Journal of English Language Teaching
Vol. 68 No. 5 (September-October 2026)
DOI: https://doi.org/10.66121/5844ae05

Mapping Linguistic Errors in English Writing: A Quantitative Study of Ninth-Grade Students

Santosh Kumar Dubey[1] & Surjoday Bhattacharya[2]

ABSTRACT

This study presents empirical evidence on linguistic errors among secondary school learners in India, using a validated error-classification framework. The study aimed to determine the frequency and distribution of linguistic errors in English writing, compare errors across school type and gender, and test the reliability of the error-coding rubric. A quantitative descriptive research design was adopted, with data collected from a stratified random sample of 185 students drawn from two secondary schools, one government and one private, in Bhadohi district, Uttar Pradesh. Students completed a standardised dictation task, and their written responses were analysed using a validated rubric classifying errors as orthographic, morphological, syntactic, lexical, and punctuation-related. Inter-rater reliability was estimated using Fleiss' Kappa (0.534–0.678) and intraclass correlation coefficients (0.756–0.845). The findings reveal a high error density (40.29%), with orthographic errors occurring most frequently, followed by syntactic and morphological errors. A Mann-Whitney U test showed no significant difference in error density by school type ($U = 3870.5$, $p = .449$) or gender ($U = 4092.50$, $p = .639$), suggesting that writing challenges are more closely linked to instructional factors than to demographic characteristics. The validated coding rubric demonstrated satisfactory reliability for structured linguistic error analysis. These findings provide empirical evidence for English language teaching in the Indian ESL context and underline the need for explicit instruction in spelling, grammar, and sentence construction, alongside diagnostic assessment practices and teacher professional development to strengthen writing proficiency.

KEYWORDS: Error Analysis; Dictation; English Writing; Orthographic Errors; Inter-Rater Reliability.

INTRODUCTION

Writing is an important skill that trains individuals to express their thoughts, emotions, and experiences systematically and effectively. It is more than a medium of communication; it is a foundation for developing reflection, analysis, logic, and creativity. Through writing, students build their language skills, vocabulary, and grammar. Writing also fosters self-confidence, learning, and introspection, and is highly significant for higher education and professional life, since reports, research papers, applications, e-mails, and other professional correspondence are all produced through writing. The development of writing skills is therefore crucial not only for educational achievement but also for personal, social, and professional success (Klimova, 2013; Bora, 2023; Lee & Schmidgall, 2020).

According to Jabeen et al. (2015) and Gjuzi and Meta (2026), error analysis and feedback are critical to students' linguistic development: identifying the errors students make helps to strengthen their writing skills and improve the quality of language pedagogy. Error analysis (EA) provides an effective method for the systematic study of linguistic structures and errors, helping researchers understand the process of language learning and the errors students commonly make. EA reveals the nature of language development and the causes underlying errors, providing a foundation on which teaching materials are developed and corrective feedback is tailored to students' differing needs (Sarkar, 2025; McDowell, 2020; Wolf, 2022; Mahdiyah et al., 2023; Heydari & Bagheri, 2012; Hassan & Rami, 2024). Internationally, a wide range of studies have examined language errors in English writing across different educational settings. Previous studies have found errors related to orthography, morphology, syntax, lexis, and punctuation among students learning English as a

[1] Assistant Professor, Dayawanti Punj Post Graduate College, Sitamarhi, Bhadohi, Uttar Pradesh, India, 0000-0002-2958-8368, santoshdubey@zohomail.in

[2] Associate Professor, DR. Ram Manohar Lohiya Government Degree College, Jaunpur, Uttar Pradesh, India, 0000-0002-2286-1501, surjoday@gmail.com

second or foreign language (Jabeen et al., 2015; Sarkar, 2025). Studies have shown that orthographic weaknesses affect sentence-construction skills, and that first-language transfer, limited vocabulary, and weak grammar are also major causes of writing errors (Jabeen et al., 2015; Bushi & Meta, 2026; Sarkar, 2025).

Research on English writing in India has grown notably over the past decade; however, numerous gaps remain. Most studies focus on university or higher-education students, while secondary-level students, for whom the foundational development of writing skills is taking place, have received comparatively little attention (Sarkar, 2025; Jabeen et al., 2015). Additionally, most Indian studies are limited to descriptive identification of writing errors, with a corresponding lack of standardised quantitative profiling based on validated error coding. Similarly, comparative analyses of writing errors by gender and school type (government and private) remain scarce, even though such studies could contribute significantly to developing inclusive, evidence-based teaching strategies (Sarkar, 2025).

A further methodological limitation concerns the reliability of linguistic error classification. Although error taxonomies are widely used, evidence of inter-rater reliability is reported comparatively rarely, making it difficult to establish the consistency and reliability of these taxonomies (Sarkar, 2025). Furthermore, dictation is globally recognised as an effective teaching method that combines listening, spelling, grammatical processing, and writing skills; however, in the Indian ESL context, its potential as a diagnostic tool for standardised quantitative analysis of language errors remains underutilised (Sarkar, 2025).

