© IJARW | ISSN (O) - 2582-1008
April 2026 | Vol. 7 Issue. 10
www.ijarw.com

AI'S DUAL IMPACT ON ACADEMIC VOCABULARY: A COMPARATIVE ANALYSIS OF STUDENT WRITING OF ENGLISH-MAJOR STUDENTS AT THE SCHOOL OF FOREIGN LANGUAGES – THAI NGUYEN UNIVERSITY

Do Thanh Mai [1], Pham Thi Phuong Thao [2]

School of Foreign Languages- Thai Nguyen University

ABSTRACT

The rapid proliferation of Large Language Models (LLMs), exemplified by AI-assisted writing tools, is reshaping the academic landscape. This study conducts an in-depth comparative analysis of Academic Vocabulary (AV) use between two corpora: Self-composed essays by 30 English-major students (Corpus A) and equivalent essays generated or edited with AI assistance (Corpus B). Employing corpus analysis methods (based on the AWL and lexical diversity metrics) combined with pragmatic analysis, the quantitative findings confirm that Corpus B exhibited a significantly higher frequency and complexity of AV ($p<0.01$) compared to Corpus A, suggesting that AI acts as an effective vocabulary amplifier. However, the qualitative analysis identified three core weaknesses in Corpus B: Collocation Errors; Contextual Inappropriacy, leading to a pseudo-voice; and Lack of Nuance. These findings not only caution against the risk of reduced cognitive engagement among learners but also hold profound implications for Applied Linguistics. The study emphasizes the necessity of shifting the teaching focus from generating complex vocabulary to critically evaluating and refining AI output, aiming to develop students' pragmatic evaluation competence.

Keyword: Academic Vocabulary, AI-Assisted Writing, Corpus Analysis, Pragmatic Competence

1. INTRODUCTION

1.1.Context and the Significance of Academic Vocabulary (AV)

In the digital age, academic writing has become a multimodal skill. AV, comprising neutral, precise, and abstract terms (such as analyze, hypothesis, paradigm), is an essential tool for constructing coherent arguments and expressing critical thinking. Coxhead's AWL (2000) remains a common benchmark for assessing AV proficiency. Mastery of AV signals that a learner has moved beyond everyday communication limits to participate in the academic discourse community. The advent of LLMs like ChatGPT-4 has created a "leap" in text generation capabilities. These tools can produce academically acceptable text in terms of grammar and vocabulary, seriously challenging traditional assessment methods and raising questions about academic authenticity.

Research Problem and GapMost initial studies on AI and writing focused on fluency and grammatical structure. However, there is a scarcity of research delving into the micro-level characteristics of language—specifically, the quality and contextual appropriateness of AV. Can AI, trained on vast academic data, convey the necessary nuance and naturalness for AV use? If students rely solely on AI, do they miss the deep learning process involved in lexical choice?

This study provides a comprehensive quantitative-qualitative analysis to assess not just the quantity but also the quality of AV utilized27.

1.2.Research Objectives

Quantitative: To compare the statistically significant differences in the frequency (token percentage), diversity (D-Measure), and complexity (AWL Sublist 5-10) of AV between Corpus A and Corpus B.

Qualitative: To analyze pragmatic and collocation errors in Corpus B to determine the extent of shortcomings in naturalness and contextual appropriateness. **Implication**: To propose a new teaching model that leverages AI as a learning tool

while strengthening students' critical evaluation competence regarding language output.

2.LITERATURE REVIEW

2.1.Theory of Academic Vocabulary and Complexity

AV is considered the bridge between colloquial and specialized language34. Coxhead's AWL (2000) consists of 570 core words. However, assessing AV goes beyond word lists to include the ability to use Academic Collocations (Hyland, 2008). Quality academic text requires not only many AV words but also their use to be grammatically accurate and semantically natural.

