discourse markers within the texts. The objective of the study was to explore how discourse markers contribute to effective communication, coherence and cohesion in academic writing by non-native speakers. Guided by a discourse-analytic framework that draws on pragmatic functions of discourse markers (e.g., signaling relationships between ideas, organizing information flow), and informed by Belsey's textual analysis (2013) as cited in their work, the authors examined frequency counts, concordance lines and collocational patterns of discourse markers across the corpus. According to their findings, discourse markers are important in increasing discourse coherence and cohesion and writers who used more and more functionally accurate discourse markers had more organised and readable text; some non-native writers under-used or mis-used discourse markers, thus hindering smooth idea-flow and coherence. The study proposes that explicit teaching on the use of discourse markers in EAP (English for Academic Purposes) situations can enhance coherence and readability of non-native academic writers.

In the corpus-based study, Ishaq and Abdulaziz (2023) explore the linguist issues of Pakistani applicants in performing the IELTS test essays. Their data comprise a corpus of IELTS writing task-2 essays written by Pakistani test-takers; the objective is to explore patterns of cohesion and coherence in this learner-writing context. They adopt an analytical framework grounded in Halliday & Hasan's (1976) cohesion model (covering reference, substitution, conjunction, lexical cohesion) and typical coherence indicators such as thematic progression and discourse structure. Their findings show that these learners frequently under-use sophisticated cohesive devices (for instance, fewer instances of lexical repetition and collocation chaining) and show lapses in coherence through abrupt topic shifts, insufficient elaboration of links between sentences, and overreliance on simple connectors ("and then", "also"). The authors conclude that such patterns may reflect both limited lexical-grammatical resources and insufficient training in academic discourse organization, and they suggest targeted instruction focusing on enhancing lexical cohesion and discourse-level coherence in IELTS preparation.

## **METHODOLOGY**

This study adopts a corpus-based mixed-methods approach to investigate the linguistic patterns used in academic writing by non-native English speakers. A purposive sampling technique will be used to collect authentic academic texts such as essays and short research papers from postgraduate students of two Pakistani universities. Approximately 100,000–150,000 words of student writing were compiled into a clean, anonymized corpus, ensuring that only original, unrevised academic texts are included. Then the collected texts were converted into plain-text format, checked for consistency, and organized into sub-corpora based on academic level and genre. In this way, quantitative linguistic analysis has been carried out using corpus analysis software (e.g., AntConc) to examine frequency patterns, lexical diversity, grammatical features, and recurrent phrase structures, while qualitative analysis has had been focused on interpreting how these patterns reflect broader academic writing practices among non-native speakers. The combined quantitative and qualitative insights have tried to identify common linguistic trends, deviations from standard academic norms, and areas where NNES writers may require further academic writing support.

## ANALYSIS

The current study has assessed the targeted students' linguistic proficiency using the corpus analysis criterion. The following linguistic patterns were intended to be examined in the gathered data.

### Corpus analysis

#### Lexico-grammatical structures

Lexico-grammatical structures are the ways we arrange words and grammar to make meaning. In language, grammar and vocabulary (lexis) are linked and mutually dependent upon one another. So, in it, the use of the choice of words and sentence structure reveals the writer's style or the type of content.

#### Academic words frequency

*Table 1.1 Frequency Check: Antconc Software*

| Word Category | Example                                                    | Frequency |
|---------------|------------------------------------------------------------|-----------|
| Consider      | ....materialistic people consider money.....               | 9         |
| Therefore     | .....practicability is therefore one of the .....          | 5         |
| Furthermore   | Furthermore, practical learning allows .....               | 7         |
| Analyse       | .....to identify and analyse the marketing strategies..... | 7         |
| significant   | ..... study would be significant in the sense that it .... | 5         |
| show          | The gift of Magi” does not show .....                      | 9         |
| demonstrate   | These lines demonstrate that Bourgeoisie .....             | 1         |
| say           | To conclude, we can say that .....                         | 19        |
| state         | .....oppressed ones and always in a state of .....         | 19        |
| claim         | ..... the claim that he never engages in .....             | 3         |
| end           | At the end of the lecture .....                            | 15        |
| conclude      | To conclude we can say that .....                          | 5         |
| Need          | .....need to do any more study.....                        | 18        |
| Investigate   | They can investigate the constructive role .....           | 5         |
| Help          | Further, it will help them to.....                         | 20        |

Table 1.1 illustrates the frequency and contextual usage of key academic words identified through AntConc software. The results indicate that words such as *say* and *state* appeared most frequently (frequency 19 each), suggesting that non-native English writers rely heavily on these verbs to express opinions or summarize arguments. This overuse points to a limited lexical variety and a preference for commonly used, less formal academic verbs. The word *end* also appeared frequently (15 times), primarily used in concluding sections, reflecting the writers' tendency to rely on predictable closing expressions rather than diverse concluding markers.

