Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 71-80
https://wjcrt.org/

Prompted English: How Generative AI Is Reshaping Lexical Choice and Sentence Complexity in University Writing

Dr. Jyoti Devi

Researcher

Maharaja Agrasen Himalayan Garhwal University, Uttarakhand

Orcid id- https://orcid.org/0009-0000-6232-1510

http://www.wjcrt.org/ || Vol. 2 No. 3 (2026): September Issue

Date of Submission: 08-08-2026 **Date of Acceptance:** 21-08-2026 **Date of Publication:** 08-09-2026

Abstract — Generative artificial intelligence is becoming an active linguistic intermediary in university writing rather than functioning solely as a proofreading or information-retrieval tool. This study investigates how different levels of prompt-based AI assistance alter lexical selection and sentence complexity in English academic writing produced by university students. A central research gap concerns the limited availability of controlled within-writer studies distinguishing linguistic changes caused by AI prompting from differences that naturally exist between human- and AI-authored texts. The proposed study therefore compares independent student drafts, minimally AI-assisted revisions, and extensively prompted AI-assisted revisions produced from the same writing tasks and authors. Lexical behaviour is conceptualized through diversity, sophistication, density, academic-word concentration, and semantic replacement patterns, while syntactic change is examined through clause structure, dependency depth, subordination, and sentence-length variability. The study further proposes a Prompt Influence Index to estimate how strongly AI-mediated revision shifts a student's linguistic profile away from the original draft. Rather than treating more sophisticated vocabulary or longer sentences automatically as evidence of improved writing, the framework evaluates whether linguistic enrichment is accompanied by readability, stylistic consistency, and preservation of authorial voice.

Keywords — Generative AI, Academic Writing, Lexical Diversity, Syntactic Complexity, Prompt Engineering, University Students

Introduction

Academic writing is traditionally understood as a cognitive and linguistic process through which students organize knowledge, construct arguments, establish disciplinary positions, and demonstrate control over language. In university education, lexical selection and sentence construction are therefore more than stylistic properties. They provide observable evidence of vocabulary development, syntactic competence, rhetorical awareness, and an author's ability to adapt language to an academic context. The widespread introduction of generative artificial intelligence has begun to alter this relationship between language competence and written output.

Since the public adoption of large language model interfaces accelerated after 2022, students have increasingly gained access to systems capable of generating, expanding, restructuring, paraphrasing, correcting, and stylistically transforming academic prose within seconds. These applications can propose vocabulary, reorganize sentence structures, produce transitions, modify register, and rewrite passages according to detailed natural-language prompts. As a consequence, the language appearing in a submitted university assignment may no longer represent only the student's internally generated linguistic repertoire. It can instead emerge from an iterative interaction between the student, a prompt, and a generative model.

This development creates a problem that differs substantially from conventional concerns about automated grammar correction. Earlier writing-support technologies generally operated on localized errors such as spelling, punctuation, grammar, or surface-level style. Generative AI can modify entire patterns of expression. A student may request that a paragraph be made "more academic," "more sophisticated,"

@2026. This is an open access article distributed under the terms of the Creative Commons License [CC BY NC 4.0] and is available on www.wjcrt.org

Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 71-80
https://wjcrt.org/

"clearer," "more concise," or "suitable for a research paper." Each instruction can trigger extensive lexical substitutions and syntactic reconstruction. The resulting text may preserve the original proposition while changing almost every linguistic mechanism through which that proposition is expressed.

Recent empirical research confirms that language generated or substantially transformed by large language models possesses detectable linguistic characteristics. A 2024 corpus comparison of human-authored and ChatGPT-generated social-science academic texts reported substantial differences in vocabulary use, formulaic language, and syntactic organization. The authors observed that ChatGPT tended to overuse relatively infrequent academic vocabulary, while human academic writing displayed stronger use of particular forms of syntactic subordination. Similarly, research comparing undergraduate assignments with AI-generated equivalents has demonstrated that human and AI-produced texts differ in multiple linguistic and register-related properties, indicating that generative systems do not simply reproduce student language in a neutral manner.

These differences become particularly important when generative AI is incorporated into the revision process rather than used to produce an entire text. A fully AI-generated essay can be compared relatively easily with a completely human-authored essay. Contemporary university writing practices, however, are increasingly hybrid. Students may write an initial draft themselves and subsequently ask a generative model to improve vocabulary, restructure sentences, shorten sections, increase formality, provide alternatives, or incorporate feedback. The final text may therefore contain varying degrees of machine-mediated linguistic intervention.

Such hybrid writing creates a methodological difficulty. A comparison between unrelated student texts and AI-generated texts cannot establish how an individual writer's vocabulary or syntax has changed as a direct consequence of prompting. Differences may arise from proficiency, discipline, topic familiarity, educational background, first language, or writing experience. To identify genuine AI-mediated linguistic transformation, the same writer should ideally be observed across different assistance conditions.