These methodological and contextual limitations underline the need for studies of English writing at the secondary level that use a validated error-classification framework and reliable quantitative analysis. Such studies can help build students' objective understanding of their linguistic difficulties, while also laying the groundwork for quality improvement in English writing, evidence-based classroom practice, and targeted teaching strategies (Jabeen et al., 2015; Bushi & Meta, 2026; Sarkar, 2025).

Considering these research gaps, the present study conducted a quantitative analysis of linguistic errors in the English writing of ninth-grade students in Bhadohi district. A standardised dictation task was used, and a validated error-coding rubric was administered. The first objective was to determine the number of linguistic errors and the error density produced by students. The second objective was to analyse the distribution of errors across the orthographic, morphological, syntactic, lexical, and punctuation categories. The third objective was to compare error patterns by gender and school type. The fourth objective was to establish the reliability of the error-classification framework using inter-rater agreement. Grounded in established error-analysis theory, this study presents context-specific empirical evidence for English language teaching (ELT) within the Indian secondary education system.

REVIEW OF LITERATURE

Linguistic EA systematically identifies and categorises language inaccuracies, providing a framework to examine writing errors among ninth-grade students. It focuses on grammatical and syntactical errors in sentence construction, revealing students' specific challenges and guiding targeted instructional interventions. Research in this area underscores that analysing grammatical rules, vocabulary use, and structural patterns is essential for developing effective strategies to improve students' English writing proficiency.

Theoretical Foundation

Error analysis (EA) has emerged as one of the most effective frameworks for understanding writing development in a second language (L2). EA holds that many errors emerge from the developmental process of language learning, in contrast to the contrastive analysis hypothesis, which held that the main cause of students' errors was first-language interference (Corder, 1967; Khalifa, 2018). Stephen Pit Corder proposed that students' errors are not merely deficiencies but systematic indicators of language development (Corder, 1967). Similarly, Larry Selinker, through interlanguage theory, clarified that learners construct a developing linguistic system distinct from both their native and target languages (Selinker, 1972).

Since then, the theoretical scope of EA has expanded. Comparative analyses of different theories of second language acquisition (SLA) show that no single theory can account for all types of student error (Khalifa, 2018). For example, according to Khalifa (2018), contrastive analysis is best suited to explaining errors arising from language transfer, whereas EA more suitably accounts for developmental errors. Similarly, universal grammar helps explain students' inherent grammatical intuitions, while monitor theory

underlines the role of conscious language monitoring (Khalifa, 2018). Lennon (2008) likewise proposed that SLA research has moved beyond predicting errors through language comparison toward understanding students' cognitive processes through interlanguage theory. Taken together, these approaches make clear that writing errors arise not only from first-language transfer, but also from linguistic knowledge, cognitive processes, learning theories, and the interactive nature of the teaching experience.

In recent years, EA has been further enriched by cognitive and sociocultural approaches. Almineeai (2025), drawing on Chomsky's Minimalist Program to explain syntactic errors, proposed that grammatical inaccuracies reflect disruptions in students' internal computational processes. Li et al. (2024), focusing on the role of artificial intelligence and sociocultural learning environments, argued that the development of writing skills is also strongly influenced by interactive feedback and technology-assisted instruction. Although the theoretical foundations of these approaches differ, they converge on the view that student errors provide important evidence for understanding language development and designing more effective teaching methods. Drawing on these approaches, this study adopts EA as its principal conceptual framework for analysing the linguistic difficulties present in the English writing of Indian secondary school students.

Methodological Approaches

Methodological approaches to linguistic EA have shifted from purely descriptive observation to more rigorous quantitative designs that use structured error taxonomies and statistical analysis (Lasaten, 2014; Isma et al., 2023). Most studies analyse essays, narratives, and descriptive compositions, which offer authentic writing samples but often introduce variability in content and complexity (Nadia, 2024; Hafizhoh et al., 2025). Researchers are increasingly adopting standardised error-classification frameworks to improve continuity and consistency (Dobrić & Sigott, 2014); however, reporting of inter-rater reliability remains comparatively rare, and few studies use standardised writing tasks, limiting the reliability of their conclusions. Although dictation is recognised as a valuable tool for integrating listening and writing processes, its use in quantitative error analysis remains limited (Felker et al., 2019; Fraca de Barrera, 2007).

Pedagogical Implications

Existing literature confirms that EA is an effective tool for improving the pedagogy of English writing. By identifying recurring linguistic errors, teachers can provide targeted feedback, design remedial teaching, and strengthen students' grammar and writing skills (Rahman et al., 2017; Hamed, 2018). Recent studies appropriately advocate for integrating error analysis into formative assessment and technology-assisted feedback to support individualised learning (Mahdiyah et al., 2023; Li et al., 2024). Despite this importance, few studies have translated these insights into teacher education and curriculum design within the context of Indian secondary schools.