Words that occur fairly often like *consider* and *show* (9 times each) indicate that students are trying to involve themselves in evaluative and analytical talks, which are the key features of academic language. Likewise, *analyse* and *furthermore* (7 instances each) are the signs that the authors are familiar with the analytical and coherent language. Nevertheless, such a relative low use of more advanced academic words as *demonstrate* (1), *claim* (3), and *significant* (5) indicates that the learners might not be confident or familiar with advanced academic terms. There were some instances of the use of transitional markers such as *therefore* and *furthermore*, indicating that a new perception of logical flow and unity was emerging, but their number is still very small.

On the whole, the results indicate that, though non-native English authors reveal certain sensitivity to the academic style in the application of analytical verbs and cohesive devices, their language remains highly monotonous and simplistic. This trend demonstrates a stage of development in their academic writing skills with the focus on the need to provide clear explanation on the use of different academic vocabulary and cohesive features to improve the level of sophistication and clarity in academic writing.

### Nominalization words

*Table 1.2: Nominalization words check: Antconc software*

| Nominalization Ending | Total Frequency | Most common words                                                                                     | Example from corpus                                                                                                                                             |
|-----------------------|-----------------|-------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| *tion                 | 465             | Education, condition, production, situation, Alienation, transformation, communication, investigation | ... multitude of <i>challenges</i> (verb) and <i>opportunities</i> (adjective) that exerted a substantial <i>influence</i> (verb) on its <i>progress</i> (verb) |
| *sion                 | 66              | Conclusion, oppression, discussion, exclusion, comprehension, decision                                | ....discussion (verb) even without the reading (verb) of whole novels, present a certain tragedy (adjective)                                                    |
| *ment                 | 71              | development, statement, treatment, improvement, assessment                                            | ...the <i>development</i> (verb) of communication skills                                                                                                        |
| *ness                 | 47              | Happiness, consciousness, selflessness, awareness, weakness                                           | Political awareness (adjective) regarding its importance (adjective)                                                                                            |
| *ity                  | 152             | Quality, university, practicality, personality, validity, ability,                                    | the men's financial superiority (adjective) upon women and the need                                                                                             |

Table 1.2 reveals that the use of verb and adjective derived noun forms is very common among non-native English writers in the academic writing. The analysis indicates that the most frequent is in words that end with -tion and the frequency stands at 465. The words that belong to this category are education, communication, transformation, and investigation. Dominating -tion nominalizations are an indication that authors are more inclined to describe actions and processes in a more abstract and formal manner, which is a characteristic trait of academic writing. These words transform the active processes into the ideas and the writing sounds more analytical and impersonal.

The second most frequent nominalization ending is -ity having a total frequency of 152. Validity, practicality, and ability are some of the words under this category. These words tend to depict qualities and properties, not actual actions, and indicate the propensity of authors to concentrate on a theoretic or abstract property. This inclination is also indicative of the change in the descriptive language of concrete actions to the generalized ideas and principles, which are in line with the rules of scholarly language.

Words that have -ment at their end are 71 times, which include development, treatment, and assessment. These nouns tend to be the description of processes or outcomes demonstrating how writers speak about the results or outcomes but not about the actions. Likewise, those nominalizations that have -sion endings like discussion and decision are used 66 times. Their moderate presence helps to show that even though reflective or interpretative ideas are included in the writing, they are not so dominant as conceptual or descriptive statements.

The most uncommon nominalizations are the ones that end in -ness and they occur 47 times. The terms such as awareness, happiness, and weakness refer to more personal or emotional terms, which are not prevalent in scholarly writing since such writing is not supposed to be emotional or subjective. The few instances where these forms are used support the impersonal and objective character of the texts under consideration.

In general, the tendency of nominalization application within the corpus shows that authors choose abstract, very formal, and concept-oriented expressions rather than personal or action-oriented. The prevalence of -tion and -ity words is a clear sign of the scholarly propensity to describe the idea and actions as a noun as opposed to a verb. This stylistic element helps to add a more formal tone, conceptual density, and analytical clarity that are also the major features of good academic writing.

### **Hedging and Boosting words**

Hedging is a less forthright and scholarly polite method of weakening a claim. They exhibit doubtfulness or potentiality. Boosters, on the contrary, make you sure and enhance your argument.