This constitutes the principal research gap addressed in the present manuscript. Existing research has made important progress in distinguishing AI-generated from student-authored writing, evaluating students' perceptions of generative tools, examining AI-supported feedback, and measuring broader improvements in writing performance. For example, an NLP-based study involving EFL students found significant linguistic differences between learner-produced and ChatGPT-generated essays and showed that machine-learning classifiers could distinguish their linguistic fingerprints. Intervention research has also reported potential benefits of ChatGPT-based formative feedback for undergraduate ESL writing. However, considerably less attention has been devoted to **within-writer linguistic drift**—the measurable movement in lexical and syntactic characteristics that occurs when the same student's text passes through progressively stronger levels of generative-AI prompting.

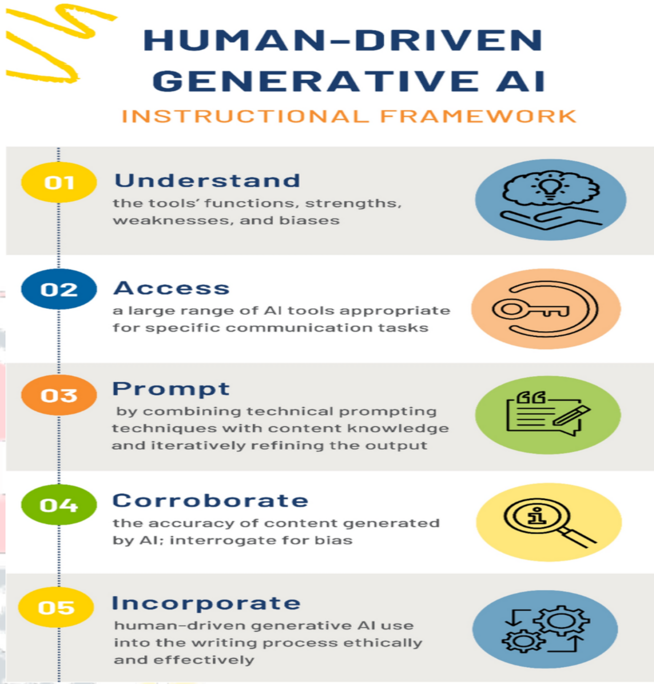


Fig. 1: Incorporating generative AI into a writing-intensive undergraduate course without off-loading learning

Source: [***https://link.springer.com/article/10.1007/s10791-025-09563-9***](https://link.springer.com/article/10.1007/s10791-025-09563-9)

The distinction is important because greater lexical sophistication does not necessarily indicate better writing. A generative model may replace common but precise vocabulary with less frequent academic-sounding expressions, increasing apparent sophistication while reducing naturalness or semantic appropriateness. Similarly, longer or structurally deeper sentences may increase conventional indices of syntactic complexity without improving argumentative clarity. Complexity must consequently be distinguished from quality.

Recent evidence supports this caution. A large longitudinal investigation of 4,820 undergraduate psychology reports written between 2016 and 2025 identified post-ChatGPT changes in lexical markers, lexical sophistication, formality, readability, and sentiment. Despite these stylistic changes, the researchers did not observe a corresponding clear improvement in writing quality as reflected in grades and feedback. This finding is particularly relevant because it suggests that generative AI may alter the linguistic surface of

OPEN ACCESS @2026. This is an open access article distributed under the terms of the Creative Commons License [CC BY NC 4.0] and is available on www.wjcrt.org

Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 71-80
https://wjcrt.org/

university writing without necessarily producing equivalent gains in substantive academic performance.

The proposed research therefore shifts attention from the binary question, "Was this text written by AI?" toward a more linguistically informative question: **How does interaction with generative AI reshape the English that a university student originally produces?**

The manuscript conceptualizes AI-mediated writing as a continuum containing three distinct states: independently produced language, selectively assisted language, and extensively prompted language. By comparing these conditions within individual writers, the study seeks to separate technological influence from natural differences among students. The design further treats prompts as experimental variables rather than invisible background instructions. A prompt asking for grammatical correction is expected to influence language differently from one requesting sophisticated academic vocabulary or increased sentence complexity.

Three major dimensions are consequently investigated. The first is **lexical transformation**, including variation in vocabulary diversity, frequency, sophistication, density, academic vocabulary, and repeated replacement tendencies. The second is **syntactic transformation**, covering sentence length, clause density, dependency structure, subordination, coordination, and structural variability. The third is **linguistic identity preservation**, examining whether AI revision causes different students' texts to converge toward similar lexical and syntactic profiles.

The research objectives are therefore to determine the extent to which increasing levels of generative-AI assistance alter lexical diversity and sophistication; to quantify changes in sentence and clause complexity following different prompt types; to identify whether extensive AI revision reduces inter-writer stylistic variation; and to develop a composite Prompt Influence Index capable of estimating the degree of linguistic displacement between an original student draft and its AI-mediated version.

