Reading & Writing - Journal of the Literacy Association of South Africa
ISSN: (Online) 2308-1422, (Print) 2079-8245

Exploring the relationship between syntactic complexity, error types, and pedagogical implications in advanced writing: A quantitative analysis

Authors:
Muhammad A. Mahmood[1]
Noshaba Bano[1]
Muhammad Imran[2,3]
Norah Almusharraf[4]

Affiliations:
[1]Department of Applied Linguistics, Faculty of Arts and Social Sciences, Government College University Faisalabad, Faisalabad, Pakistan

[2]Educational Research Lab, College of Sciences and Humanities, Prince Sultan University, Riyadh, Saudi Arabia

[3]Department of English, School of Humanities, Education and Social Sciences, Khazar University, Baku, Azerbaijan

[4]Educational Research Lab, College of Sciences and Humanities, Prince Sultan University, Riyadh, Saudi Arabia

Corresponding author:
Muhammad Imran,
mimran@psu.edu.sa

Dates:
Received: 16 June 2025
Accepted: 16 Jan. 2026
Published: 17 Feb. 2026

Background: This study investigates the relationship where common linguistic errors may act as constraints on the attainment of syntactic complexity and explores the resulting pedagogical implications in advanced writing.

Objective: The exploration of error patterns has been conducted to analyse sentence structure and instructional strategies in the advanced writing teaching process. The main objective of the research is to find a connection between the common linguistic errors that might serve as barriers to syntactic complexity being achieved, and a discussion of consequent pedagogical implications in advanced writing.

Method: The investigation employs a quantitative approach in which the essays of 100 undergraduate students are analysed, and the relationship between the different features of syntactic complexity is examined. The features of syntactic complexity in the research studied include the mean number of words before the verb (SYNLE), the average number of modifiers per noun phrase (SYNNP), and sentence-to-sentence similarity. In contrast, the types of linguistic errors analysed include noun errors, form errors, and punctuation mistakes. The syntactic complexity features are drawn from the Coh-Metrix software, which provides measures of text cohesion and coherence.

Results: The study reveals a significant negative correlation between noun and form errors and syntactic complexity. These errors, in particular, lead the writers to simplify their sentence structures, resulting in shorter pre-verbal segments and a drop in cohesion and coherence. In contrast, other error types, such as lexis and punctuation errors, have very weak or no correlation with the features of syntactic complexity.

Conclusion: Highlighting noun mistakes and grammar errors as writing instruction issues is one of the most significant findings these results point to. They also indicate that the inclusion of syntactic complexity metrics in automated writing systems may help provide very nuanced feedback, thus allowing the writers to work on their sentence sophistication and the coherence of their texts.

Contribution: The study has a major impact on research in second language writing and syntactic analysis, as it empirically demonstrates that noun and form errors are associated with sentence complexity and textual cohesion. In addition, it has led to the drawing of practical insights for writing training and the creation of automated feedback tools aimed at refining syntax and improving coherence in advanced second language writing.

Keywords: syntactic complexity; noun errors; syntactic cohesion; advanced writing; writing instruction; pedagogy.

Introduction

The ability to write well is an essential requirement for both academic and professional success, especially in situations that require advanced linguistic and cognitive abilities (Biber et al. 1998). Syntactic complexity, an indicator of advanced writing, is a significant feature that characterises

How to cite this article: Mahmood, M.A., Bano, N., Imran, M. & Almusharraf, N., 2026, 'Exploring the relationship between syntactic complexity, error types, and pedagogical implications in advanced writing: A quantitative analysis', *Reading & Writing* 17(1), a587. https://doi.org/10.4102/rw.v17i1.587

high-quality writing and plays an important role. It shows the writer's ability to create complex sentence structures by using conjunctions, nesting, and a mix of syntactic patterns (Ahmad et al. 2024; Lu 2010; Rashid et al. 2024). Thus, syntactic complexity is very important for building and clarifying the two aspects of good writing: its logicality and engagement (Crossley & McNamara 2016).

On the other hand, despite its significance, many university students who attend English as a second language (ESL) classes do not reach the expected level of syntactic complexity in their writing. On the contrary, they mostly use basic sentence structures, which limit the overall quality of their texts (Ambreen, Ozdemir & El-Dakhs 2023; Housen, Kuiken & Vedder 2012; Polyanskaya et al. 2024.). The use of simplicity is not a purposeful stylistic choice but rather arises from linguistic difficulties, such as noun and verb errors and a lack of knowledge of advanced syntactic features. The main part of second language acquisition (SLA) research is devoted to the interactions and influences of complexity, accuracy, and fluency (CAF) on each other (e.g. complexity leading to drop in accuracy); however, this study takes an inverted viewpoint: it posits that high rates of linguistic errors act as a restraining factor that lowers an author's willingness and capacity to use complex syntax.

The disparity between the necessity of syntactic complexity and students' reliance on simpler structures underscores the need for a clear understanding of the factors that hinder their writing development. The past research has always pointed to linguistic errors, notably nouns and verbs, and improper punctuation, as the major factors preventing writers from achieving the desired syntactic sophistication (Ahmad et al. 2024; Ferris 2002; Truscott 2007). Errors of this kind affect students' sentence fluency and coherence, gradually eroding their confidence and forcing them to avoid complex sentences altogether. The presence of syntactic complexity in advanced writing has been widely researched, but the interaction between linguistic errors and syntactic simplicity has received very little attention (Ortega 2012). As a result, the current study intends to fill this lacuna by examining the extent to which linguistic errors dictate the use of complex syntactic structures and how this, in turn, impacts the cohesion and coherence of academic writing.

This disconnection between the need for syntactic complexity and students' reliance on simpler structures raises essential questions about the factors impeding their writing development. Research has consistently identified linguistic errors, such as nouns, verbs, and punctuation errors, as key barriers to syntactic sophistication (Ahmad et al. 2024; Ferris 2002; Truscott 2007). These errors disrupt sentence fluency and coherence, ultimately eroding students' confidence and compelling them to avoid complex constructions altogether. Although existing studies have extensively explored the role of syntactic complexity in advanced writing, far less attention has been given to the interplay between linguistic errors and syntactic simplicity (Ortega 2012). Therefore, this study aims to address this gap by investigating how linguistic errors influence the use of complex syntactic structures and, consequently, affect the cohesion and coherence of academic writing.