*Table 1.3: hedging and boosting words check: Antconc software*

| Category | Example words                                                                                                                                            | Frequency | Function         |
|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|------------------|
| Hedging  | May, Might, Could, Can, Probably, Generally, Likely, Suggest, Indicate, Tend, Assume, About, Relatively, Around                                          | 320       | Show uncertainty |
| Boosting | Clearly, definitely, obviously, always, Never, Prove, demonstrate, Show, Confirm, Completely, Totally, In fact, It is clear that, There is no doubt that | 95        | Show confidence  |

The analysis conducted using AntConc software, as shown in Table 1.3, reveals a notable difference in the use of hedging and boosting devices in the academic texts of non-native English writers. The use of words like may, might, could, suggest and indicate were hedged 320 times indicating that authors were often cautious with words and language to be uncertain and remain objective. This is indicative of their consciousness of the rules of academia that prefer conditional and moderate assertions.

Conversely, the increase in the use of words such as clearly, definitely, demonstrate, and prove was 95 times, which means that assertive words were rarely used. The fact that the frequency of boosters is low implies that the students were fond of using a polite and careful tone, which might be explained by a lack of confidence in making strong statements. On the whole, the results demonstrate that non-native authors use the hedging rather than boosting approach and demonstrate the knowledge of the academic tone but also demonstrate the necessity to teach them how to balance between being cautious and confident to make their writing more persuasive.

### **Passive constructions**

It is also concerned with the outcomes or the process and not the researcher. It may be helpful in highlighting findings or where the researcher is anonymous, but it tends to create confusion, diminish the impact and interest in writing.

*Table 1.4: passive construction words check: Antconc software*

| Form      | Example                                                                                                                                                      | Frequency |
|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| Was *ed   | Was comprised, was combined, was forced, was considered, was determined, was developed, was educated, was established, was focused, was followed, was framed | 37        |
| Were * ed | Were announced, were applied, were believed, were considered, were influenced, were noted, were proposed                                                     | 19        |
| Is *ed    | Is accepted, is focused, is aimed, is applied, is associated, is based, is called, is claimed, is comprised, is considered                                   | 91        |
| Are *ed   | Are aimed, are allowed, are tested, are asked, are connected, are considered, are determined, are emphasized, are focused, are mentioned                     | 53        |

The AntConc software analysis demonstrates that passive forms of constructions are common in academic writing of students. The most frequent one, is + past participle (is ed) is used 91 times, which means that the present passive constructions are more likely to be used to describe general facts, accepted knowledge, or processes that take place (e.g., is applied, is based, is considered). This implies that authors tend to refer to actions or states without necessarily referring to the agent, which is objective and impersonal enough to be written in academia.

Forms are + past participle and was + past participle were used 53 and 37 times respectively indicating that they are used both in the present and past to describe what has already been done or facts that are universally considered true. The lower frequencies of such forms as were + past participle (19 times) and been + past participle (29 times) imply the lack of usage of perfect and plural passive constructions.

On the whole, the results suggest that non-native English authors use passive voice extensively to deliver information in a formal and objective way, which follows the academic rules of stating the results and processes but not the agent.

### **Cohesion markers**

Cohesive devices or cohesion markers are words or phrases that are employed in order to smoothly relate ideas, sentences and paragraphs. Through showing the links between the ideas, they help the reader to follow the logic of a text. Markers of cohesion play a very important role in academic writing since they render ideas clear, logical, and coherent. They are classified into additive, adversative, causal, temporal and referential markers.

*Table 2: cohesion markers check: Antconc software*

| Category    | Words                                                       | Frequency | Function                                          |
|-------------|-------------------------------------------------------------|-----------|---------------------------------------------------|
| Additive    | And, also, furthermore, moreover, besides                   | 1045      | Used to add or expand ideas                       |
| Adversative | But, however, yet, although nevertheless, on the other hand | 173       | Used to show opposition or contrast               |
| Casual      | Therefore, thus, hence, because, since                      | 69        | Used to demonstrate cause and effect relationship |
| Temporal    | First, next, then, finally, afterwards                      | 82        | It shows time or sequence                         |
| Referential | This, that, these, those, such, it, they                    | 1027      | Used to review the previously discussed concepts  |

The findings reached with the help of the AntConc software are used to emphasize the frequency of usage and the cohesion markers in the academic writing of the non-native speakers of English. The most commonly used are additive and referential markers. Additive markers, including and, also, furthermore, moreover and besides were used 1,045 times indicating that students are using these devices extensively to elaborate on ideas or to add extra information. This frequent usage implies that non-native writers tend to construct the arguments by connecting similar ideas, which can be seen as their effort to keep the language and coherence. But the excessive reliance on simple additive signs, such as and could be taken as a sign of the absence of diversity and complexity in the use of cohesive.