Previous studies have significantly advanced the understanding of linguistic errors; however, three important gaps remain. First, few studies have examined secondary-level students within the Indian ESL context. Second, very few studies have established the reliability of error classification using formal inter-rater agreement measures (Dobrić & Sigott, 2014). Third, standardised quantitative studies comparing patterns of linguistic errors by gender and school type remain rare. Addressing these gaps, this study applies a validated error-coding rubric with established inter-rater reliability to analyse linguistic errors among ninth-grade students in Uttar Pradesh, providing reliable evidence to inform English language pedagogy and evaluation.

OBJECTIVES

- To quantify the overall error rate and error density in English written texts produced by the students after dictation.
- To classify and measure the frequencies of major error categories: orthographic, morphological, syntactic, lexical, and punctuation.
- To compare error patterns across subgroups (school type: private and government, and Gender).
- To compute inter-rater reliability (IRR) for the error-coding rubric.

HYPOTHESES

H1: In dictation-based writing of grade nine, orthographic errors are found more frequently as compared to syntactic errors.

H2: Students from private schools exhibit lower error density than students from government schools.

H01: There is no significant difference in error density between male and female students.

H3: The error-coding rubric yields substantial inter-rater agreement.

OPERATIONAL DEFINITIONS

In this study,

- Error Density refers to the following calculation: Error Density = (Total Number of Errors ÷ Total Number of Words) × 100
- Orthographic Error refers to any deviation from standard English orthographic norms (spelling and script mistakes) that results in the inaccurate representation of words in written form.
- Morphological Error refers to the incorrect use of inflectional markers (gender, number, case, or tense).
- Syntactic Error refers to mistakes in word order, word agreement, or sentence structure.
- Lexical Error refers to incorrect word choice (semantic error) or the use of an inappropriate register.
- Punctuation Error refers to incorrect or missing punctuation marks.

RESEARCH DESIGN

The present study adopted a quantitative research approach and a descriptive, cross-sectional design to systematically analyse linguistic errors in the English writing of ninth-grade students. A quantitative approach was suitable because the study aimed to statistically analyse the frequency, distribution, and density of different categories of writing errors, and to compare error patterns using statistical techniques (Creswell & Creswell, 2018). The descriptive design categorised linguistic errors without manipulating any variable, allowing an accurate picture of students' writing performance to be obtained (Cohen et al., 2018).

This study was conducted at a single point in time, making it cross-sectional in nature. Students' written responses were collected through a standardised dictation task and analysed using a multidimensional error-coding rubric covering orthographic, morphological, syntactic, lexical, and punctuation errors. Inter-rater agreement was estimated using ICC and Fleiss' Kappa before the main analyses to establish the reliability of the conclusions. Descriptive statistics were used to summarise the frequency and distribution of linguistic errors, while the Mann-Whitney U test was applied to examine differences in error density by gender and school type.

Overall, the research design provides a rigorous and reliable framework for profiling linguistic errors quantitatively, aligned with the study's objectives within the context of ESL in Indian secondary schools.

Population, Sampling, and Participants

The target population for this study comprised ninth-grade students from government and private secondary schools in Bhadohi, Uttar Pradesh. Stratified random sampling was used to collect the data. Schools were first categorised as government or private; based on administrative permission, one government and one private secondary school were then selected, and students were selected at random from the ninth grade in each school.

The final sample comprised a total of 185 students, both male and female, with 109 from the government school and 75 from the private school. Ninth-grade students were selected because, at this level, they have accumulated substantial exposure to formal English language teaching, while their writing skills remain in a formative phase; this class was therefore considered appropriate for studying linguistic errors in English writing.

The sample size was suitable and adequate for non-parametric group comparisons using the Mann-Whitney U test, as well as for descriptive statistical analysis (Cohen et al., 2018; Creswell & Creswell, 2018).

Inclusion and Exclusion: Students who were present and provided consent to participate were included and asked to write the dictated passage. Students who were absent on the day of testing, or who declined consent, were excluded from the study.

Research Instruments

Data were collected using a standardised English dictation task and a validated multidimensional error-coding rubric. The dictation was constructed so that spelling, morphology, syntax, vocabulary, punctuation, and other linguistic features could all be assessed, and all participants received the same linguistic input. Compared with open-ended writing tasks, standardised dictation minimises variation arising from differences in content selection, writing style, and subject knowledge (Felker et al., 2019; Lu et al., 2020).

Students' written responses were analysed using a structured error-coding rubric adapted from established error-analysis frameworks (Dobrić & Sigott, 2014), which classifies errors into five major

categories: orthographic, morphological, syntactic, lexical, and punctuation.

Before data analysis, operational definitions and coding guidelines were developed for each error category to ensure consistency and minimise subjective judgement during error identification. Two independent raters coded the students' responses, and the reliability of the rating rubric was estimated using Fleiss' Kappa and the Intraclass Correlation Coefficient (ICC). Both statistics indicated satisfactory inter-rater agreement (see Tables 1 and 2), reinforcing the objectivity and reliability of the linguistic error analysis.