Syntactic complexity is a major source of difficulty for students at university, which, in turn, affects the coherence and cohesion of their writing. Besides, linguistic errors, such as inaccurate verb use and incorrect noun agreement, hinder the fluency of students' sentences (Bodnar & Kadlowec 2018; Ellis & Barkhuizen 2005). Subsequently, students' writing loses depth and clarity, becoming simpler. The basic cause of this problem is the lack of instructional emphasis on features such as number of modifier-based phrases in a noun phrase (SYNNP) and syntactic length of specified word units (SYNLE), which are essential for the development of syntactic complexity. Thus, without intensive interventions, students will remain stuck in a cycle of linguistic errors and syntactic simplicity, which in the end will result in their writing being no more than at a basic level and the expected development in writing proficiency being completely blocked (Bodnar & Kadlowec 2018; Housen & Simoens 2016).

In this study, syntax and writing quality are closely tied, offering insights into how linguistic challenges influence students' writing practices (Rababah, Rababah & Naji Al-Khawaldeh 2024; Imran, Almusharraf & Dalbani 2024).

This study's inversion of the error-complexity relationship contributes a novel perspective to the CAF framework in SLA. While CAF typically measures how writers balance these three constructs, our approach empirically demonstrates that a deficit in accuracy (specifically noun and form errors) directly constrains the potential for complexity.

Additionally, implications for teaching have been extracted from this research. It offers considerable ideas about various methods that can be used in teaching. One of the methods that is discussed is the need for integrated teaching strategies, which would simultaneously focus on linguistic errors and encourage the use of complex sentence structures through parameters such as the SYNLE and SYNNP (Busch 2012).

Consequently, this would lead to greater text-based cohesion and coherence, and eventually the students would reap the benefit in terms of improved academic and professional work performance.

Finally, this article contributes to the field of SLA by establishing a clear connection between linguistic error analysis and the development of syntactic complexity. This foundation provides the basis for further work on innovative ways to develop writing pedagogy so that students are equipped to develop what would be considered advanced writing content.

Literature review

Syntactic complexity posits the sophistication of sentence structures by statements of increasing subordination, clause

embedding, and syntactic diversity. Studies (Biber et al. 1998; Jagaiah, Olinghouse & Kearns 2020; Shah 2024) have concluded that it is widely acknowledged as indicative of advanced writing in various professional and academic contexts, with a high profile of cognition and language. The often-used statistics to quantify syntactic complexity in texts are the number of words on average before the main verb (SYNLE) and the number of modifying phrases in a noun phrase (SYNNP) (Graesser, Wiemer-Hastings & Wiemer-Hastings 2008; Takeshi 2018).

Linguistic errors present similar barriers to attaining syntactic complexity; they come in diverse categories, such as those related to nouns, verb forms, punctuation, lexis, pronouns, connectors, adverbials, and noun phrases (Ahmad et al. 2024; Jagaiah et al. 2020). It has also been shown that these mistakes hinder the flow of sentences and force authors to simplify their sentence structures to reduce the likelihood of mistakes (Ferris 2002; Truscott 2007; Kyle & Crossley 2018.). Noun errors, for example, could be errors in agreement, plurality, or misplaced words which would lead to the incapacity to use complex noun phrases and, thus, result in no promotion of fluency in sentences (Hyland 2019). Likewise, verb tense and morphology errors could result in fewer examples to form a dependent structure in the subordination since they will have less provision of the subordinating conjunctions thereby completely preventing subordination (Ellis & Barkhuizen 2005). Pronoun or connector errors will make coherence between sentences very weak and, hence, rob the writers of the good usage of the advanced discourse markers (Afzaal et al. 2022; Nation 2008). This reflects the creation of a negative feedback loop where writers due to fear of errors refrain from using complex structures which in turn cuts off their ability to build up syntactic complexity down the feedback chain, thereby reducing overall cohesion and coherence.

One of the suggestions considered promising for future studies is to combine the advanced linguistic theories with insights from technology to get an understanding of the mechanisms behind these interactions. The focus of recent efforts in second language (L2) writing has been on the involvement of automated tools and machine learning models in correcting linguistic errors. The progression in AI-assisted tools has provided very detailed information on errors especially for ESL learners (Afzaal et al. 2019; Almashy et al. 2024). Moreover, they have been recognised as having potential in identifying problems of cohesion at the discourse level characterised as being higher order, thus deterring the possibility of detection by traditional linguistic analyses (Zakaria & Aryadoust 2024). Recently, there have been studies on the influence of cultural differences and mother tongue (L1) oral languages on the development of syntactic complexity over the years. A relatively new study has proved this (Imran & Ain 2019).

First-language interference has been shown to significantly shape error patterns in ESL writing, as Ahmad et al. (2024) purport. These findings fit most suitably with earlier research by accenting the need for culturally responsive pedagogical strategies. This viewpoint supports earlier discoveries by stressing the necessity of culturally relevant pedagogy. This study accepts the previous findings and sees the broader debate because of a theoretical and technological approach to ESL writing improvement.

In an attempt to ground or support the investigation, three theoretical perspectives lay the foundation for explaining how syntactic complexity, linguistic errors, and writing quality are interrelated. Firstly, the Cognitive Load Theory (CLT) by Sweller (1994) defines the construction of complex syntactic structures as cognitively taxing for writers, particularly in L2 learners. We argue that frequent, high-impact linguistic errors (e.g., in noun and form structures) significantly increase the extrinsic cognitive load. To manage this increased load, a process identified as a compensatory strategy, writers are forced to divert resources away from complex structural planning and deliberately opt for simpler sentence structures.

For example, to avoid subordinate clauses or embedded constructions, writers can focus on the accuracy of grammar. So, learners can write more complexly if they are taught the right intervention method to develop complex sentence patterns.