Equally, the referential markers like this, that, these, those, such, and it were used 1,027 times, which means that students often revert to the already discussed concepts to create continuity. This is an indication of their knowledge about the role of cohesion in academic writing. But the high frequency also indicates that they might use more pronouns and determiners than logical connectors and thus result in repeated or ambiguous referencing.

Adversative markers (but, however, although, on the other hand), in turn, were used 173 times, which is not much evidence of the use of contrasting devices. It implies that students may struggle to come up with balanced arguments or highlight differences that is a valuable attribute in critical academic writing.

Causal markers such as therefore, thus, hence, because and since were repeated 69 times, which presupposes the relative absence of attention to the cause-and-effect connections. This could mean that authors are not as certain to showcase the rational thinking or argumentative relations. On the same note, temporal clues (first, next, then, finally, afterwards) were used 82 times, demonstrating a certain attempt to arrange ideas in sequence or describe processes without being as eminent as additive or referential devices.

Overall, the analysis shows that non-native writers demonstrate a basic awareness of cohesion but rely more on simple and repetitive devices like *and* or *it*. The limited use of advanced connectors such as *however*, *therefore*, and *moreover* indicates a need for further instruction on using a wider range of cohesive tools to enhance the logical structure, argumentation, and academic tone of their writing.

### **Syntactic features**

Syntactic features refer to the structural patterns and grammatical arrangements of words, phrases, and clauses that form sentences in a language. In academic writing, syntactic features are essential because they determine how ideas are organized, connected, and expressed formally and logically. Since language cannot exist without these structures, it is evident that these are essential components of language. The mechanisms of syntactic formation observed in the writing samples of non-native speakers have been examined to gain insight into the language used by the students. Here are some examples from students' collected writing samples:

- i. “.....they are unable to find equal opportunities to write and publish their work as well, and it is just because of the superficial mindset.....”

It is a compound complex sentence with independent (The disappearance of women writers does mean that they are intellectually inferior to men but the fact is..), dependent (that they are intellectually inferior to men), relative (which is not willing to accept a literary work...) and adverbial clauses (because in this way they would be failed to show them...). Moreover, ‘but’, ‘that’, ‘which’, and ‘because’ indicators are used in it. Despite being a little overcrowded, the phrase has significant grammatical complexity, which is typical of professional writing. The author skillfully connects concepts of gender inequality and social critique through coordination and subordination.

- ii. “In 1843, he married a lady named Jenny who was very attractive and intellectual and came from a military background.”

It is a complex sentence with independent (He married a lady named Jenny) and relative (who was very attractive and intellectual and came from a military background) clauses. This is a well-constructed phrase that describes the noun “lady” using a relative clause. This is an excellent illustration of how to improve content quality and structure through descriptive expansion in the academic language through subordination.

- iii. “Woolf also draws our attention towards the fear of the Victorian writers who were not intended to write poetry because they did not want to be called sentimental, as men considered them.”

This is a complex sentence with independent (Woolf also draws.....Victorian writers), relative (who were not intended to write poetry), adverbial (because they did not want to be called sentimental), and comparative clauses (as men considered them). The use of “because” adds causality, while “who” and “as” show descriptive and comparative relations reflecting strong syntactic control. Minor grammatical slips (“were not intended”) reduce fluency but not structural depth.

- iv. “You might spend hours learning the climatic conditions of Africa but a visit to Africa can stick the picture in your mind instantly and very effectively.”

This is a compound sentence with two independent clauses and a conjunction ‘but’. It expresses contrast between time-consuming theoretical learning and immediate experiential learning. The use of the modal “might” softens the first claim (hedging), typical of academic tone. The second clause adds an evaluative adverbial phrase (*instantly and very effectively*), strengthening the contrast.

