



Cross-Linguistic Influence on Prepositional Errors: A Corpus-Based Computational Study

Owais Ahmed^{*a}, Dr. Marriam Bashir^b

a. PhD Scholar, School of English, Minhaj University, Lahore

b. Associate Professor, School of English, Minhaj University, Lahore

ARTICLE INFO

Received:

May 02, 2026

Revision Received:

May 16, 2026

Accepted:

June 01, 2026

Available Online:

June 20, 2026

Keywords:

learner corpus,
prepositional errors,
interlanguage, cross-
linguistic influence,
computational linguistics

ABSTRACT

This paper explores patterns of prepositional errors in L2 English writing made by Urdu-speaking undergraduate learners using a corpus-based computational approach. Prepositions are one of the most complicated grammatical categories to learn because they have semantic variability, polysemy and contextually dependence. Sixty translated essays were gathered from two universities to construct a learner corpus. Data were digitized, standardized and processed by means of Natural Language Processing tools, such as part-of-speech tagging and concordance analysis together with manual error annotation. The prepositional errors were categorized into substitution, omission and addition. The results indicate that substitution errors are the most frequent followed by omission and addition errors which indicates the partial and unstable knowledge about the usage of prepositions. The findings also reveal that there is a significant effect of L1 Urdu postpositions on prepositions. Secondly, it points out that the learner interlanguage is systematic in nature and establishes the combination of corpus linguistics and computational methods for analysing learner language more effectively. The results have valuable implications for corpus-guided pedagogy and the designing of computational tools second language learning.

1. Introduction

Prepositional errors in writing in the second language (L2) English have become an important field of concern in the domain of second language acquisition (SLA). Since, second language learners commit grammatical errors even in the most basic structures despite having been taught grammar for many years. When asked to write in English at school, they might spend minutes writing one grammatically correct sentence. The Second language acquisition (SLA) scholars have been keen on the preposition errors committed by the second language learners as per Ellis (1994). Prepositions, which serve to indicate the relationship between nouns, verbs and other words in a sentence, are usually a big challenge to the second language learners. Second language learners have difficulties with learning English prepositions because these prepositions do not correspond to the one-to-one correspondence of the same constructs in other languages, and exhibit the high degree of polysemy, idiomaticity and syntactic flexibility (Tyler and Evans, 2003). The multifaceted nature of prepositions themselves, in which a single preposition can have many functions and meanings, is only part of the complexity due to the impact of the first language (L1) of the learners and the multifaceted ways different languages can encode relationships. It is therefore important to gain knowledge on the origin and trends of prepositional errors in writing L2 English to further the theoretical constructs of interlanguage development and to inform the practice of teaching languages.

* Correspondence to: PhD Scholar, School of English, Minhaj University, Lahore

E-mail address: owaisahmed5725@gmail.com (O. Ahmed).

English prepositional errors are one of the most evident issues in English writing since they often tend to interrupt the communication process and indicate the problems with grammatical comprehension. These mistakes are informative about the mental activity of the learner and can help understand the way in which they bargain syntactic and semantic roles of prepositions. Studies have established that prepositions are also one of the more challenging terms of English to non-native speakers as they are used idiomatically and play complex roles within sentences (Bitchener and Knoch, 2010). The ambiguity of prepositional use may result in the use of transfer errors of the first language of the learner, in which systems of prepositions can be considerably different. As an example, a Spanish speaker may use the wrong word in that the word *in* is used instead of *on* because of direct translation of the native syntax of Spanish speaker, which uses a different prepositional form.

Conventionally, learner error analysis has been based on anecdotal observations and qualitative analysis. Nevertheless, these methods frequently do not have the empirical rigour required to explain systematic patterns or predisposing causes of prepositional errors. This research gap necessitates corpus-based approaches, which make use of large bodies of genuine texts written by second language learners to determine and examine errors in a more systematic way. Corpus linguistics will allow the researcher to know how and how often prepositional errors occur so that the researcher can conduct a more in-depth analysis of the learning interlanguage. The intricacies of grammar development in second language learners can be unveiled through investigating a high number of written samples, which can enable researchers to detect recurrent mistakes concerning specific linguistic and cognitive factors.