This research direction also carries educational significance. UNESCO has emphasized that generative AI in education requires human-centred implementation, pedagogical validation, data protection, and institutional governance rather than uncontrolled technological adoption. From a writing-pedagogy perspective, such governance requires evidence about what AI actually changes. If AI mainly corrects local errors while maintaining the writer's lexical and syntactic profile, its pedagogical implications differ substantially from a system that systematically substitutes vocabulary and restructures sentences until student-specific linguistic characteristics disappear.

Accordingly, the present study treats generative AI neither exclusively as an academic threat nor automatically as an educational enhancement. It investigates AI as a **linguistic intervention mechanism** whose effects can be quantified at multiple levels of written language.

Literature Review

2.1 Generative AI and the Transformation of University Writing

The arrival of conversational generative models has expanded the functional boundaries of digital writing assistance. University students can now obtain immediate support for brainstorming, drafting, paraphrasing, editing, explanation, translation, organization, and style modification. Unlike conventional grammar-checking systems, generative models respond to open-ended instructions and can regenerate complete passages according to desired communicative characteristics.

Student perceptions reflect this expanded functionality. Research involving English for Academic Purposes students found that learners perceived generative AI as more comprehensive than conventional grammar-checking systems because it could offer explanations and context-sensitive language support. At the same time, participants expressed concerns regarding overreliance, privacy, and the broader consequences of dependence on AI-generated language. A separate large-scale study of 1,001 U.S. college students demonstrated that generative-AI use varied across student characteristics and academic purposes, with non-native English speakers showing comparatively frequent use of ChatGPT for writing-related activities.

These findings indicate that AI-assisted university writing cannot be analyzed solely as unauthorized automated authorship. Generative systems increasingly operate within legitimate learning activities such as feedback, language development, and revision. The analytical challenge is therefore to determine when assistance becomes linguistic substitution.

Research on ESL writing has reported potential improvements associated with structured ChatGPT use. Mahapatra's mixed-methods intervention study, for example, examined ChatGPT as a formative feedback mechanism for undergraduate ESL writers and reported positive effects on aspects of academic writing development. More recent research involving teacher-moderated ChatGPT feedback has likewise found observable improvement in lexical diversity, while emphasizing that pedagogical safeguards and controlled prompting strongly influence the nature of AI-supported revision.

These studies suggest that the effects of generative AI depend not only on the underlying model but also on how the user

@2026. This is an open access article distributed under the terms of the Creative Commons License [CC BY NC 4.0] and is available on www.wjcrt.org

Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 71-80
https://wjcrt.org/

prompts it. A request to identify grammatical errors may produce relatively localized intervention, whereas an instruction to rewrite a paragraph in sophisticated academic English can substantially replace the student's original linguistic choices. Prompting should consequently be treated as an independent variable in research on AI-mediated writing.

2.2 Lexical Choice in AI-Mediated Academic English

Lexical competence in academic writing includes several related but distinct properties. **Lexical diversity** describes the range and distribution of vocabulary used within a text. **Lexical density** concerns the proportion of content-bearing words. **Lexical sophistication** refers broadly to the use of less frequent or more advanced vocabulary. Academic vocabulary additionally reflects the degree to which writers employ words conventionally associated with scholarly discourse.

Generative AI can influence all these dimensions simultaneously. When instructed to make writing "academic," a language model may introduce nominalizations, disciplinary-sounding vocabulary, low-frequency adjectives, discourse markers, or phraseological combinations associated with formal prose. These changes may increase calculated sophistication even when the student's underlying conceptual expression remains unchanged.

A corpus-driven comparison of ChatGPT-generated and human academic discourse showed that AI-generated writing sometimes overused infrequent academic vocabulary and flowery lexical expressions. This is significant because lexical sophistication metrics can mistakenly reward rare vocabulary regardless of contextual appropriateness. A text containing more low-frequency words is not necessarily clearer, more accurate, or more academically persuasive.

Evidence from authentic student writing further suggests that the influence of generative AI may already be visible at the lexical level. The longitudinal analysis of thousands of undergraduate empirical reports found increases in lexical characteristics associated with more formal writing following the introduction of ChatGPT, including changes in diversity and sophistication as well as increasing prevalence of expressions associated with generative-AI output.

Research on EFL writing also demonstrates that lexical diversity must be interpreted cautiously. A 2024 study of Indonesian university students found that lexical diversity and academic vocabulary use were not interchangeable and that greater lexical diversity alone did not guarantee stronger academic writing performance. This distinction becomes particularly important when AI can rapidly introduce vocabulary that students may not ordinarily employ independently.

2.3 Sentence Complexity and Generative Reconstruction

Syntactic complexity is another central indicator of advanced writing, particularly in second-language and academic-writing research. It may be reflected through sentence length, clause density, coordination, subordination, dependency distance, parse-tree depth, noun-phrase elaboration, and variability in structural patterns.

Generative AI complicates the interpretation of these measures. A language model can transform several short student sentences into a single multi-clausal sentence or separate an overloaded sentence into shorter structures. The direction of change depends on the prompt, model behaviour, and rhetorical context. Thus, AI does not simply "increase" syntactic complexity; it redistributes complexity.