Secondly, the Processability Theory proposed by Pienemann (1998) states that the internal development stage determines whether relative clause structures are used appropriately. Errors in grammar, vocabulary, and syntax are signs of students' limited processing abilities. The theory further describes a case in which very early learners have difficulty with relative clauses and connectors and therefore rely heavily on simple sentences. This scenario clearly shows the relationship between linguistic errors and writers' tendency to simplify syntax in difficult writing situations.

Finally, Text Model 1 (Kintsch 1988) considers grammatical composition a crucial aspect of improving text readability and understanding. The text models suggest that variety and cohesion in syntax are of utmost importance in creating good texts. Writers, however, in the case of lacking cohesion due to errors in the language, may still hold clarity, but at the same time forget the complexity of the syntax. A case in point is that when writers avoid discursive markers, such as 'however' or 'moreover', they may also fail to notice logical linkages that could have strengthened overall coherence.

Considering the shortcomings unearthed in the preceding studies, the current research has been driven by the questions mentioned. These inquiries aimed to examine the interaction between syntactic complexity and the occurrence of linguistic errors, while also developing tools that could be put into practice to help correct errors in language teaching.

Research questions

- What is the relationship between syntactic complexity features (e.g., SYNLE, SYNNP) and different linguistic error types in advanced writing?
- The query focuses on the correlation between various measures of syntactic complexity and the different types of errors as the mainstay of our statistical analysis.
- How do noun and form errors influence the syntactic sophistication and cohesion of written texts?
- This question ponders the balance required to stick to the rules of grammar while gaining a satisfying level of complexity in the sentence structure.
- Are errors related to lexis, punctuation, and prepositions significant predictors of syntactic complexity?
- This question is set to unravel whether less dominant error types still affect the overall syntactic structure, thereby broadening the domain of error influence.
- How can writers correct nouns and form errors using compensatory techniques, and how does correcting these linguistic errors in writing instruction improve syntactic complexity and coherence?

The question connects the empirical results to the educational intervention, firmly positioning them within the context of possible ways to enhance writing quality through targeted error-correction strategies.

Research methods and design

The researcher adopts a quantitative research methodology for the study, aiming to explore the relationships between various categories of writing errors and syntactic complexity. The research primarily aims to identify trends and correlations between the two variables to provide a quantitative understanding of the impact of linguistic errors on sentence structure.

Research approach

The methodology is built on the following elements. First, a correlational study is adopted to determine the direction and strength of the correlation between error types (e.g., noun errors and form errors) and syntactic complexity (e.g., SYNLE and SYNNP). This method is particularly fitting for the research purpose, as it guides the researcher in observing significant trends in how errors affect the syntactic complexity of writing.

Data collection

Sampling and population

The study made use of purposive sampling whereby participants were selected according to their L1 for the central purpose of analysing L1's influence on ESL writing errors. In very broad terms, 500 students were purposely selected from an approximate university-wide population of 3000 students to include all the major linguistic groups on campus, namely Urdu (191), Punjabi (163), Saraiki (132), Pashto (10), and Balochi (4) speakers. Even if the numbers in the sample do not even out across these groups, the variation reflects the actual demographic proportions within the university and is, therefore, representative for the purposes of this study.

The subjects were requested to voluntarily provide crucial background information regarding gender, mother tongue, and discipline of study, which would, in principle, give their descriptive essays a cue for error categorisation regarding L1 influence. This design ensured representation of each L1 group, even those with small populations such as Pashto and Balochi, in the analysis of the study. Thus, the purposive strategy served in fulfilling the aim of the study, which was to identify and compare L1-specific patterns of error in ESL writing. The data set consists of 100 essays collected from university students across Punjab Province, Pakistan.

Students consented to the use of their written materials in the study and thus upheld ethical issues.

Coh-Metrix is the software tool employed for analysing the acquired syntactic complexity features from the essays. This text analysis tool assesses linguistic parameters such as sentence length, syntactic cohesion, and lexical variation into the generation of features that give quite a precise aspect of syntactic complexity. Errors in essays, such as nouns, punctuation, prepositions, and form, are extracted and classified by the automated tools used for error detection. Such tools ensure a consistent and correct annotation for the whole data set.

Annotation

The annotation process is based on OpenAI Prompts and follows the recognised categories (noun, form, punctuation, and preposition errors) while properly quoting the International Center for Leadership in Education (ICLE) framework.

Data analysis

Data analysis was performed using SPSS (Statistical Package for the Social Sciences). The calculations were carried out in a way that the participants perceive them as trustworthy and reliable.

Correlation coefficient

The Pearson correlation was used to assess the strength and type of relationship between the independent variables (syntactic complexity features) and the dependent variables (error types). The coefficients range from –1 to 1: a negative coefficient indicates an inverse relationship, whereas a positive coefficient denotes a direct relationship.

Significance testing

After calculating the coefficients for variables, the investigator used p-values for the determination of significance. Variables were considered to have a significant relationship if their p-values were below the traditional significance level of $p < 0.05$.

Data visualisation

Scatter plots and correlation matrices graphically picture the relationships between syntactic complexity features and error types. These show the model by correlating standard deviations.

Research design

The study used a non-experimental, correlational design to determine the association between syntactic complexity and types of errors in learner essays. It is vital to mention that the research design is based on models without any treatment or manipulation of variables.

Independent variables

The predictors or independent variables of this study are syntactic complexity features (SYNLE, SYNNP, SYNMED, and SYNSTRUT) characterised in brief as follows:

- SYNLE: This is the average number of words between a noun and a verb, pointing at the level of structure complexity of the sentence.
- SYNNP: This is the average number of words in a noun phrase, indicating the internal complexity and informational density of noun phrases.
- SYNMED: This metric denotes the lowest count of swaps, erasures, and insertions that are required to transform one sentence into another, considering their parts of speech distributions.
- SYNSTRUT: The variable here denotes the extent of syntactic connection affecting any pair of two successive sentences, thus proclaiming the text's overall syntactic variety and consistency.