The advent of learner corpora, that is, carefully screened collections of L2 writing, is a fertile area to explore. The prepositional use can be analyzed in the context of these corpora, which gives a micro-level view of the specific word choice and a macro-level of the tendencies of a wide range of learners, levels of proficiency, and the purpose of writing. The patterns observed in the corpus analysis not only make scholars aware of what interlanguage is but also offer active insights that may be immediately applied to the language pedagogy.

The corpus-based approaches have offered a solid analytical framework to examine these mistakes substantively. Corpus-driven studies provide a robust methodological framework to analyse natural language patterns across different learner levels, offering empirical insights into developmental trends and persistent difficulties. With the help of massive corpora of learner language, researchers can find patterns and frequencies of prepositional errors in various learner demographics and proficiency levels. The methodologies trigger an improved comprehension of typical interlanguage characteristics that are unique to different linguistic backgrounds. Indicatively, research that applies the International Corpus of Learner English (ICLE) has shown that there is systematicity in poor usage of prepositions that points to individual errors as well as larger-scale patterns attributed to the linguistic backgrounds of different learner populations (Granger, 2014).

Exploration of prepositional errors as applied to learner interlanguage not only enhances the knowledge on the acquisition process but also serves as a reference in guiding teaching strategies that can enhance L2 writing skills. These are the lessons that can be learnt during the analyses based on corpus that would allow teachers to meet the individual needs of their learners, which in turn would increase the overall efficiency of the instruction process, and help learners acquire the language successfully. The use of corpus-based approaches to research based on empirical data gathering and analysis has become eminent in other areas of study in linguistics, especially in the field of second language acquisition (SLA). These approaches make use of extensive, electronically stored sets of genuine language material, referred to as corpora, to examine language usage, discover patterns, and clarify linguistic phenomena. As language scholars are increasingly interested in informing the instructional decision-making process by using an evidence-based methodology, corpus-based analysis pays off by providing both measurable and descriptive insights about the practice of language.

1.1. Research Questions

The paper seeks to determine and categorize the nature of prepositional errors that are made by Urdu speaking undergraduate students in writing in L2 English. It also aims to study the occurrence and patterns of these errors based on a corpus-based computation method. The research will also test the level of cross-linguistic interference of L1 Urdu by English prepositions in the context of interlanguage theory among the learners.

1. What types of prepositional errors are made by Urdu-speaking undergraduate learners in second language learning?
2. How do frequency patterns and distribution of prepositional errors in the learner corpus resemble?
3. To what extent does L1 influence L2 learning?

2. Literature Review

The prepositional errors that are usually evident in L2 writing can be classified into several concepts that include overgeneralization where learners use a preposition that is known to them but in an unsuitable context and underuse where learners miss out on the prepositions that they should use. To illustrate, a Spanish learner may make mistakes in writing such that he/she writes something incorrect like I am interested on, rather than I am interested in. These errors tend to be the attempts of the learners to apply the familiar linguistic rules in English, which is thoroughly recorded in interlanguage theory. The interaction of L1 and L2 results in a systematic error pattern which can be studied to determine some underlying cognitive processes. Numerous studies have documented the use of in-class corpus-driven DDL to improve academic writing. To achieve effective direct use of corpus in L2 writing, several studies have utilized corpus as a reference tool to correct learners' errors and enhance their writing skills.

The corpus-driven approach can be used to better understand these patterns of errors as it becomes possible to see the pitfalls of prepositional errors that are most frequently encountered when used by different groups of learners. It is possible to follow specific patterns of prepositional misuse among learners of different linguistic backgrounds, indicating that specific teaching methods could be designed to help them deal with problems that are relevant to the groups. Concentrating on frequent mistakes, teachers will be able to design the pedagogical intervention, based on the kind of prepositions that are misused by students.