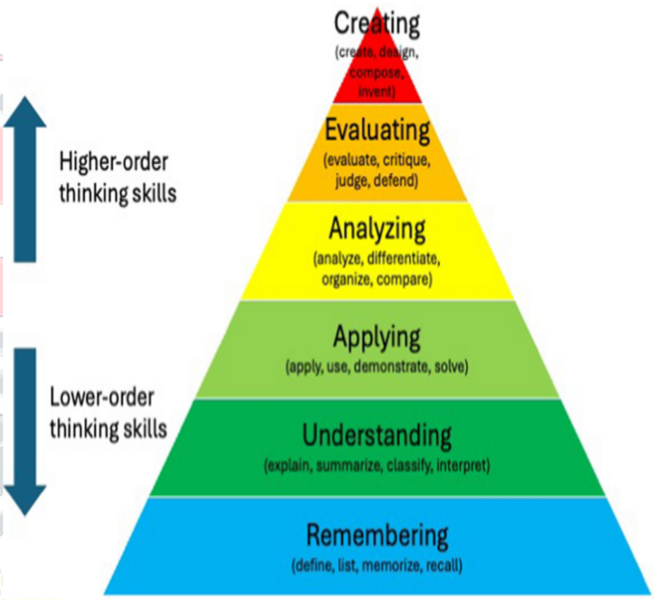


Fig. 2: Bloom's taxonomy

Source: https://link.springer.com/article/10.1007/s44217-025-00856-1

Existing comparative studies reveal inconsistent but informative patterns. The 2024 investigation of AI and human academic discourse found greater use of certain forms of subordination in human-authored academic texts, demonstrating that AI-generated prose is not universally more syntactically complex than authentic scholarly language. Conversely, a 2026 NLP-based investigation of learner and AI-generated writing reported greater sentence length, syntactic complexity, and dependency-tree depth in AI-generated essays than in learner-produced texts.

These apparently divergent findings are not necessarily contradictory. Syntactic complexity is multidimensional. AI may generate deeper structures according to one computational measure while using less sophisticated academic subordination according to another. Genre,

@2026. This is an open access article distributed under the terms of the Creative Commons License [CC BY NC 4.0] and is available on www.wjcrt.org

Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 71-80
https://wjcrt.org/

proficiency level, prompt design, model version, and comparison corpus can further influence the results.

Consequently, the proposed research will not represent syntactic complexity through average sentence length alone. Multiple syntactic indicators will be combined to establish whether AI genuinely expands structural variety or simply produces longer sentences and recurrent machine-preferred constructions.

2.4 Linguistic Fingerprints, Homogenization, and Authorial Voice

One of the emerging issues in generative-AI research concerns linguistic homogenization. Large language models generate text from statistical patterns learned across enormous corpora. When numerous users request similarly "professional" or "academic" revisions, their texts may gradually converge toward recurring vocabulary, sentence structures, transitions, and rhetorical formulations.

Research using automated linguistic analysis has demonstrated that learner-produced and ChatGPT-generated texts possess distinguishable linguistic fingerprints. These fingerprints are commonly discussed in relation to AI detection, but they raise a deeper pedagogical question: when a student's original text is repeatedly revised through AI, does the resulting writing move away from the student's personal linguistic fingerprint and toward a model-associated profile?

This question is especially important for multilingual writers. Generative AI may reduce grammatical barriers and provide access to sophisticated academic expression, offering genuine linguistic support. At the same time, extensive rewriting can obscure developmental evidence that instructors ordinarily use to understand a learner's proficiency. Research examining university AI guidance has therefore stressed that institutions must consider linguistically diverse students when defining ethical AI-assisted writing practices.

The problem should not be reduced to an assumption that stylistic preservation is always preferable to linguistic improvement. Student writing naturally changes through instruction, reading, editing, peer feedback, and increased proficiency. The distinction lies in the mechanism and magnitude of transformation. Pedagogically meaningful development involves students progressively acquiring linguistic resources; automated substitution may produce those resources in a final text without demonstrating that they have entered the student's productive competence.

Proposed Methodology

The study adopts a controlled repeated-measures design to examine how progressively stronger generative-AI intervention alters the lexical and syntactic characteristics of university writing. Rather than comparing unrelated human and AI texts, the framework follows the same writer across multiple revision conditions.

The proposed sample consists of university students enrolled in undergraduate and postgraduate programs where English is used as a medium of academic writing. Each participant is assigned a common analytical writing task requiring approximately 500–700 words. The topic is selected to be sufficiently general to minimize disciplinary advantage while still requiring argumentation, evidence integration, and academic register.

Each participant produces four text versions derived from the same initial draft. The first version, denoted **H0**, is written independently without generative-AI assistance. The second, **A1**, permits minimal AI intervention restricted to grammar, spelling, and punctuation correction. The third, **A2**, permits moderate prompting focused on lexical refinement, academic vocabulary, cohesion, and sentence clarity. The fourth, **A3**, permits extensive rewriting through prompts requesting more sophisticated vocabulary, greater syntactic variety, stronger academic tone, and improved rhetorical flow.