Dependent variables

Dependent variables are those errors that occur in students' essays:

- Noun errors: These include incorrect placement of singular and plural nouns, subject–verb agreement, object–pronoun agreement, and noun–predicate agreement.
- Form errors: These contain incorrect verb forms, tense variations, and grammatical structures.
- Punctuation errors: A category for problems such as the absence of punctuation or the wrong positioning of punctuation, etc.
- Preposition errors: Mistakes made concerning the usage of prepositions, which, in turn, inhibit the clarity and smoothness of sentences.

Ethical considerations

Data collection was preceded by obtaining informed consent from all participants, which served as assurance of voluntary participation, confidentiality, and the right to withdraw at any time. An application for ethical approval was made to the Ethics Review Committee of the Government College University Faisalabad and ethics consent was received on 09 January 2025. The ethics approval number is GCUF/ERC/536-A.

Results

Syntactic complexity features and error types

This part provides a narrative of the detection's outcome regarding the correlation between various syntactic complexity features and different types of written errors. The results demonstrate certain trends, revealing the extent to which different categories of linguistic errors can affect the overall syntactic complexity and quality of written texts.

Mean words before main verb (SYNLE)

Significant negative correlations:

- Noun errors (–0.233870, $p < 0.05$): A noteworthy negative correlation was found between the number of words before the main verb and noun errors. This indicates that when a writer makes noun errors (e.g. regarding agreement or pluralisation), they tend to use shorter or simpler sentence structures. Consequently, sentences with noun errors are shorter and simpler, with fewer words before the main verb.
- Errors of form (–0.227137, $p < 0.05$): Errors of form, including incorrect verb forms and tense mistakes, have also been identified as having a very strong negative correlation with SYNLE. Such mistakes force the authors to simplify their sentences, which, in turn, reduces syntactic complexity. This simplification in the text can impact cohesion and coherence.
- The sum of errors (–0.208003, $p < 0.05$): The number of errors, including noun, form, punctuation, and preposition errors, shows a weak negative correlation with SYNLE. Writers who produce more errors tend to simplify their sentences; probably, they want to avoid making more mistakes. This obviously would cause the impoverishment of syntax as well as an overall reduction in cohesion and coherence of the text.

All these points are significant contributions of noun and form errors that encourage the writer to syntactic simplification. Thus, writers struggling with these error types tend to use simpler structures that prevent further mistakes at the cost of overall cohesion and coherence in their writing. A negative correlation can be seen in Figure 1.

Average modifiers per noun phrase (SYNNP)

A weak negative correlation was found between lexis errors (related to vocabulary) and several modifiers per noun phrase (–0.114556). This indicates that lexis errors tend to hamper the complexity of noun phrases a little but have little or no effect on the syntactic configuration, for they are not statistically significant.

So, lexical errors hardly affect the noun phrases' syntactic features.

The sum of errors (–0.110798): The combined number of errors also establishes a weak negative association with SYNNP. Thus, the frequency of errors greatly impacts, but the functions of errors fall short regarding the complexity of noun phrases. Furthermore, the weak association indicates that the overall number of errors does not significantly affect the syntactic complexity within noun phrases.

The inferences drawn from these results indicate that the writers' lexical and grammatical misunderstandings can limit their use of synonyms. However, the phrases will still be grammatical even if some are simple. It is the minimal amount of syntactic variety that leads to the observed small negative correlation between the two variables: error rate and complexity of noun phrase form.

Sentence-to-sentence syntax similarity (SYNSTRUT)

Positive correlations include:

- Nominal errors (0.210022): The positive correlation between nominal errors and sentence-to-sentence syntactic similarity indicates that nominal errors tend to occur together with repetitive sentence structures due to writers simplifying their syntax to avoid further errors. As a result, the same findings shrink general syntactic diversity and thus the final coherence of the text.
- Adverb errors (0.150036–0.196245): Adverb errors were found to be positively correlated with sentence-to-sentence syntactic similarity. This means that adverb-related errors have a greater impact on the unification of sentences than on the linguistic variation among them. In this way, the sentence-level consistency is increased even though syntactic diversity, which plays a significant role in cohesion and coherence, is lost.

This result reflects the limitation of sentence variety due to particular error types, nominal and adverb errors groups, the positive correlation depicted in Figure 2, and the sliding similarity of these errors mean that the corresponding sentences are not so different from each other in their structure and thus are not that complex in syntactic terms, hence affecting the overall unity of the text.

Figure 2 shows that nominal errors correlate positively with syntactic similarity (0.210022), and adverb errors show a milder positive correlation (0.17314). This means that authors who make such mistakes are more likely to use the same syntax, resulting in fewer variations and, hence, less diversity and coherence in the text.

Minimal edit distance (SYNMED)

There was a weak correlation between the minimal edit distance measures (SYNMEDpos, SYNMEDwrd, SYNMEDlem) and lexis errors, such as forms, which carried a very slight implication of their effect on syntactic cohesion. The correlations are weak, which means that the measures of sentence similarity remain almost unaffected by the presence of isolated error types.

Indeed, lexis and form errors have a slight negative impact on sentence cohesion but do not result in a major reduction in the text's syntactic diversity or cohesion, making them unnoticeable. In fact, it might even be the case that the very slight effects of these errors on sentence similarity serve to reinforce the assertion that more complex errors, such as noun and form errors, might lead to the most significant switch in syntax and, consequently, a reduction in cohesion.

In short, it reveals that some forms of errors, particularly noun and form errors, are closely related to simple syntactic structures. Compromising due to these errors forces writers to use fewer complex sentences, which could reduce the probability of writing mistakes, but would not strengthen the cohesion and coherence of the text. However, lexis and adverb errors have a more limited relation to syntactic phenomena and so could have a more limited effect on the overall complexity and cohesion of the text.

These results further highlight the need for linguistic error in writing instruction, especially in those concerning nouns or verb forms, as students gain incisive syntactic complexity and cohesion in writing. Future studies may investigate specific intervention methods to reduce such forms of errors and encourage more sophisticated syntactic structures for improving writing proficiency at the university level.