Although a dictation task engages listening, spelling, written transcription, and several other aspects of language, it was selected because the primary objective of this study was not to evaluate the broader aspects of writing, but to obtain a standardised and comparable linguistic sample for quantitative error profiling. The interpretation of the study's conclusions should therefore be understood within this methodological context. Future studies could combine authentic writing tasks with dictation to obtain a more comprehensive evaluation of writing proficiency.

Table 1. Error-coding rubric

Category	Code	Description	Example	Research Relevance
Orthographic (OR)	OR1	Spelling error	Wrong → Rong	Shows writing accuracy gaps
	OR2	Diacritic error	Discipline → Discepline	Indicates orthographic awareness
Morphological (MO)	MO1	Gender/number error	Boy … Beautiful	Reflects grammatical issues
	MO2	Tense/aspect error	Wrong auxiliary usage	Shows temporal/aspect difficulty
Syntactic (SY)	SY1	Word order error	Misplaced sentence parts	L1 interference, sentence structure
	SY2	Subject–verb agreement	Subject–verb mismatch	Shows concord errors
	SY3	Fragment/run-on	Overextended sentence	Indicates cohesion issues
Lexical/ Semantic (LX)	LX1	Wrong word choice	Inaccurate vocabulary	Limited lexical depth
	LX2	Collocation/register	Misuse of expression	Shows pragmatic errors
Punctuation/ Capitalization (PU)	PU1	Missing/misused punctuation	Incorrect pause	Weak discourse conventions
	PU2	Quotation/dash misuse	Incorrect punctuation	Stylistic/orthographic awareness

Reliability and Validity

The results of the inter-rater reliability analysis for each error category indicate a generally high level of agreement and reliability among raters. Orthographic (OR) errors showed a Fleiss' Kappa (K) of 0.642, reflecting substantial agreement among raters, and an ICC of 0.823, indicating good reliability. Morphological (MO) errors had a K of 0.587, corresponding to moderate agreement, and an ICC of 0.789, also reflecting good reliability. Syntactic (SY) errors yielded a K of 0.534 (moderate agreement) and an ICC of 0.756 (good reliability), suggesting acceptable consistency in coding. Lexical/Semantic (LX) errors recorded a K of 0.612 (substantial agreement) and an ICC of 0.812 (good reliability). Punctuation (PU) errors achieved the highest agreement, with a K of 0.678 (substantial agreement) and an ICC of 0.845, indicating very good reliability. Overall, the findings demonstrate that the error-coding rubric is reliable, with most categories exhibiting substantial inter-rater agreement and good intraclass consistency, supporting the trustworthiness of the data-coding process in this study.

Table 2. Fleiss Kappa result (Categorical Agreement)

Code	K=	Agreement	ICC	Reliability
OR	0.642	Substantial agreement	0.823	Good Reliability
MO	0.587	Moderate agreement	0.789	Good Reliability
SY	0.534	Moderate agreement	0.756	Good Reliability
LX	0.612	Substantial agreement	0.812	Good Reliability
PU	0.678	Substantial agreement	0.845	Good Reliability

The administration of the dictation task was standardised across several parameters, including narration style, dictation speed, participant characteristics (9th-grade students), and testing environment. This uniformity helped maintain internal validity by controlling potential extraneous variables that could affect student performance. The error-coding rubric was developed based on well-defined linguistic categories and refined through a pilot test, thereby ensuring construct validity of the instrument. Furthermore, the rubric and the dictation passage were reviewed by three English language experts, which strengthened the content validity of both the rubric and the task materials.

Procedure of Data Collection

Data collection took place during regular school hours after obtaining the necessary permissions. Standardised English dictation was administered in comparable classroom settings in both schools to ensure consistency. Students were informed of the objective and guidelines of the writing task before the dictation began.

The dictation passage was read at a controlled pace to give students adequate time to transcribe it accurately. All participants received uniform language input, minimising variation arising from content selection, writing style, and subject knowledge. Response sheets were collected immediately after completion and coded anonymously to maintain confidentiality.

Students' response sheets were evaluated by two independent raters using the validated multidimensional error-coding rubric. Disagreements were resolved through discussion based on predetermined error-coding guidelines, after which the final coded dataset was prepared for statistical analysis.

DATA ANALYSES

Data were analysed using descriptive and inferential statistics. First, the linguistic errors in students' writing were classified into five categories — orthographic, morphological, syntactic, lexical, and punctuation. The overall distribution of linguistic errors was then calculated in terms of frequency, percentage, and error density for each category.

The reliability of the error-rating rubric was tested using Fleiss' Kappa and ICC to assess the consistency of error coding between the two raters. The resulting scores indicated satisfactory inter-rater agreement, confirming the reliability of the coding process.

Before conducting inferential statistical analyses, the distribution of error density was tested for normality. As the data did not meet the assumptions of a normal distribution, the Mann-Whitney U test was applied to compare error density by gender (male and female) and school type (government and private). All statistical analyses were performed using SPSS, with statistical significance set at $p < .05$.