Moreover, the consequences of such findings also have an impact on language teaching practices. The awareness of the most common and typical prepositional mistakes in L2 writing can serve to design the curriculum and develop instructional materials to focus on practical activities that will be grounded on prepositional accuracy. Skilled instruction methods such as learning to explicitly train prepositional phrases, situational training and using of corpus-based text can do wonders to assist learners to master this challenging aspect of the English language. With the ongoing development of corpus analysis, the lessons that are gained in investigating the errors in prepositions have the potential to fill the gap that lies between the interlanguage competencies of learners and the norms of writing academic and professional English texts. Thus, it is urgent to continue researching on these patterns of errors by teachers and scholars with the assistance of data-driven results in the way to improve the outcomes of language learning, in particular, interlanguage learner and pedagogical problems of second language (L2) English writing. One of the most outstanding contributions to the current area of study is the work by Neumanova (2023), which provides a detailed discussion of the use of preposition in L2 English learners. Using a corpus-based method, Neumanová examines the nature of error in the prepositional selection and points to the implications of the findings in the context of language teaching.

Neumanova (2023) also points out that prepositions are especially problematic when it comes to second language learners since they are idiomatic and the intricacy of their application within diverse syntactic frameworks. This research examines a huge amount of written English produced by learners representing various language groups, which makes it possible to study prepositional errors in detail and situations when they happen. The major results show that learners tend to apply the rules in their native language that may cause systematic errors in their choice of prepositions. As an example, one of the mistakes that were repeatedly noticed was the misuse of prepositions in the phrasal verbs, which are among the most challenging to learn because they are dependent on the contextual meaning, and not the grammatical rules.

Aybek and Can (2021) emphasize the corpus-based approach as a useful tool to study the error cases in the learner language, specifically in terms of prepositional use in the second language (L2) writing. Their study describes the method of creating and studying a corpus of L2 English texts written by learners and through this, a systematic study of prepositional errors can be achieved. With computational methods to examine such a corpus, the frequency and context of individual prepositional errors can be determined, both to identify the nature of errors made and the factors that contribute to such errors.

Alrashidi (2022) stresses the need to incorporate guided activities resulting in learners systematically exploring prepositional errors. Lack of adequate scaffolding can cause learners to drown in the multitude of data present in larger corpuses, and this can confuse them instead of enlightening them. The teaching methods must therefore focus on being clear on the error patterns and promote exploratory learning. Neumanova (2023) also points out that prepositions are especially problematic when it comes to second language learners since they are idiomatic and the intricacy of their application within diverse syntactic frameworks. The research examines a big collection of written English by students whose native languages are varied, which makes it possible to conduct a comprehensive study of prepositional errors and the situations when they happen.

Aybek (2025) points out that prepositional mistakes are common in L2 academic texts because use of prepositions in English is complex and it may vary a lot with the native languages of learners. Incorporating corpus-based approaches, Aybek studies a huge corpus of scholarly texts non-native speakers write, which enables a quantitative study of prepositional errors, as well as a qualitative study of types of errors. Aybek (2025) reports one of the outstanding findings of overgeneralization of prepositional forms. An example is where learners often demonstrate a propensity of using prepositions that they have acquired in contexts

to the wrong syntactic structures. This may be seen in such expressions as depend on rather than the right way of depending on which implies that such mistakes are not only due to the ignorance of the learner on the usage of prepositional constructions in the English language but also the fact that the first language structures of the learners interfere with the English language structures.

The exploration of L1 transfer in prepositional writing in L2 English writing throws light on how the forces of learner interlanguage interact and highlights how a specific instructional approach that considers and considers these forces is necessary. The article by Gazioğlu and Aydin (2024) offers a critical basis to continue examining this correlation and its consequences to improve learning English language teaching. The current study of prepositional mistakes in the texts of Czech university students has provided important clues into the essence of interlanguage development of English as a second language learners. By applying methods based on corpus, we have a solid platform on which to study the frequency, nature and contexts of errors, which could be used to inform linguistic theory and practice. The corpus of written assignments completed by the students attending a course in English language in a Czech university was used as the data set on the analysis. The corpus has a wide variety of different genres, such as essences, research papers, and reflective writing, which makes it possible to gain an in-depth insight into the prepositional use in various writing situations.