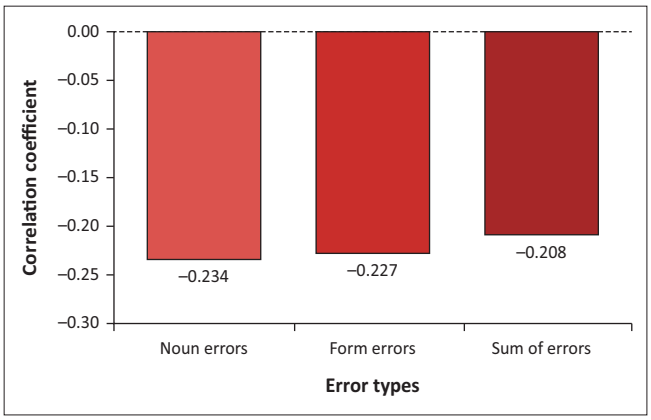

FIGURE 1: Negative correlation between key error types and SYNLE (mean words before main verb) ($p < 0.05$ for all).

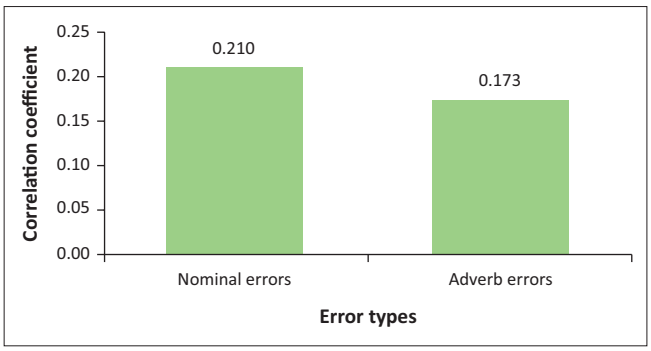

FIGURE 2: Positive correlation between nominal and adverb errors and sentence-to-sentence syntactic similarity (SYNSTRUT).

The correlation between metrics and error types

As shown in Figure 3, this section investigates how specific features of syntactic complexity correlate with various kinds of written errors. It reveals that some types of errors greatly correspond with syntactic complexity while others have little relevance. Those relationships could lead researchers to know how errors react to sentence structure, cohesion, and coherence.

Influence of noun and form errors

Noun error (–0.233870; $p < 0.05$): Strong negative correlation gives an inference that noun errors will be more in number with a decrease in words before the main verb. A person making errors about nouns is said to avoid long and complicated sentences and adopt a pre-verbal segment, shortened to reduce the chances of making more mistakes. Thus, these errors induce less syntactically complex sentences and reduce cohesion and coherence. Writers will more likely turn to simple sentences when faced with complex noun-related problems of agreement or pluralisation, although it also shortens sentences.

Form errors (–0.227137; $p < 0.05$): A negative correlation also holds for form errors like incorrect forms of verbs or tense errors. Writers experiencing these errors will form shorter sentences whose number of words before the verb is as few as possible, as they do not want such sentences to be longer than the normal length measured by the number of words before the main verb. The direct relationship between simplification and textual cohesion lies in its impact on syntactic complexity.

The sum of errors (–0.208003; $p < 0.05$): The overall number of errors negatively correlates with SYNLE and supports the idea that the increased rate of errors leads to simpler sentence structures. Writers use basic sentence structures to avoid many mistakes, which reduces syntactic complexity and textual cohesion.

The errors mainly lie in the use of nouns and forms, and this is what drives simplified structures in sentences, compromising syntactic complexity to avoid a full compounding mistake. Here, the text's cohesion and coherence suffer.

Minimal impact of other error types

Lexis errors (–0.114556, $p > 0.05$): In this particular case, the relationship between lexis errors and syntactic complexity characteristics, particularly with SYNNP, which indicates average modifiers per noun phrase, is slightly negative but not significant. This means that lexical errors have a negligible effect on the syntactic structures. There might be large differences in word choices, but ways do not greatly affect the complexity or cohesion of sentence construction.

Punctuation errors, likewise, are not strongly correlated with any of the syntactic features; hence, their implication concerning sentence structure is rather slight. Punctuation errors do create problems in understanding and fluency at the level of individual sentences, but not in terms of overall syntactic complexity and general uniformity to the same extent.

Sentence similarity and error types

SYNSTRUT (sentence-to-sentence syntax similarity): Sentence similarity characteristics didn't reveal much error-type sensitivity, implying that those measurements are somehow divorced from particular linguistic inexactness. Therefore, these features will be typically stylistic among the sentences and consistent inside the sentence, which will not usually be influenced by localised grammatical errors. Hence, no concern is attached to nouns, form, lexis, punctuation, or any other prevalent errors concerning similarity from sentence to sentence.

These indicate that even though errors such as lexis, punctuation, and prepositions might damage clarity and coherence at the sentence level, they hardly touch overall syntactic complexity or consistency. On the contrary, noun and form errors seem to play a greater role in determining sentence structure.

We present here a summation of our statistical analysis, which mapped the correlations of various parameters of syntax complexity to linguistic error types in texts of an advanced level. Using Pearson's pair correlation analysis, we assessed the relationship of parameters such as the average of words before the principal verb (SYNLE) and the number of modifiers per noun phrase (r) to error types such as noun, form, and lexis errors. The summation of what has been gathered is shown in Table 1.

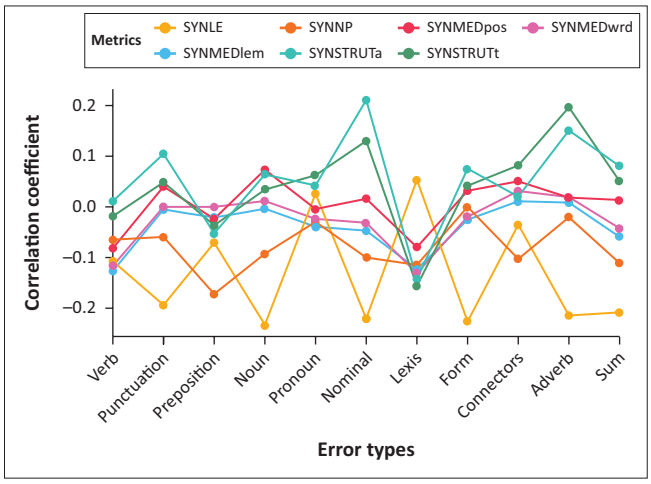

FIGURE 3: Correlation between metrics and error types.