Objective 1

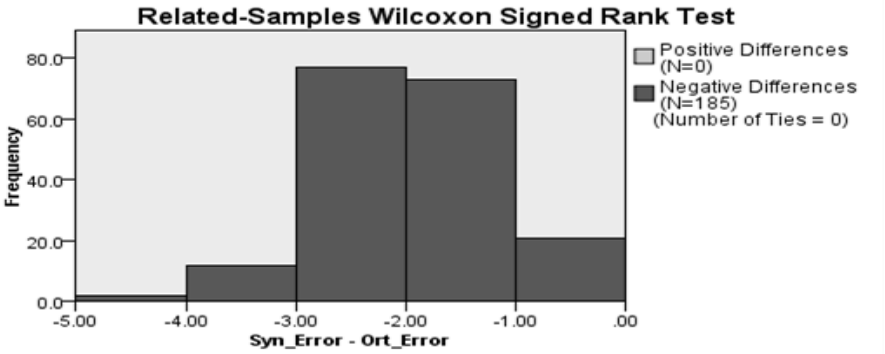
The first objective of this study was to quantify the overall error rate and error density in the English written texts produced by students after dictation.

The descriptive statistics (see Table 3) reveal that although students produced a consistent volume of text (mean total words = 130.53, SD = 13.16), they demonstrated a profoundly high and troubling level of inaccuracy in their written dictation, with an average error rate of 40.29%, meaning that approximately 40 out of every 100 words contained an error. This indicates a fundamental and widespread struggle with the foundational skills of spelling, grammar, or auditory processing required for the task. Furthermore, the exceptionally high standard deviation (30.06) and range (111) for total errors, substantially larger than the variability in word count, indicate that this poor average performance masks extreme internal disparity: the group is not uniformly weak, but instead spans a wide range of abilities, from a few relatively proficient writers to a significant number of students producing text with an overwhelming density of errors. The data therefore point to a dual challenge: a systemic issue with written accuracy across the entire cohort, compounded by a severe inequality in linguistic competence that calls for differentiated instruction rather than a uniform remedial approach.

Table 3. Error rate

	N	Range	Mean	Std. Deviation
Total Words	185	81.00	130.5297	13.16494
Total Error	185	111.00	52.5892	30.05672
Valid N	185			

An alternative hypothesis (H1) was formulated stating that, in dictation-based writing among grade-nine students, orthographic errors occur more frequently than syntactic errors. Tests of normality for the orthographic and syntactic error data indicated non-normal distributions. Consequently, a Wilcoxon Signed-Rank test was conducted to compare the frequency of orthographic and syntactic errors in students' dictation-based writing. The results showed a statistically significant difference between the two error types, Z = 716.94, p < .001, with orthographic errors occurring more frequently than syntactic errors, supporting the hypothesis (see Figure 1).



Total N	185
Test Statistic	.000
Standard Error	716.994
Standardized Test Statistic	-11.998
Asymptotic Sig. (2-sided test)	.000

Figure 1 Wilcoxon sample rank test

Objective 2

The second objective of this study was to classify and measure the frequencies of the major error categories (see Table 4). The descriptive statistics provide a quantitative overview of the frequency and distribution of six major error categories — Orthographic (OR), Morphological (MO), Syntactic (SY), Lexical (LX), Punctuation (PU), and Syntactic-Morphological (SYM) — in the written texts produced by the sample (N = 185). The analysis reveals a clear hierarchy of error prevalence and significant variability in student performance. The data indicate that orthographic errors (M = 14.17, SD = 7.99) are the most frequent category, constituting the primary challenge for this cohort. This is followed closely by syntactic (M = 11.71, SD = 7.77) and morphological errors (M = 11.26, SD = 6.83), which also exhibit substantial means. The high standard deviations for these three categories — each nearly half the value of their respective means — point to considerable heterogeneity within the sample, suggesting that while these areas are problematic on average, there is a wide disparity in mastery, with some students making very few errors and others struggling significantly. The large ranges (OR range = 30, SY range = 27, MO range = 24) further corroborate this pronounced variability, indicating that the gap between the lowest- and highest-performing students in these linguistic domains is substantial (see Table 4).

Table 4. Descriptive Statistics for error categories

Categories	Range	Mean	Std. Deviation
OR Error	30.00	14.1730	7.99132
MO Error	24.00	11.2595	6.83404
SY Error	27.00	11.7081	7.77127
LX Error	19.00	10.0757	5.19874
PU Error	9.00	3.4865	1.73865
SYM Error	7.00	1.8865	.97973
Valid N (listwise)	185		

In contrast, lexical errors (M = 10.08, SD = 5.20) show a moderately high mean but a comparatively lower standard deviation and range, implying a more consistent, yet still problematic, level of difficulty with word choice across the population. The categories of punctuation (M = 3.49, SD = 1.74) and syntactic-morphological errors (M = 1.89, SD = 0.98) are the least frequent error types. Their low means, small standard deviations, and restricted ranges suggest that these are areas of relative strength for the cohort, with most students demonstrating a more uniform and competent grasp of these mechanics (see Table 4).