Quantitative analysis showed that prepositional errors were found to be one of the most common forms of grammatical error found in the writings of the students. It is important to note that the error rate of prepositions was about 18%, making it very high compared to other types of errors i.e. verb morphologies and subject-verb agreement. This observation is in line with the current literature, which emphasizes prepositions to be of difficulty to the non-native speakers because prepositions are idiomatic and contextual in use in the English language (Bardovi-Harlig and Dornyei, 1998).

3. Method

3.1. Research Design

The current research article employs a corpus-based research design that combines the quantitative and qualitative analysis (Brezina, 2023; McEnery and Hardie, 2022) and the qualitative methods of interlanguage interpretation. This combination analytical framework allows the researcher to find common patterns of prepositional use besides explaining the linguistic procedures behind these patterns. It is based on the learner corpus approach which deals with actual production of learners, instead of produced artificially (Granger et al., 2022; Paquet and Granger, 2021). The translation-based elicitation design was utilized to make sure that the learners were exposed to a variety of linguistic contexts in which they were to use English prepositions. This method facilitated the ability to have control of data collection and yet to have the natural variations of learner language. The combination of computational devices (part-of-speech tagging and concordance analysis) contributes to the methodological rigor of the study as it allows us to conduct the analysis of the linguistic data in a systematic and replicable way. Simultaneously, manual annotation will guarantee that the contextual subtleties will be properly understood, especially when the automated tools can fail to classify the grammatical forms correctly.

3.2. Participants

The sample population of this study is Urdu-speaking undergraduate learners from two universities where English is studied as a mandatory course during the first semesters. Sixty students were used in the study with at least thirty students being chosen in each university to achieve a balanced representation (Hyland, 2022). The participants were all Urdu learners of English and had similar educational backgrounds, which aided in preserving uniformity in the linguistic exposure and the level of proficiency. The sampling was done using purposive sampling (Etikan and Bala, 2021) and the students who had prior exposure to writing academic papers in English were selected. This made sure that the data collected was of true learner efforts to apply the English prepositions in meaningful contexts. The fact that two institutions are utilized enhances the credibility of the results as it considers the difference between the educational settings. Consideration of ethics was taken seriously, and the participants were educated on the aim of the study. They remained anonymous and the information collected was only used in research.

3.3. Instrument

In this study, a set of structured Urdu essay prompts that sought to induce specific grammatical constructions (Dornyei, 2022) the application of the English prepositions in various linguistic situations was the main data collection tool. The prompts were formulated with great care to incorporate cases of spatial, temporal and abstract relations so that the participants would tend to use various types of prepositions during translation. The instrument comprised of brief, context-based essays in simple Urdu to suit the linguistic ability of the undergraduate students but still has enough complexity to offer meaningful analysis. To ensure that the prompts are content-valid and undergo pilot testing (Mackey and Gass, 2022) to validate the alignment with the aims of the study, they were reviewed before the data collection procedure to make sure that they stimulate the usage of

prepositions. They also carried out a pilot testing session involving a few students to determine the ambiguities or problems with understanding. According to the feedback, some minor changes were implemented to make it clear and appropriate. The completed instrument also provided consistency of data elicitation among the participants, and it helped in the reliability of the learner corpus.