TABLE 1: Correlations between key syntactic complexity features and linguistic error types (selected results).

Syntactic feature	Error type	Correlation coefficient (*r*)	*p*
SYNLE (mean words before main verb)	Noun errors	-0.233	< 0.05
	Form errors	-0.220	< 0.05
	Sum of errors	-0.200	< 0.05
SYNNP (Modifiers per noun phrase)	Lexis errors	-0.110	> 0.05*
	Sum of errors	-0.110	> 0.05*
Syntactic similarity (SYNSTRUT)	Nominal errors	0.210	Not specified
	Adverb errors	0.170	Not specified

*, Not significant.

Joined with the first research question, the variety of error types, the balance in syntax structures, and form and noun errors were found to have a strong intertwined relationship. These concerns connect our comprehension of the importance of the following errors in nouns and form for sentence cohesion.

Compensatory strategy in writing

The analysis states that writers often employ strategies to maintain grammatical correctness despite noun and form errors. The process of doing so involves rephrasing the sentence structure by either truncating the pre-verbal segments or making the latter part of the sentence simpler. Thus, it would be less severe for the writer to construct sentences. The results align well with CLT, which postulates that writers simplify their formulations to manage cognitive resources more efficiently. However, the use of this technique does not reduce the errors that follow, even though it may sometimes reduce sentence cohesiveness or cause the different parts of a sentence not to conform to each other in a unified manner.

Lexical and punctuation errors

These types of errors do not significantly increase syntactic complexity. Among these error types, lexis and punctuation showed a weak correlation with the analysed syntactic features, suggesting they are not at the same level as noun and form errors.

Sentence similarity and error independence

The manner of errors could render them captive to the very holiness of semantics. They could just as well obscure the entire theme of the content from the readers.

SYNSTRUT

The features of sentence-to-sentence syntactic similarity are not easily changed by the different error types, which shows the independence of these measures from the various inaccuracies in linguistics. The errors might cause a loss of clarity in individual sentences, but this does not affect the extent to which identical sentences occur throughout the text. These measures are likely to indicate greater stylistic or writing factors that are not specifically related to certain errors.

The occurrence of error independence does not merely signify a reduction of uniformity across the sentences but rather typifies the complexities of textual coherence.

Discussion

Practical applications

Educational interventions: Target noun and form errors

Writing instruction for syntactic complexity is centred around noun and form error correction. Such a unique position is taken by teaching programmes that aim to make students aware of such errors, thereby enhancing their syntactic sophistication and overall coherence. Writing classes should focus on reducing noun and form errors, as these directly influence students' syntactic complexity. As depicted in Figure 4, these interventions comprise exercises aimed at enhancing sentence structure by avoiding oversimplification using longer, more complex sentences, and students are encouraged to participate in these activities.

The most significant impact on syntactic complexity comes from noun and form errors, which account for 80% of that complexity; thus, these two areas of writing instruction should be a primary focus. Sentence similarity has a lower degree of impact at 50% and lexical and punctuation errors at 30%. This suggests that although they affect clarity and cohesion, these errors do not produce differences in the complexity of sentence structures.

Computational tools: Automated grammar tools

Along with the error-detection, more precise feedback should be offered through the integration of syntactic complexity features in automated writing systems. The grammar-checking software not only catches errors but also evaluates various aspects of syntactic complexity, such as SYNLE, SYNNP and SYNSTRUT. These complexity measures, incorporated into automated systems, will provide feedback to help writers improve the syntactic sophistication and cohesion of their writing.

Furthermore, such systems will make predictions about the construction of more complex and varied sentence structures. This will eventually help the individual writer reduce errors while simultaneously increasing syntactic diversity and textual coherence.

Implementing class activities and strategies

Among the different targeted strategies, the main goal is to correct noun and form errors during writing classes for ESL learners. This may involve taking stated errors to a certain extent, thereby helping reduce errors, while increasing the difficulty of syntax and cohesion. Possible human activities are explained next.

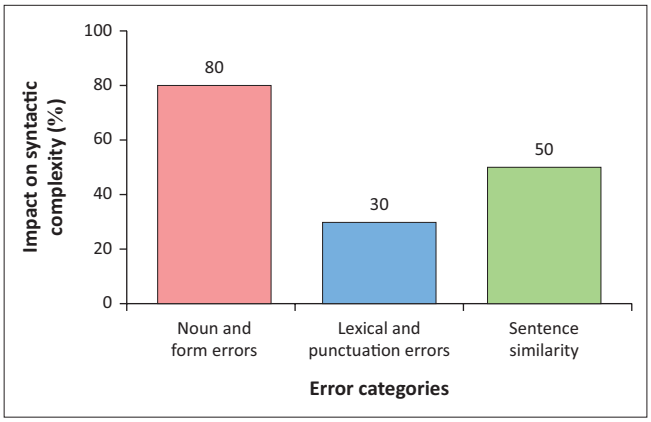

FIGURE 4: Focus areas in educational writing interventions.

Noun-phrase extension exercises: Bring in simple noun phrases to the class and ask students to expand the phrases with modifiers, relative clauses, and other syntactic structures. For instance, start with: 'The teacher explained the topic'. The expanded version might be: 'The experienced teacher whom her students greatly respected explained the intricate topic in a detailed manner'.

This technique helps the student use elaborate noun phrase constructions, enabling them to cope with errors in noun agreement and placement.

Error correction workshops: Students receive different text samples and must mark and correct all errors involving nouns and verbs. After doing so, students might still have these common irregularities in mind, which in turn will help them gain a better understanding of the relationship between grammatical accuracy and syntactic complexity.