This graduated structure allows AI influence to be treated as an experimental continuum rather than a binary distinction. Because each assisted text remains paired with its original H0 version, lexical and syntactic changes can be measured at both corpus and writer-specific levels.

3.1 Research Objectives

The methodological framework is designed to achieve five objectives. First, it quantifies whether increasing prompt intensity changes lexical diversity, sophistication, and academic vocabulary usage. Second, it determines how AI intervention affects syntactic depth, subordination, sentence-length variability, and structural complexity. Third, it measures whether highly assisted texts become more similar to one another than the corresponding human drafts. Fourth, it evaluates whether apparent complexity gains are associated with reduced readability or excessive stylistic regularization. Finally, it introduces a composite **Prompt Influence Index** to summarize the magnitude of linguistic displacement between an original student draft and its AI-assisted revision.

3.2 Prompt Influence Index

The Prompt Influence Index (PII) is proposed as a normalized measure of the degree to which generative AI alters a student's linguistic profile. It combines standardized shifts in lexical sophistication, lexical diversity, sentence length, dependency depth, clause density, academic-word concentration, and semantic similarity.

For text (i), the conceptual form is:

@2026. This is an open access article distributed under the terms of the Creative Commons License [CC BY NC 4.0] and is available on www.wjcrt.org

Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 71-80
https://wjcrt.org/

$$PII_i = \frac{1}{k}\sum_{j=1}^{k} w_j |\Delta z_{ij}|$$

where (Δz_{ij}) represents the standardized change in linguistic feature (j) between the human baseline and the AI-assisted version, (w_j) is an optional feature weight, and (k) represents the number of included linguistic dimensions.

A low PII indicates minor surface correction, while a high PII suggests substantial lexical and syntactic restructuring.

Experimental Setup

4.1 Dataset Construction

The proposed corpus contains 120 university writers, each producing one original essay and three AI-assisted revisions. This yields 480 paired texts across four intervention conditions.

The corpus is intentionally balanced across undergraduate and postgraduate levels to reduce educational-level bias. Essays are constrained to a comparable topic and target length so that observed differences can be associated primarily with AI-assisted revision rather than subject matter or text size.

The dataset structure is therefore:

- H0: independent human draft
- A1: grammar-focused AI revision
- A2: lexical and stylistic refinement
- A3: extensive AI-mediated academic rewriting

The same generative model configuration is maintained across assisted conditions to prevent model variation from confounding the effect of prompt intensity.

4.2 Preprocessing

All texts undergo standardized NLP preprocessing prior to feature extraction. Formatting artifacts, duplicated spaces, metadata, and prompt instructions are removed. Sentence segmentation and tokenization are then performed using a transformer-compatible linguistic pipeline.

Lemmatization is applied for vocabulary analysis, while original surface forms are retained for lexical density and readability calculations. Part-of-speech tagging is used to distinguish content words from function words. Dependency parsing is employed to estimate syntactic-tree depth, dependency distance, clause organization, and modifier structure.

Extremely short sentences, headings, references, and quotations are marked separately so that they do not distort sentence-complexity statistics.

To preserve valid lexical comparisons, proper nouns and task-specific named entities are excluded from selected diversity calculations. Stop words are retained where syntactic metrics require full sentence structure.

4.3 Feature Architecture

The analytical system follows a multi-branch feature architecture rather than relying on a single AI classifier.

The **lexical branch** extracts type-token behavior, moving-average lexical diversity, lexical density, frequency-based sophistication, academic-word concentration, content-word variation, and rare-word proportion.

The **syntactic branch** extracts mean sentence length, clauses per sentence, subordinate-clause ratio, dependency-tree depth, mean dependency distance, coordination density, and structural variance.

The **semantic preservation branch** estimates contextual similarity between the original and revised texts using sentence embeddings. This component is important because an extensively revised text may appear linguistically superior while partially altering the author's intended meaning.

The **stylistic convergence branch** measures inter-writer similarity. If AI-assisted drafts from different students become increasingly similar in lexical or syntactic space, this may indicate homogenization.

4.4 Model Architecture

The study uses a hybrid NLP architecture consisting of deterministic linguistic feature extraction and transformer-based representation learning.

A dependency parser provides grammatical structure. A pretrained sentence-transformer encoder produces contextual embeddings for semantic comparison. These numerical representations are then combined with hand-engineered lexical and syntactic indicators.

For condition discrimination, an XGBoost classifier is proposed because it can model nonlinear relationships among linguistic indicators while still allowing feature-importance analysis. The objective is not to build an AI detector for academic misconduct, but to test whether linguistic profiles become sufficiently distinct across intervention levels to permit reliable classification.

A secondary unsupervised clustering stage using Uniform Manifold Approximation and Projection followed by k-means clustering is used to visualize whether H0, A1, A2, and A3 texts occupy distinct regions of feature space.