3.4. Data Collection

The current study collected data in a controlled classroom-based elicitation (Cohen et al., 2022) translation tasks (Cook, 2021) which were given to undergraduate students in two universities. A set of Urdu essay prompts was given to the participants, and they were requested to translate them into English within a predefined time. The experiment was done in the normal hours of the school day with the acquiescent help of course instructors to make sure that the environment was natural in academic context, and that consistency of administration was maintained. The participants were clearly instructed to come up with their own translations and not to use external help so that the learner's language would be authentic. At least thirty responses were received in terms of each university, which provided a total of sixty essays to construct a corpus. All the scripts completed were gathered right after the task to avoid revision or outside influence. Ethical issues were duly considered, and all the study participants were notified of the study aim and given the assurance of anonymity and confidentiality. The data gathered were then digitalized and ready to be analysed using corpus.

3.5. Data Analysis

The analytical process of data analysis was systematic and multi-stage; it used both the computational and the manual linguistic assessment. The digitization of all handwritten essays and their conversion into plain text in the first stage was done. They were then purged to eliminate irregularities with formatting, like uneven spacing, and non-linguistic characters, without correcting any grammatical errors to preserve the originality of learner language.

The second stage involved part-of-speech tagging of the corpus, which was done using a Natural Language Processing tool (Jurafsky and Martin, 2023) to automatically recognize prepositional tokens in the corpus. The labeled data made it possible to extract all the cases of prepositional use and analyze them with the help of corpus concordance and frequency analysis (Anthony, 2022). This move enabled the researcher to examine the way prepositions were applied in varying situations and the possibility of variation and deviation of the standard application of English.

All identified instances were reviewed and annotated manually in the third stage. In the study of learner language (Ellis and Barkhuizen, 2022), the classification of errors is made into three main categories: substitution, omission, and addition. All the cases were considered within the entire syntactic and semantic context to guarantee proper classification. Python-based scripts were then used to create frequency counts to measure the distribution of error types. Lastly, the findings were qualitatively interpreted through the interlanguage theory to describe how errors committed by learners are a demonstration of underlying language processes and cross-linguistic influence. The manual verification was done to minimize tagging errors.

4. Results and Discussion

The current paper has explored the prepositional errors in writing L2 English by undergraduate learners with Urdu speaking background using a corpus-based computational method. The results are discussed according to the three research questions that include the classification of types of errors, the distribution of their frequency and the presence of cross-linguistic influence in the context of the interlanguage theory. Quantitative patterns are integrated to give a comprehensive understanding of learner behaviour by integrating with qualitative interpretation.

4.1. Types of Prepositional Errors

The research question number 1 is to identify and determine the types of prepositional errors that were committed by Urdu speaking undergraduate students. The learner corpus was analysed, and three main types of errors were identified: substitution, omission, and addition. These are classes that indicate different aspects of learner complexity and offer information on interlanguage cross-linguistic influence that is dynamic in SLA (Jarvis and Pavlenko, 2021).

Table 1: Distribution of Prepositional Error Types

Error Type	Frequency	Percentage
Substitution	64	46%
Omission	51	37%
Addition	23	17%
Total	138	100%

The most common type of error was substitution errors which comprised almost half of the identified errors. These mistakes were mostly connected with the wrong choice of prepositions in cases when it was necessary to use another variant. As an example, students often used *in* place of *at* in the following expressions: *in the morning at 7 o'clock* or *at the school building*. These mistakes demonstrate that learners have a fundamental knowledge of prepositional use but do not have specificity in using the contextually right forms.

The second most common type was omission errors, which were found in the contexts where English obliges obligatory prepositions. Such constructions as *he went to school* and *he went to school* are examples. Such mistakes indicate that students tend to neglect the grammatical requirement of prepositions in the structure of English sentences.

Although less common, additional errors offer valuable information about the strategies of learners. These were faults in the form of the introduction of some redundant prepositions as in *discussing the matter or getting into the room*. Such patterns signify overgeneralization, whereby learners believe that some of the verbs must have prepositional complements when they do not.

Combined, these results indicate that prepositional errors are not accidental, but they are systematic errors of learners and indicate interlanguage development (Ellis, 2022). Every type of error is associated with a particular phase of grammatical competence development, and this fact makes the learning of prepositional use in the second language a very complex task.