Sentence combining exercises: Students are invited to merge several short sentences into a single sentence with more content, for example:

- Simple: She read the book. The book was interesting. She learned a lot.
- Complex: She read an interesting book and learned a lot from it.

This exercise makes the student a master at forming complex subordinate constructions and developing good syntactic variety.

Peer feedback sessions: Small peer feedback sessions function as a gradual process in which students read each other's papers and point out errors in the noun and form outline. The feedback serves as a reminder to students, enabling them to conduct a more informed and critical reading of each other's papers and to be more trusting in the process.

By embedding these strategies within writing instruction, educators may enable their students to identify linguistic problems and subsequently exist in a world of syntactic complexity.

Figure 5 represents the effectiveness of different modes of writing to promote the occurrence of syntactic complexity, marking sentence combining (90% occurrence) and noun-phrase extension (85%) as important. Error correction workshops (75%) and peer feedback (80%) may also perform important tasks in polishing writers, improving their grammatical accuracy, and structural variation in writing.

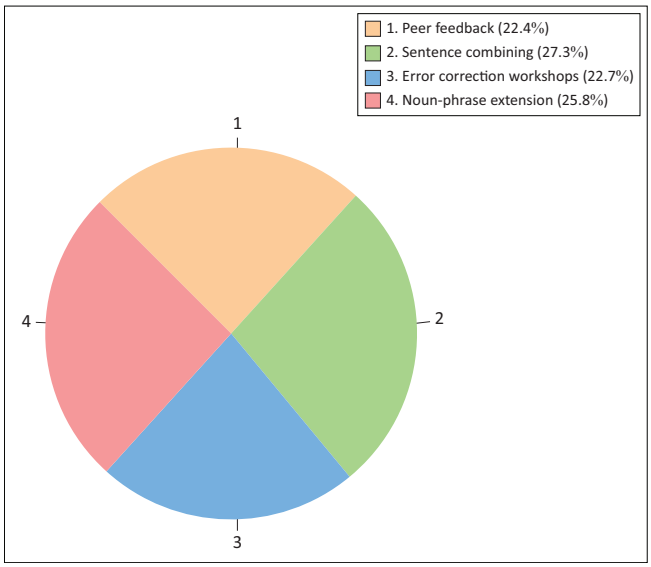

FIGURE 5: Effectiveness of writing strategies in syntactic complexity.

Conclusion

To recap, the present study has much to say about the typology of errors and the relationship between syntactic complexity in advanced writing, thereby providing a better understanding of some of the difficulties that academic writers encounter during the writing process. It has been observed that noun and form errors are largely associated with simplified sentence structures. On the other hand, errors due to lexis, punctuation, and prepositions show an almost negligible impact on syntactic complexity. This implies that errors in nouns and verbs restrict the writer's ability to create or comprehend complex sentences, which in turn affects the sources of cohesion and coherence in their writing. These implications point out that the approach should tackle the specific error types associated with writing, for instance the error types of nouns and form, in writing instruction. Therefore, focusing on these areas will empower teachers to create greater syntactic sophistication, enhancing the overall quality of writing. Moreover, the findings reveal that automated writing systems require incorporating measures such as syntactic complexity to provide a broader feedback dimension beyond error detection alone. Such systems would, in turn, improve the user's writing skills by allowing the user to continually refine their syntactic patterns and embed greater cohesion into the writing. In the long run, future studies might consider the possibilities of having direct interventions for the reduction of noun and form error rates, and the effect on the syntactic complexity that might be taken along the way; thus, these would be the factors that very indirectly still support advanced writing proficiency development.

Acknowledgements

The authors would like to thank Prince Sultan University for financial and technical support.

Competing interests

The authors reported that they received funding from Educational Research Lab, Prince Sultan University, which may be affected by the research reported in the enclosed publication. The authors have disclosed those interests fully and have implemented an approved plan for managing any potential conflicts arising from their involvement. The terms

of these funding arrangements have been reviewed and approved by the affiliated university in accordance with its policy on objectivity in research.

CRediT authorship contribution

Muhammad A. Mahmood: Conceptualisation, Formal analysis, Project administration, Data curation, Supervision. Noshaba Bano: Formal analysis, Investigation, Writing – original draft, Software. Muhammad Imran: Methodology, Visualisation, Writing – review & editing, Funding acquisition. Norah Almusharraf: Formal Analysis, Data curation, Resources, Writing – review & editing, Supervision. All authors reviewed the article, contributed to the discussion of results, approved the final version for submission and publication, and take responsibility for the integrity of its findings.

Funding information

This work was supported by the Educational Research Lab, Prince Sultan University.

Data availability

All data are available in this article; however, for more details, the data will be available on request from the corresponding author, Muhammad Imran.

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. They do not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The authors are responsible for this article's results, findings, and content.

References

Afzaal, M., Hu, K., Chishti, M.I. & Imran, M., 2019, 'A study of Pakistani English newspaper texts: An application of Halliday and Hasan's model of cohesion: A discourse analysis', *International Journal of English Linguistics* 9(5), 78–88. https://doi.org/10.5539/ijel.v9n5p78

Afzaal, M., Imran, M., Du, X. & Almusharraf, N., 2022, 'Automated and human interaction in written discourse: A contrastive parallel corpus-based investigation of metadiscourse features in machine-human translations', *SAGE Open* 12(4), 21582440221142210. https://doi.org/10.1177/21582440221142210

Ahmad, M., Mahmood, M.A., Siddique, A.R., Imran, M. & Almusharraf, N., 2024, 'Variation in academic writing: A corpus-based investigation on the use of syntactic features by advanced L2 academic writers', *Journal of Language and Education* 10(3), 25–39. https://doi.org/10.17323/jle.2024.21618

Almashy, A., Ahmed, A.S.M.M., Jamshed, M., Ansari, M.S., Banu, S. & Warda, W.U., 2024, 'Analyzing the impact of CALL tools on English learners' writing skills: A comparative study of errors correction', *World Journal of English Language* 14(6), 657–667. https://doi.org/10.5430/wjel.v14n6p657

Ambreen, F., Ozdemir, B. & El-Dakhs, D.A.S., 2023, 'Saudi undergraduate EFL learners' attitudes toward using textism in their academic writing', *Asian-Pacific Journal of Second and Foreign Language Education* 8(1), 38. https://doi.org/10.1186/s40862-023-00213-5

Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E., 1998, *Longman grammar of spoken and written English*, Longman, London.