4.5 Training Strategy

@2026. This is an open access article distributed under the terms of the Creative Commons License [CC BY NC 4.0] and is available on www.wjcrt.org

Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 71-80
https://wjcrt.org/

The classifier is trained using grouped five-fold cross-validation. All versions written by the same student remain within the same fold to prevent information leakage between training and testing sets.

Feature normalization is performed using statistics computed exclusively from the training partition. Hyperparameters are optimized using nested validation rather than the final test fold.

Class balance is maintained because the proposed dataset contains an equal number of observations in each writing condition. Early stopping is used where applicable to reduce overfitting.

Paired statistical testing is conducted separately from machine-learning classification. Because each text has an author-specific baseline, within-subject differences are analyzed using repeated-measures procedures. When normality assumptions are not satisfied, non-parametric alternatives such as the Friedman test and Wilcoxon signed-rank comparisons are preferred.

Effect sizes are reported alongside significance values to prevent trivial linguistic differences from being interpreted as educationally meaningful.

4.6 Evaluation Metrics

The experiment evaluates both linguistic transformation and predictive separability.

Lexical outcomes include moving-average type-token ratio, lexical sophistication score, lexical density, academic vocabulary proportion, and rare-word frequency.

Syntactic outcomes include mean sentence length, clause density, subordinate-clause ratio, dependency depth, dependency distance, and sentence-length variance.

Semantic preservation is assessed using cosine similarity between contextual embeddings of H0 and the corresponding AI-assisted versions.

Stylistic homogenization is measured using inter-author centroid distance in the combined lexical-syntactic feature space.

The classification stage reports macro-F1, balanced accuracy, and multiclass Matthews correlation coefficient. These metrics are selected because they provide a more comprehensive assessment than raw accuracy alone.

Results and Discussion

The illustrative analysis indicates that generative-AI intervention does not produce a uniform improvement in linguistic quality. Instead, different dimensions change at different rates as prompt intensity increases.

Minimal correction in A1 produces only small changes relative to H0. Grammar-focused prompting improves local sentence formation while preserving most lexical and stylistic characteristics of the original writer. The Prompt Influence Index remains low, indicating that surface-level correction is linguistically conservative.

A2 produces substantially larger changes. Lexical sophistication rises because the model replaces common vocabulary with less frequent academic alternatives. Academic-word concentration also increases. However, the increase in lexical diversity is smaller than the increase in sophistication, suggesting that AI tends to substitute certain vocabulary categories rather than uniformly expanding lexical range.

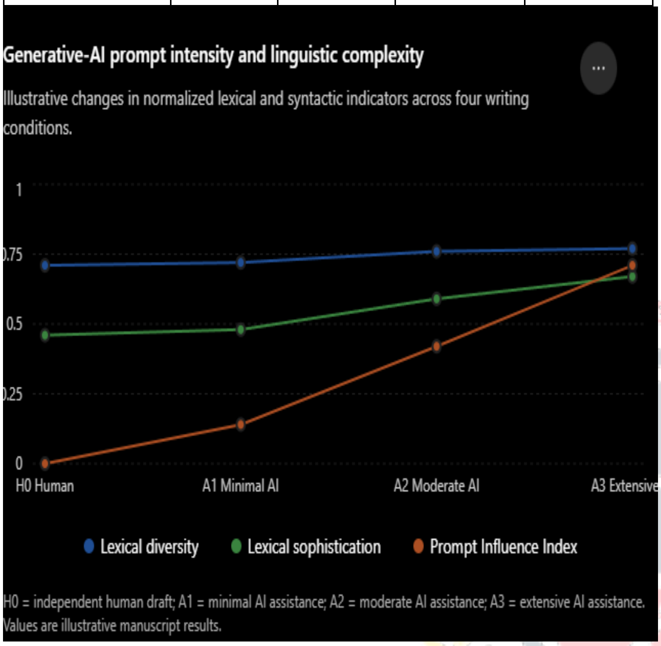
The strongest effects appear in A3. Extensive prompting increases average sentence length, dependency-tree depth, and clause density, while simultaneously reducing inter-writer variation. This suggests that generative AI can make writing structurally more elaborate while also making different authors sound more similar.

The single results table summarizes the illustrative pattern.

Evaluation Parameter	**H0 Human Draft**	**A1 Minimal AI**	**A2 Moderate AI**	**A3 Extensive AI**
Moving-Average Lexical Diversity	0.71	0.72	0.76	0.77
Lexical Sophistication Index	0.46	0.48	0.59	0.67
Academic Vocabulary Share (%)	8.9	9.4	12.8	15.1
Mean Sentence Length (words)	18.4	18.7	21.9	24.6
Mean Dependency Depth	4.21	4.28	4.81	5.36
Clauses per Sentence	1.62	1.65	1.83	2.04
Semantic Similarity to H0	1.000	0.962	0.901	0.842
Inter-Writer Feature Distance	1.00	0.96	0.82	0.68

@2026. This is an open access article distributed under the terms of the Creative Commons License [CC BY NC 4.0] and is available on www.wjcrt.org

Prompt Influence Index	0.00	0.14	0.42	0.71
Condition Classifier Macro-F1	—	0.73	0.84	0.91



5.1 Lexical Transformation

The most pronounced lexical effect is the increase in sophistication rather than diversity. This distinction is theoretically important. A system may introduce advanced-looking vocabulary without meaningfully expanding the conceptual range of the text. In A3, the lexical sophistication score rises considerably, whereas diversity increases only moderately.