In conclusion, the findings delineate a clear profile of written-accuracy challenges. The predominant difficulties lie in foundational aspects of written form (orthography) and sentence-level grammar (syntax and morphology), with significant individual differences. Lexical errors present a secondary, more consistent concern, while punctuation and combined syntactic-morphological elements are the least problematic. This hierarchy is critical for directing pedagogical intervention, suggesting that instructional resources should be prioritised toward spelling, grammatical agreement, and sentence structure before addressing finer points of lexical precision or punctuation.

Objective 3

The third objective of this study was to compare error patterns across subgroups (school type – private and government – and gender). Under this objective, an alternative hypothesis (H2) was formulated stating that students from private schools exhibit lower error density than students from government schools. The Mann–Whitney U test, a non-parametric alternative to the independent-samples t-test, was employed because the error data violated the assumption of normality, indicating that the distribution of total errors within groups did not meet parametric requirements.

Table 5. Mann-Whitney U test: Grouping variable: school types

Ranks					**Test Statistics**	
	School Type	**N**	**Mean Rank**	**Sum of Ranks**		**Total Error**
Total Error	Private	76	89.43	6796.50	Mann-Whitney U	3870.500
	Govt.	109	95.49	10408.50	Wilcoxon W	6796.500
	Total	185			Z	-.758
					Asymp. Sig. (2-tailed)	.449

The analysis yielded a significance value of p = .449, well above the conventional threshold of .05. This implies that the probability of the observed difference in error distributions arising purely by chance is approximately 44.9% (see Table 5). Consequently, the null hypothesis of no group difference cannot be rejected. The findings therefore provide no statistically significant evidence that students from private and government schools differ in their total error rates on dictated texts. The directional hypothesis, which posited lower error density among private-school students, is not supported by the data. Although the negative Z-score suggests a minimal tendency in the expected direction, the effect is insufficiently consistent or robust to reach statistical significance. Hence, the variation in error counts between school types can be regarded as negligible.

A null hypothesis (H01) was also formulated under Objective 3, stating that there is no significant difference in error density between male and female students. Because the data violated the assumption of normal distribution, the non-parametric Mann–Whitney U test was used to examine whether male and female students differed significantly in their total error counts. The analysis yielded U = 4092.50, Z = –0.469, p = .639 (two-tailed). As this significance value is well above the conventional alpha level of .05, the result indicates no statistically significant difference in error density between male and female students (see Table 6).

Table 6. Mann-Whitney U test: Grouping variable: gender type

Ranks					**Test Statisticsa**	
	Sex	**N**	**Mean Rank**	**Sum of Ranks**		**Total Error**
Total Error	Male	87	94.96	8261.50	Mann-Whitney U	4092.500
	Female	98	91.26	8943.50	Wilcoxon W	8943.500
	Total	185			Z	-.469
					Asymp. Sig. (2-tailed)	.639

Thus, the null hypothesis of no difference between groups cannot be rejected. The direction of the Z-score suggests a very slight tendency favouring one group, but the effect is negligible and not systematic. In sum, the data provide no evidence that sex influences the number of errors produced in the dictated texts, indicating that the observed variation is likely due to chance rather than a systematic gender effect.

Consistent with prior research, the findings suggest that error production is shaped more by instructional quality, learning strategies, and contextual factors than by biological sex. Pedagogically, this implies that error-reduction interventions should be applied uniformly across genders, focusing on grammar instruction, syntactic practice, and individualised feedback. Overall, sex does not appear to be a key determinant of writing accuracy; instead, learner-specific and educational factors play a more critical role. Future research with larger and more diverse samples may further clarify these influences.

Objective 4

The fourth objective of this study was to compute inter-rater reliability for the error-coding rubric. Under this objective, a hypothesis (H3) was formulated stating that the error-coding rubric yields substantial inter-rater agreement. This was assessed using Fleiss' Kappa and the Intraclass Correlation Coefficient (ICC). Results demonstrated strong agreement across all error categories. Orthographic (k = 0.642, ICC = 0.823) and punctuation errors (k = 0.678, ICC = 0.845) showed substantial agreement and high reliability. Morphological (k = 0.587, ICC = 0.789), syntactic (k = 0.534, ICC = 0.756), and lexical/semantic errors (k = 0.612, ICC = 0.812) also exhibited moderate-to-substantial agreement with good reliability (see Table 2). These findings, with all ICC values exceeding the 0.75 threshold for good reliability (Koo & Li, 2016), confirm the study's fourth hypothesis (H3) and validate the rubric as a robust and consistently applicable instrument for error classification.