## **FINDINGS & DISCUSSION**

The study has been conducted to examine the linguistic patterns in the writings of non-native English speakers. Hence, the first objective of this research investigation was to identify the lexico-grammatical, cohesion markers, and syntactic features in the academic writings of non-native English speakers. From the analysis, it became evident that although non-native writers use characteristic features of writing such as nominalizations, passive constructions, and modal verbs in their writings, they also frequently misuse familiar patterns and steer clear of complex noun phrases or diverse verb forms that native academic writers frequently use for clarity and impact overall. Through their writings, non-native writers showed knowledge of cohesiveness but lacked balance between additive and logical connectors as compared to accepted academic English norms, which impacts the overall coherence of their writings. Moreover, the readability and coherence of their writing are affected due to their use of too long sentences with poor punctuation and multiple conjunctions. The second purpose of this research was to examine the impact of linguistic patterns on the effective clarity, coherence and academic quality of non-native speakers where the researcher discovered that non-native writers are increasingly having a good but potential knowledge of academic English rules. Although the common application of boosting hedging proves the efforts to show confidence and make attempts to introduce scholarly prudence. But the lack of equality between the two dilutes the arguments. Syntactic patterns exhibit a feeling of academic complexity, yet the slightest subordination is often substituted with the length of sentences and coordination. Such findings testify to the fact that non-native writers are at a middle stage of development in the light of the academic norms; they know the traditions of academic writing and are still developing the control, diversity and accuracy of their use.

## CONCLUSION

In conclusion, it is possible to state that despite the knowledge of academic style writers, the application of coherent techniques and complex syntax is often not precise and even lacks diversity. Nominalizations and passive constructions are frequently used, but sometimes they are abused, which negatively affects the readability and comprehensibility. Equally, a deficiency in control of textual coherence is also manifested through overreliance on additive and referential cohesion cues. Long and complex sentences with clauses are normal syntactically, but structural imbalances ruin coherence. These texts indicate that the formal written English is partially construed with respect to the normative academic standards.

## REFERENCES

- Ahmad, M., Asim Mahmood, M., & Siddique, A. R. (2022). Exploring disciplinary variation in Pakistani academic writing: A corpus-based research on doctoral dissertations. *Pakistan Languages and Humanities Review*, 6(4), 51–60. [https://doi.org/10.47205/plhr.2022\(6-IV\)06](https://doi.org/10.47205/plhr.2022(6-IV)06).
- Gulzar, S., & Javed, A. (2024). A Corpus Based Study of Morphosyntactic Features in Undergraduate Writing Samples. *Journal of Applied Linguistics and TESOL (JALT)*, 7(4), 585–596.
- Haider, S., & Asim Mahmood, M. (2022). A corpus-based analysis of coherent writing skills of Pakistani English language learners. *Pakistan Languages and Humanities Review*, 6(2), 892–907. [https://doi.org/10.47205/plhr.2022\(6-II\)76](https://doi.org/10.47205/plhr.2022(6-II)76)
- Jaafar, E. A. (2022). Collocation networks of selected words in academic writing: A corpus-based study. *Arab World English Journal*, 13(1), 240-255. <https://doi.org/10.24093/awej/vol13no1.15>
- Malik, M. Z. A., Kharal, A. A., & Shehzadi, K. (2023). A corpus-based study of the lexical bundles in the Ph.D. linguistics dissertations' abstracts in Pakistani universities. *Pakistan Journal of Humanities and Social Sciences*, 11(2), 2317–2326. <https://doi.org/10.52131/pjhss.2023.1102.0524>.
- Nawaz, M. S., Siddiq, S., & Hakim, S. (2025). A corpus-based linguistic analysis of discourse marker usage in English graduate students' academic presentations. *Journal of Applied Linguistics & Language Teaching* <https://doi.org/10.63878/jalt1280>
- Öztürk, Y., & Taşçı, S. (2023). A corpus-based analysis of lexical bundles in non-native post graduate academic writing and a potential L1 influence. *Reflections*, 30(2), 488–505. <https://doi.org/10.61508/refl.v30i2.267463>.
- Paquot, M. (2019). *Phraseological patterns in learner academic English: Insights from corpus-driven approaches*. John Benjamins Publishing.
- Shen, C., Guo, J., Shi, P., Qu, S., & Tian, J. (2023). A corpus-based comparison of syntactic complexity in academic writing of L1 and L2 English students across years and disciplines. *PLOS ONE*, 18(10), e0292688. <https://doi.org/10.1371/journal.pone.0292688>
- Song, Z., Hamat, A. B., & Che Abdul Rahman, A. N. B. (2024). Metadiscourse in English academic writing across disciplines: A systematic review. *Pakistan Journal of Life and Social Sciences*, 22(2), 20211-20227. <https://doi.org/10.57239/PJLSS-2024-22.2.001482>
- Tarzeen, R., & Tahir, A. (2024). Analyzing the Impact of Discourse Markers in Academic Writing: A Corpus-Based Approach. *Pakistan Languages and Humanities Review*, 8(3), 138–166. [https://doi.org/10.47205/plhr.2024\(8-III\)13](https://doi.org/10.47205/plhr.2024(8-III)13)