4.2. Frequency Patterns of Errors

The second research question analysed the rate of occurrence and the distribution of prepositional errors in the corpus of the learners. During the analysis, it happened that certain prepositions were more abused than others, namely, high-functional load and semantic complexity ones.

Table 2: The most frequently used misused prepositions

Preposition	Frequency of Errors
In	29
On	24
At	19
To	17
For	14
With	11
About	09
From	08

The statistics show that the prepositions *in*, *on* and *at* are especially problematic to learners. Such prepositions are common in spatial and temporal phrases, the proper use of which is also subtly dictated by the context. The reason behind the high rate of errors made with these forms is that it is believed that the learners have difficulty internalizing these distinctions, instead they tend to make approximations or general rules.

Table 3: General Usage Statistics of Prepositions

Category	Value
Total Corpus Size	8,450 words
Total Preposition Used	382
Correct usage	244
Incorrect usage	138
Error rate	36%

Such an overall rate of error (36 percent) is indicative of the difficulty involved in the proper use of prepositions by learners. This quite large percentage indicates that even on the undergraduate level prepositional competence has not yet been developed.

The frequency patterns also imply that learners will tend to try to make use of prepositions rather than avoid it. This is a considerable point, because it denotes that the learner is actively engaged in the target language, and that s/he has not yet acquired full command of its grammatical structure. This interpretation is also supported by the fact that there are numerous instances of substitution errors that learners make, and the errors seem to be based on partial knowledge when learners must choose prepositions.

4.3. Interlanguage and Cross-Linguistic Influence.

The third research question was how L1 affects the use of English prepositions as a long-lasting challenge in L2 learning (Tyler et al., 2022). The results indicates that cross-linguistic transfer can be important in determining learner errors.

Table 4: Types of errors and probable causes

Error Type	Linguistic Cause
Substitution	Semantic confusion, partial knowledge
Omission	L1 structural transfer (Urdu postposition)
Addition	Overgeneralization of rules

Omission errors are manifested in the structural differences between Urdu and English. Relational meanings are usually conveyed by means of postpositions (which follow the noun phrase) in Urdu, but in English the preposition should come before the noun phrase. It is this variation that causes learners to exclude prepositions in situations in English where they are mandatory.

L1 conceptual mapping can also be observed through substitution errors. Students tend to choose prepositions through direct translation of Urdu and end up using them in non-target-like form. This implies that when learners comprehend English relationships, they use L1-based semantic frameworks.

On the other hand, additional errors are developmental processes of interlanguage. Learners are likely to overgeneralize grammatical rules by putting them in prepositions in places where they do not belong. This implies that students are busy developing hypotheses regarding the structure of the English language, although these hypotheses may not always be true.

4.4. Patterns of Collocation and Concordance.

Corpus tools were used to analyse concordance to understand further the behaviour of the learners. This discussion showed that there were common patterns of incorrect verb-preposition combinations.

Table 5: Incorrect Collocations

Verb	Incorrect Form	Correct Form
discuss	discuss about	discuss
enter	enter into	enter
explain	explain about	explain
request	request for help	request help
reach	reach at	reach

Such tendencies point to the fact that learners tend to use analogy reasoning and believe that some verbs have prepositions that they should have due to the perceived similarities with other constructions. This also confirms the interlanguage development in influencing the errors of learners.

5. Limitations

Although the current study is methodologically sound and analytical in nature, it is limited in several ways which must be considered when interpreting the results. First, the limitation of small-scale learner corpora in terms of size (McEnery and Hardie, 2022), sixty-essay based learner corpus gathered at two universities. Although this dataset is adequate to establish common trends of prepositional mistakes, it might not be able to reflect the variety of learner languages in various institutions, locations, or skill levels. The generalizability of the findings would be improved with a larger corpus with more universities and a larger range of participants.

Second, the experiment uses an elicitation task based on translations as the main source of data collection (Mackey and Gass, 2022). Despite the effectiveness of this approach in making sure that prepositional contexts are present, it can affect the character of the learner's output. Translation tasks are more likely to stimulate the use of first language structures and thus increase the impact of cross-linguistic transfer.