Bodnar, C. & Kadlowec, J., 2018, 'Initial validation of a technical writing rubric for engineering design', *International Journal of Engineering Pedagogy (iJEP)* 8(1), 81–91. https://doi.org/10.3991/ijep.v8i1.7728

Busch, B., 2012, 'The linguistic repertoire revisited', *Applied Linguistics* 33(5), 503–523. https://doi.org/10.1093/applin/ams056

Crossley, S.A. & McNamara, D.S., 2016, 'The role of lexical and syntactic complexity in text comprehension', *Language Testing* 33(1), 3–18. https://doi.org/10.1177/0265532215601881

Ellis, R. & Barkhuizen, G.P., 2005, *Analyzing learner language*, Oxford University Press, Oxford.

Ferris, D.R., 2002, 'Treatment of error in second language writing classes: The writing teacher's perspective', *TESOL Quarterly* 36(1), 28–45. https://doi.org/10.2307/3588352

Graesser, A.C., Wiemer-Hastings, P. & Wiemer-Hastings, K., 2008, 'Constructing inferences and relations during text comprehension', in U. Cress, J. Moskaliuk & H. Jeong (eds.), *Text representation: Linguistic and psycholinguistic aspects*, pp. 249–272, John Benjamins, Amsterdam.

Housen, A., Kuiken, F. & Vedder, I., 2012, 'Complexity, accuracy, and fluency in second language acquisition', *Journal of Second Language Writing* 21(1), 1–3.

Housen, A. & Simoens, H., 2016, 'Introduction: Cognitive perspectives on difficulty and complexity in L2 acquisition', *Studies in Second Language Acquisition* 38(2), 163–175. https://doi.org/10.1017/S0272263116000176

Hyland, K., 2019, *Second language writing*, Cambridge University Press, Cambridge.

Imran, M. & Ain, Q., 2019, 'Effects of non-native instructors' L1, beliefs and priorities on pronunciation pedagogy at secondary level in district Rajanpur, Pakistan', *Journal of Language and Cultural Education* 7(2), 108–121.

Imran, M., Almusharraf, M. & Dalbani, H., 2024, 'Future of writing in higher education: A review of prospective use of artificial intelligence and ChatGPT in academia', in M.M. Asad, P.P. Churi, F. Sherwani & R. Bin Hassan (eds.), *Innovative pedagogical practices for higher education 4.0: Solutions and demands of the modern classroom*, pp. 108–125, CRC Press, Boca Raton FL, USA. https://doi.org/10.1201/9781003400691

Jagaiah, T., Olinghouse, N.G. & Kearns, D.M., 2020, 'Syntactic complexity measures: Variation by genre, grade-level, students' writing abilities, and writing quality', *Reading and Writing* 33, 2577–2638. https://doi.org/10.1007/s11145-020-10057-x

Kintsch, W., 1988, 'The role of knowledge in discourse comprehension: A construction-integration model', *Psychological Review* 95(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Kyle, K. & Crossley, S.A., 2018, 'Measuring syntactic complexity in L2 writing: Comparing multiple indices and finding differences', *International Journal of Applied Linguistics* 28(2), 316–336. https://doi.org/10.1111/ijal.12223

Lu, X., 2010, 'Automatic analysis of syntactic complexity in second language writing', *International Journal of Corpus Linguistics* 15(4), 474–496. https://doi.org/10.1075/ijcl.15.4.02lu

Nation, I.S., 2008, *Teaching ESL/EFL reading and writing*, Routledge, New York, NY.

Ortega, L., 2012, 'Interlanguage complexity', in B. Kortmann & B. Szmrecsanyi (eds.), *Linguistic complexity: Second language acquisition, indigenization, contact*, pp. 127–155, De Gruyter, Berlin.

Pienemann, M., 1998, *Language processing and second language development: Processability theory*, John Benjamins, Amsterdam.

Polyanskaya, L., El-Dakhs, D.A.S., Tao, M., Chu, F. & Ordin, M., 2024, 'The effect of writing script on efficiency and metacognitive monitoring in inferential word learning', *Metacognition and Learning* 19(2), 593–607. https://doi.org/10.1007/s11409-024-09380-3

Rababah, L.M., Rababah, M.A. & Naji Al-Khawaldeh, N., 2024, 'Graduate students' ChatGPT experience and perspectives during thesis writing', *International Journal of Engineering Pedagogy (iJEP)* 14(3), 22–35. https://doi.org/10.3991/ijep.v14i3.48395

Rashid, S., Malik, S., Abbas, F. & Khan, J.A., 2024, 'Pakistani students' perceptions about knowledge, use and impact of artificial intelligence (AI) on academic writing: A case study', *Journal of Computers in Education* 4, 1–26. https://doi.org/10.1007/s40692-024-00338-7

Shah, S.R.A., 2024, 'Investigating syntactic complexity in Qatari EFL learners' essays: A mixed methods study', *Pakistan Languages and Humanities Review* 8(3), 515–532. https://doi.org/10.47205/plhr.2024(8-III)47

Sweller, J., 1994, 'Cognitive load theory, learning difficulty, and instructional design', *Learning and Instruction* 4(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Takeshi, K.A.T.O., 2018, 'Refinement of L2 complexity measurement model in writing: A narrative review', *Data Science in Collaboration* 2, 1–10.

Truscott, J., 2007, 'The effect of error correction on learners' ability to write accurately', *Journal of Second Language Writing* 16(4), 255–272. https://doi.org/10.1016/j.jslw.2007.06.003

Zakaria, A. & Aryadoust, V., 2024, 'A scientometric analysis of applied linguistics research (1970–2022): Methodology and future directions', *Applied Linguistics Review* 15(4), 1595–1619. https://doi.org/10.1515/applirev-2022-0210