CONCLUSION

This study quantitatively analysed the linguistic errors of ninth-grade students using a multidimensional error-coding framework. The findings revealed that orthographic errors were the most prevalent, followed by syntactic and morphological errors, indicating that students have not yet achieved adequate competency in basic English writing and continue to face significant difficulties in these areas. The absence of statistically significant differences by gender (male and female) or school type (government and private) suggests that students' writing difficulties are more closely related to shared teaching and learning experiences than to demographic characteristics. These conclusions are consistent with prior studies, which have highlighted orthographic and grammatical errors as persistent challenges for students learning English as a second language (Lasaten, 2014; Rahman et al., 2017; Isma et al., 2023).

Methodologically, this study demonstrates that the combined use of a standardised dictation task, a validated error-coding rubric, and satisfactory inter-rater reliability provides a systematic, objective, and reliable method for identifying and profiling linguistic errors. The study's conclusions enrich the research literature on error analysis in the context of ESL in Indian secondary schools, and provide empirical evidence to inform English writing pedagogy, diagnostic assessment, and teacher education.

Educational Implications

The findings of this study carry several important implications for English language pedagogy and learning.

- First, the abundance of orthographic, syntactic, and morphological errors indicates a need for greater focus on orthography, grammar, sentence construction, and editing skills within the secondary-level English curriculum. Writing pedagogy should integrate language-focused activities so that students' linguistic accuracy and overall writing ability can be developed together.
- Second, teachers should adopt diagnostic error analysis as a classroom assessment practice, enabling them to provide corrective feedback on recurring linguistic errors. Constructive writing activities, peer review, and revision-based practices may further improve students' linguistic accuracy and writing fluency.
- Third, linguistic error analysis, diagnostic assessment, and evidence-based feedback should be incorporated into both pre-service and in-service teacher education programmes. Such professional development will better equip teachers to analyse students' writing, identify their learning needs, and design remedial teaching.
- Fourth, the validated error-coding rubric developed in this study offers a practical and reliable framework for continuous, formative assessment. Schools may use such diagnostic tools to monitor students' writing progress, evaluate teaching effectiveness, and identify

students who require additional academic support.

Limitation of the Study

This study has some limitations. First, it was conducted in only two secondary schools in a single district of Uttar Pradesh, which may limit the generalisability of the conclusions. Second, the analysis of linguistic errors was based on standardised dictation, which also draws on listening, orthographic, and transcription skills; the conclusions should therefore be understood as reflecting linguistic accuracy within this specific task. Third, because the study used a cross-sectional design, students' writing development could not be evaluated over time.

Recommendation for Future Studies

Future studies should include larger and more diverse samples drawn from different academic environments and regions. For a more comprehensive evaluation, essays, narratives, and argumentative writing should be used alongside standardised dictation. Furthermore, longitudinal and mixed-methods studies should be conducted, as these could better illuminate the relationship between linguistic errors and classroom teaching practices.

REFERENCES

Almineeai, A. (2025, January). Understanding syntactic deviations of sentence derivation in student English: A study within the minimalist program. Forum for Linguistic Studies, 7(1), 960–974.

Bora, P. (2023). Importance of writing skills to develop students' communication skills. Journal for Research Scholars and Professionals of English Language Teaching, 7(35), 2–6. https://doi.org/10.54850/jrspelt.7.35.009

Bushi, J., & Meta, A. (2026). Error correction and its role in foreign language learning. European Journal of Contemporary Education and E-Learning, 4(2). https://doi.org/10.59324/ejceel.2026.4(2).06

Ci-hui, L. (2003). Error analysis and reflections on its application to foreign language teaching. https://doi.org/10.3969/j.issn.1674-2346.2003.02.016

Dobrić, N., & Sigott, G. (2014). Towards an error taxonomy for student writing. Zeitschrift für Interkulturellen Fremdsprachenunterricht, 19(2), 111–118.

Engka, R., Aderlaepe, & Miliha, L. (n.d.). A study of grammatical errors in students' writing explanation text of eleventh-grade students at Senior High School 10 Kendari. Journal of Teaching of English, 9(4).

Felker, E. R., Ernestus, M., & Broersma, M. (2019). Evaluating dictation task measures for the study of speech perception.

Fraca de Barrera, L. L. (2007). El procesamiento psicolingüístico del dictado y sus implicaciones para la enseñanza de la lengua escrita. Revista de Pedagogía, 28(82).

Government of India. (2011). Census of India 2011. Government of India.

Haded, M. S. (1998). The merits of exploiting error analysis in foreign language teaching and learning. RELC Journal, 29(1), 55–65. https://doi.org/10.1177/003368829802900104

Hadke, S. K. (2021). Hindi language: Its contemporary importance and aspects. International Journal of Academic Research and Development, 6(4), 69–74.

Hafizhoh, P. N., Riznanda, W. A., & Amalia, H. (2025). Students' use of simple present tense in their descriptive writing: An error analysis. Journal of English Language Learning, 9(1), 878–898.