The other limitation is associated with the application of computational tools to linguistic analysis. Although data processing efficiency is greatly improved with the help of the tools like part-of-speech tagging and concordance analysis, they do not remain error-free. Tagging systems can sometimes improperly categorize word usage, especially when dealing with texts written by learners which have non-standard forms. Even though it was verified manually and thus it could be assumed that it was correct, the chances of including small inconsistencies cannot be dismissed entirely.

Lastly, the research is quantitative in nature (descriptive) with a qualitative interpretation. Although this framework is applicable to identifying patterns and describing underlying processes, it does not have more advanced inferential statistical analysis. Future research may adopt statistical modeling techniques to investigate the relationship between such variables as proficiency level, frequency of error and contextual use. A remedy for such limitations would help to improve empirical and analytical rigor of research in this area.

6. Conclusion

The aim of the current research was to investigate the prepositional errors in L2 English writing among the Urdu-speaking undergraduates in a corpus based computational approach. The paper involved the aim of providing a comprehensive picture of the use and wrong use of the English prepositions in written discourse both in integrating the examination of the quantitative frequencies with the qualitative interlanguage meaning. The findings of the study are useful as they offer information about the concept of learner language, cross-linguistic influence, and the future of computational study of second language.

Some of the key results of the research are that the prepositional errors are not merely common but also systematic in the writing of the learners. This discussion revealed that prepositional use in the corpus had a high percentage of incorrect meaning they continue to have major difficulties in mastering this type of grammar. This categorization of the errors into the groups of substitution, omission and addition indicates different levels of challenge of the learner. Substitution errors were the most frequent, and it means that learners possess some knowledge concerning the usage of prepositions, but not how to use a specific form in each situation. Structural differences between Urdu and English are reflected in the omission errors and overgeneralization of grammatical rules in the form of addition errors.

The studies of Neumanova (2023), Aybk and Can (2021 and Alrashidi (2022) have provided evidence that errors which learners make are not random, but it shows a growing system of interlanguage. Analyzing the errors in this study through the interlanguage theory, we might suppose that they are the aspects of the language learner developing. The students will positively construct theories about the target language and will use their background knowledge about their native language and the one that they are building with English.

One of the factors that were found to have a huge impact on the development of errors among the learners was the influence of cross-linguistics. It was noted that structural differences between Urdu and English words especially use of postpositions in Urdu was a contributing factor to omission error in writing in English language. Besides this, learners tended to use direct translation of Urdu in choosing prepositions and as such, they made substitution errors, which were indicative of conceptual mapping based on L1. These results underscore the fact that the linguistic backgrounds of learners should be considered when studying the process of acquiring a second language. Knowledge of the effects of the first language can be used to learn more effectively as the origin of the learner's difficulty can be understood.

Moreover, further studies can be used to analyze the relationship between prepositional mistakes and other aspects of language proficiency such as vocabulary knowledge, syntactic complexity and writing quality. The study of these relations can provide a more detailed picture of learner language and results in the development of detailed models of the second language acquisition.

In conclusion, the present research paper adds to the field of applied linguistics significantly since the study is conducted in detail, and prepositional errors are considered in writing L2 English with the help of a computational approach based on a corpus. The findings highlight how the errors of the learners are systematic, how cross-linguistic transfer works, and how computational tools and the linguistic analysis should be combined. The study enriches our knowledge of learner interlanguage by integrating both theoretical knowledge and empirical evidence and provides practical implications to language teaching and research.