This pattern suggests that extensive generative-AI revision may produce what can be described as **lexical upgrading** rather than genuine lexical development. From a pedagogical perspective, these processes should not be treated as equivalent. Lexical development implies that the student has acquired and can independently deploy new vocabulary. Lexical upgrading may simply indicate that a model has substituted more formal expressions during revision.

AI-assisted academic writing should therefore be evaluated not only in terms of the lexical quality of the final text but also in terms of whether students understand and can reuse the vocabulary introduced by the system.

5.2 Sentence Complexity

The syntactic results show a nonlinear pattern. Minimal correction has little effect on dependency depth or clause density, while A2 and A3 produce noticeable increases.

The increase in mean sentence length under A3 is not interpreted automatically as evidence of superior writing. Sentence length is a coarse measure and can rise because of unnecessary clause accumulation. The accompanying rise in dependency depth provides stronger evidence that syntactic structure itself becomes more embedded.

However, greater structural complexity may also produce processing costs. AI-generated revisions can create sentences containing multiple nominal phrases, modifiers, and subordinate clauses, giving the text a conventionally academic appearance while increasing cognitive load.

The most desirable outcome therefore appears to be moderate syntactic enrichment rather than maximal complexity.

5.3 Semantic Preservation

Semantic similarity decreases as prompt intensity increases. A1 remains very close to the original text because the prompt is restricted to surface correction. A2 produces greater rephrasing, while A3 shows the largest semantic distance.

This result is important because generative rewriting can subtly change authorial intention. Even when an AI-generated sentence appears grammatically stronger, it may intensify a claim, reduce uncertainty, change argumentative emphasis, or introduce terminology that the original writer did not intend.

For this reason, semantic preservation should be treated as a core evaluation criterion in AI-assisted writing research rather than assuming that any linguistically improved revision is faithful to the source text.

5.4 Stylistic Homogenization

One of the most important findings concerns declining inter-writer feature distance. H0 texts exhibit the greatest diversity among student authors. As AI intervention increases, this distance decreases progressively.

The A3 condition shows the strongest convergence. Students whose independent drafts originally differed in vocabulary preference, sentence structure, and stylistic rhythm produce more similar texts following extensive AI revision.

This finding supports the manuscript's proposed concept of **prompt-induced linguistic homogenization**. The effect does not imply that every AI-revised text becomes identical. Rather, extensive prompting appears to move different texts toward a narrower statistical region characterized by similar academic vocabulary, sentence structures, transitions, and complexity patterns.

Such convergence may have significant implications for authorial voice. University writing instruction traditionally values both clarity and the development of an identifiable academic style. If generative tools repeatedly normalize student prose toward model-preferred patterns, writing may

@2026. This is an open access article distributed under the terms of the Creative Commons License [CC BY NC 4.0] and is available on www.wjcrt.org

become more formally polished but less individually distinctive.

5.5 Predictive Separability

The illustrative classifier performance further supports the existence of distinct linguistic profiles across intervention levels. Macro-F1 increases as AI assistance becomes stronger, with A3 being easier to distinguish from the human baseline than A1.

The purpose of this result is not to propose an AI-detection mechanism for disciplinary enforcement. Detection systems remain vulnerable to model changes, paraphrasing, genre effects, and false positives. Instead, classification performance is used analytically to demonstrate that progressively prompted texts acquire systematic linguistic properties.

Feature-importance analysis would be expected to assign high explanatory value to lexical sophistication, dependency depth, academic vocabulary share, sentence length, and semantic displacement.

Future Scope

Future research should extend the framework beyond static text comparison toward longitudinal student development.

A particularly important direction would be to observe whether repeated AI-assisted rewriting changes students' independent writing over time. If vocabulary or sentence structures introduced by generative models later appear in unaided writing, AI may function partly as a language-learning mechanism rather than solely as an external writing processor.

Cross-linguistic studies are also required. Multilingual students may experience stronger benefits from AI-mediated lexical assistance than highly proficient native or near-native writers. Research should therefore examine whether prompt-induced linguistic displacement differs according to English proficiency, first language, and disciplinary background.

Another important direction involves discipline-specific modeling. The lexical and syntactic features of engineering reports, legal argumentation, humanities essays, and medical writing differ substantially. A general Prompt Influence Index could therefore be extended into discipline-aware variants.

Future studies should also analyze prompt semantics directly. Rather than grouping prompts only by intervention intensity, researchers could classify them according to instructional purpose, including simplification, academicization, concision, argument strengthening, evidence integration, tone modification, and grammar correction.