Hamed, M. (2018). Common linguistic errors among non-English major Libyan students' writing. Arab World English Journal, 9(3), 219–232. https://doi.org/10.24093/awej/vol9no3.15

Hamuddin, B. (2015). Bringing EA into ELT's classroom: Rising students' awareness to avoid errors. ELT Lectura, 2(2). https://doi.org/10.31849/ELT-LECTURA.V2I2.464

Hassan, M. K., & Rami, Z. (2024). Exploring nuances in second language acquisition: An error analysis perspective. Migration Letters. https://doi.org/10.59670/ml.v21i4.7319

Heydari, P., & Bagheri, M. S. (2012). Error analysis: Sources of L2 learners' errors. Theory and Practice in Language Studies, 2(8). https://doi.org/10.4304/TPLS.2.8.1583-1589

Isma, A., Rasmin, L. O., & Samsudin. (2023). Decoding the challenges: A study of English writing errors among EFL students. GLENS: Global English Insights Journal, 1(1), 1–9.

Jabeen, A., Kazemian, B., & Mustafai, M. S. (2015). The role of error analysis in teaching and learning of second and foreign language. Education and Linguistic Research, 1(2), 52–57. https://doi.org/10.5296/elr.v1i1.8189

Khalifa, M. F. (2018). Contrastive analysis, error analysis, markedness theory, universal grammar and monitor theory and their contributions to second language learning. International Journal of Linguistics, 10(1), 12–45. https://doi.org/10.5296/ijl.v10i1.12479

Klimova, B. F. (2013). The importance of writing. Paripex – Indian Journal of Research, 2(1), 9–11.

Koo, T. K., & Li, M. Y. (2016). A guideline for selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lasaten, R. S. (2014). Analysis of errors in the English writing of teacher education students. Researchers World, 5(4).

Lee, S., & Schmidgall, J. (2020). The importance of English writing skills in the international workplace (Research Memorandum).

Lennon, P. (2008). Contrastive analysis, error analysis, interlanguage. In Introduction to Applied Linguistics (pp. 51–60).

Li, X., & Kane, F. (2016). Error analysis and CFL learning in Northern Ireland: Error processing by beginners. In Sinolingua (pp. 244–259).

Li, Z., Huang, Q., & Liu, J. (2024). ChatGPT analysis of strengths and weaknesses in English writing and their implications. Applied Mathematics and Nonlinear

Sciences, 9(1), 1–15. https://doi.org/10.2478/amns-2024-0912

Lu, Y., Gales, M. J., & Wang, Y. (2020). Spoken language grammatical error correction. In Proceedings of Interspeech 2020. https://doi.org/10.21437/INTERSPEECH.2020-1852

Mahdiyah, Z., Lunawati, E. S., & Monika, R. S. (2023). Peran analisis kesalahan berbahasa dalam meningkatkan kemampuan menulis. Jurnal Yudistira, 2(1), 52–58. https://doi.org/10.61132/yudistira.v2i1.380

McDowell, L. (2020). Error analysis: A methodological exploration and application. JALT Postconference Publication, 19(1). https://doi.org/10.37546/JALTPCP2019-53

Nadia. (2024). The analysis of students' grammatical errors in recount text writing. Journal of English Language Teaching and Linguistics, 9(3), 257–267. https://doi.org/10.21462/jeltl.v9i3.1424

Nainggolan, D., Simarmata, R. O., & Napitupulu, F. D. (2025). An error analysis of inflectional affixes in writing analytical exposition text on students at SMAN 3 Medan. Action Research Journal Indonesia, 7(2), 875–892.

Phettongkam, H. (2013). Error analysis and its implications in communicative English language teaching. Thammasat Review, 16(3), 96–108.

Pristha, J. M., Jamil, S., & Fatahillah, W. (2025). Grammatical error analysis in writing narrative texts: The case of Grade X students of SMA Negeri Gunung Meriah, Deli Serdang. Tuwah Pande: Jurnal Ilmu Pendidikan dan Pengajaran, 4(1), 149–160.

Qing, Y. (2025). Morphological error analysis in junior high school English writing. Advances in Education, 15(3), 595–599. https://doi.org/10.12677/ae.2025.153444

Rahman, M. A., Sainu, M. N., & Asfah, I. (2017). Revisiting the linguistic errors in the writing of the students majoring in English. Advances in Social Science, Education and Humanities Research, 149, 158–160.

Saragi, C. N., Silalahi, D. P., & Saragi, D. M. (2025). Surface-level errors in EFL writing: A taxonomic study of junior high school students' descriptive texts in Grade IX SMP Negeri 23 Medan. Jurnal Basataka, 8(1), 241–246.

Sarkar, S. (2025). The role of error analysis in teaching and learning of English as a second language: A case study in Jawahar Navodaya Vidyalaya.

Waldi, M. R., Zulandha Melani, B. B., Lail, H., & Nurtaat, L. (2024). An analysis of conjunction errors in recount text: A study at the 10th grade of SMAN 1 Gerung. Pendas: Jurnal Ilmiah Pendidikan Dasar, 9(4).

Wolf, C. (2022). Error analysis. In The Routledge handbook of translation and methodology. Routledge. https://doi.org/10.4324/9781003034704-20