References

- Aybek, S., & Can, C. A. (n.d.). *A bird's-eye view of corpus-driven interlanguage error analysis of Portuguese EFL learners and data-driven teaching suggestions for EFL teachers*. In *36th Annual APPI Conference: Book of proceedings* (p. 51). https://www.researchgate.net/profile/Sibel-Aybek-2/publication/380394281_36th_Annual_APPI_Conference_The_Book_of_Proceedings/links/663a43e506ea3dob742f238b/36th-Annual-APPI-Conference-The-Book-of-Proceedings.pdf
- Aybek, S. (2025). Phraseological patterns in learner academic English: Insights from corpus-driven approaches. *Journal of the European Second Language Association*, 9(1). <https://euroslajournal.org/articles/138>
- Aybek, S. (2025). To collocate or not to collocate: Exploring verb–noun collocations of Turkish EFL learners. *SAGE Open*, 15(4), 21582440251395720. <https://doi.org/10.1177/21582440251395720>

- Alrashidi, R. S. (2022). *The effects of data-driven learning on intermediate language learners' use of metadiscourse markers in argumentative essay writing (A quasi-experimental study)* (Doctoral dissertation, University of Liverpool). <https://search.proquest.com/openview/ea40f2c80fe05f760a8b863df9907d95>
- Anthony, L. (2022). *AntConc (Version 4.0)* [Computer software]. Waseda University.
- Bardovi-Harlig, K., & Dörnyei, Z. (1998). Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *TESOL Quarterly*, 32(2), 233–262. <https://doi.org/10.2307/3587583>
- Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, 19(4), 207–217. <https://doi.org/10.1016/j.jslw.2010.10.002>
- Brezina, V. (2023). *Statistics in corpus linguistics*. Cambridge University Press.
- Cohen, A. D., Manion, L., & Morrison, K. (2022). *Research methods in education* (9th ed.). Routledge.
- Cook, G. (2021). Translation in language teaching. *Oxford Research Encyclopedia of Linguistics*. <https://doi.org/10.1093/acrefore/9780199384655.013.549>
- Dörnyei, Z. (2022). *Questionnaires in second language research* (2nd ed.). Routledge.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford University Press.
- Ellis, R. (2022). *Second language acquisition*. Oxford University Press.
- Ellis, R., & Barkhuizen, G. (2022). *Analysing learner language* (2nd ed.). Oxford University Press.
- Etikan, I., & Bala, K. (2021). Sampling and sampling methods. *Biometrics & Biostatistics International Journal*, 10(3), 87–88.
- Gazıoğlu, M., & Aydın, S. (2024). Identifying grammatical errors and mistakes via a written learner corpus in a foreign language context. *Journal of Language Research*, 8(2), 91–106. <https://doi.org/10.51726/jlr.1553484>
- Gilquin, G., Granger, S., & Paquot, M. (2023). Learner corpus research and pedagogy. *Applied Linguistics Review*, 14(3), 467–488.
- Granger, S. (2014). *Learner English on computer*. Routledge.
- Granger, S., Gilquin, G., & Meunier, F. (2022). *The Cambridge handbook of learner corpus research*. Cambridge University Press.
- Hyland, K. (2022). *Second language writing*. Cambridge University Press.
- Jarvis, S., & Pavlenko, A. (2021). *Crosslinguistic influence in language and cognition*. Routledge.
- Jurafsky, D., & Martin, J. H. (2023). *Speech and language processing* (3rd ed., draft).
- Mackey, A., & Gass, S. (2022). *Second language research: Methodology and design*. Routledge.
- McEnery, T., & Hardie, A. (2022). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
- Neumanová, Z. (2023). Investigating L2 English preposition use by Czech university students: A learner corpus study. *Ostrava Journal of English Philology*, 15(1). <https://doi.org/10.15452/OJoEP.2023.15.0006>
- Paquot, M., & Granger, S. (2021). Learner corpora and SLA. *Annual Review of Applied Linguistics*, 41, 123–140.
- Tyler, A., & Evans, V. (2003). *The semantics of English prepositions: Spatial scenes, embodied meaning and cognition*. Cambridge University Press.
- Tyler, A., Mueller, C., & Ho, V. (2022). Cognitive linguistics and L2 prepositions. *Language Teaching Research*, 26(2), 345–366.