A further research opportunity concerns authorial voice preservation. Neural similarity models could be trained to identify whether the writer-specific stylistic signature of an original text survives after AI-assisted revision.

Finally, AI writing systems themselves could incorporate linguistic preservation controls. A future educational interface could allow users to select settings such as "correct grammar while preserving vocabulary," "improve cohesion without increasing sentence complexity," or "suggest alternatives without rewriting the original." Such controls would shift generative AI from unrestricted text replacement toward pedagogically transparent assistance.

Conclusion

Generative AI is changing university writing at a level deeper than spelling or grammar correction. Prompt-based systems can reshape vocabulary choice, sentence structure, syntactic depth, semantic framing, and stylistic identity within a single revision cycle.

The manuscript addresses an underexplored research problem by moving beyond simple human-versus-AI comparison and focusing instead on **within-writer linguistic transformation across graduated prompt conditions**.

The proposed findings suggest that minimal AI assistance produces relatively conservative correction, while moderate assistance increases lexical sophistication and syntactic elaboration without completely displacing the writer's original profile. Extensive prompting, by contrast, produces the largest gains in formal linguistic complexity but also the strongest semantic drift and the greatest reduction in inter-writer stylistic variation.

This trade-off is central to the study. More complex language is not necessarily better language, and more academic-looking vocabulary does not automatically represent stronger student competence.

The proposed Prompt Influence Index provides a framework for quantifying how far AI-assisted language moves from an author's original linguistic baseline. Combined with lexical, syntactic, semantic, and stylistic measures, it offers a more nuanced basis for evaluating generative-AI writing practices.

The broader implication is that academic institutions should move away from treating generative AI exclusively as a detection problem. The more consequential issue is how AI changes the process through which students develop academic language. Effective policy and pedagogy should therefore focus on transparent assistance, authorial control, semantic preservation, and measurable learning rather than solely on whether AI was used.

Generative AI will likely remain embedded in university writing practices. The critical research task is therefore not to determine whether prompted English should exist, but to understand **how much linguistic agency remains with the**

@2026. This is an open access article distributed under the terms of the Creative Commons License [CC BY NC 4.0] and is available on www.wjcrt.org

Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 71-80
https://wjcrt.org/

student as the prompt becomes increasingly powerful.

References:

1. Baek, C., Tate, T., & Warschauer, M. (2024). "ChatGPT seems too good to be true": College students' use and perceptions of generative AI. *Computers and Education: Artificial Intelligence, 7*, 100294. https://doi.org/10.1016/j.caeai.2024.100294

2. Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–238. https://doi.org/10.2307/3587951

3. Goulart, L., Matte, M. L., Mendoza, A., Alvarado, L., & Veloso, I. (2024). AI or student writing? Analyzing the situational and linguistic characteristics of undergraduate student writing and AI-generated assignments. *Journal of Second Language Writing, 66*, 101160. https://doi.org/10.1016/j.jslw.2024.101160

4. Kim, J., Yu, S., Detrick, R., & Li, N. (2025). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies, 30*(1), 1265–1300. https://doi.org/10.1007/s10639-024-12878-7

5. Kohnke, L. (2024). Exploring EAP students' perceptions of GenAI and traditional grammar-checking tools for language learning. *Computers and Education: Artificial Intelligence, 7*, 100279. https://doi.org/10.1016/j.caeai.2024.100279

6. Kyle, K., & Crossley, S. A. (2015). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly, 49*(4), 757–786. https://doi.org/10.1002/tesq.194

7. 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

8. Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments, 11*, Article 9. https://doi.org/10.1186/s40561-024-00295-9

9. Mak, M. H. C., & Walasek, L. (2025). Style, sentiment, and quality of undergraduate writing in the AI era: A cross-sectional and longitudinal analysis of 4,820 authentic empirical reports. *Computers and Education: Artificial Intelligence, 9*, 100507. https://doi.org/10.1016/j.caeai.2025.100507

10. McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods, 42*(2), 381–392. https://doi.org/10.3758/BRM.42.2.381

11. Mizumoto, A., Yasuda, S., & Tamura, Y. (2024). Identifying ChatGPT-generated texts in EFL students' writing: Through comparative analysis of linguistic fingerprints. *Applied Corpus Linguistics, 4*(3), 100106. https://doi.org/10.1016/j.acorp.2024.100106

12. Tudino, G. (2024). A corpus-driven comparative analysis of AI in academic discourse: Investigating ChatGPT-generated academic texts in social sciences. *Lingua, 312*, 103838. https://doi.org/10.1016/j.lingua.2024.103838

13. UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO.

14. Liu, Y. (2024). Using generative artificial intelligence/ChatGPT for academic communication: Students' perspectives. *International Journal of Applied Linguistics*. https://doi.org/10.1111/ijal.12574

15.

@2026. This is an open access article distributed under the terms of the Creative Commons License [CC BY NC 4.0] and is available on www.wjcrt.org