2.2.AI and Scaffolding Theory in Writing

Scaffolding Theory (Vygotsky) posits that learners can achieve more complex tasks with expert assistance. AI is viewed as a powerful scaffolding tool (Molnar, 2023), helping learners overcome barriers in grammar and fluency. However, digital scaffolding risks a boomerang effect: if support is excessive, it can reduce cognitive effort, leading to technological reliance and impeding deep learning.

2.3.Pragmatics and Language Authenticity

Pragmatics studies how language is used in context43. In academic writing, pragmatic appropriateness is crucial: an AV word might be grammatically correct but wrong in tone or inappropriate for the genre44. Previous studies (Hyland, 2019) indicated that non-native speakers often struggle with authenticity and personal voice45. The danger of AI is its tendency to produce a synthetic voice, which may be fluent but lacks the personal mark and pragmatic subtlety.

3.METHODOLOGY

3.1.Participants and Design

30 fourth-year English-major students at the College of Foreign Languages – Thai Nguyen University participated49. A paired comparison research design was used. Each student contributed two essays (Corpus A - Self-composed; Corpus B - AI-Assisted) on the same topic, which helped control for individual writing competence.

3.2.Data Collection Procedure

Corpus A: Students completed an academic essay within a time limit (90 minutes) without using any AI or translation tools, aiming to capture pure writing competence.

Corpus B: Students were required to use LLMs (e.g., GPT-4) to generate a draft or edit the language, and then finalize the essay. This simulated real-world AI use. The total corpus collected comprised 60 essays, totaling approximately 45,000 words.

3.3.Corpus Analysis Tools

Lextutor's Vocabulary Profiler (AWL): Used to quantify AV frequency (Token Percentage) and classify by Sublist.

AntConc: Used to calculate the collocates of the 20 most frequent AV words in each corpus.

SPSS: Used to perform a Paired Samples t-test to compare quantitative metrics between the two groups.

3.4. Qualitative Analysis

Two senior lecturers independently analyzed 30 representative essay pairs. The analysis focused on: Collocation Errors: Miscombinations of AV words with other nouns/verbs/adjectives (e.g., establish a problem instead of identify/address a problem).

Contextual Errors (Register/Tone): AV use that was overly formal or too strong compared to a neutral academic context.

Overuse of Flowery Language: Insertion of unnecessarily complex phrases that reduced the clarity of the text.

4. FINDINGS

4.1.Detailed Quantitative Results

Metric	Corpus A (Self-composed, N=60)	Corpus B (AI-Assisted, N=60)	t-value	p-value	Conclusion
AWL Frequency(%)	M=6.48 (SD=1.12)	M=9.85 (SD=1.56)	t(59)=−14.28	< 0.001	AI significantly increased AV frequency
Diversity (D-Measure)	M=0.45 (SD=0.03)	M=0.47 (SD=0.04)	t(59)=−1.95	55	No statistically significant difference
Complex AV Ratio (Sublist 5-10)	M=1.21 (SD=0.55)	M=2.53 (SD=0.88)	t(59)=−8.91	< 0.001	AI increased use of high-level vocabulary

Quantitative results analysis table between Corpus A & Corpus B

The results show that AI is a highly effective tool for amplifying the volume of academic vocabulary (AV). The significant increase in the frequency and complexity of AV confirms that AI easily retrieves less common words. However, the finding that lexical diversity did not significantly increase (D-Measure $p>0.05$) is crucial: AI does not generate a broader range of academic vocabulary, but rather tends to repeat a certain set of the most complex vocabulary.

2. Qualitative Results: Pragmatic and Collocation Errors Qualitative analysis revealed three main qualitative error types in Corpus B largely absent from Corpus A.

A. Collocation Errors These errors indicate a lack of understanding of the usage norms of academic English vocabulary, even when the individual word is correct.

Example 1 (Corpus B): "***The researcher formulated a new methodology to procure the data***." (Analysis: Procure (to obtain/acquire) is overly formal and typically used for goods, unnatural with data; the natural collocations are collect/obtain the data.)

Example 2 (Corpus B): "***It is important to enact the theory when writing***." (Analysis: Enact (to legislate/stage) does not collocate with theory; the natural collocations are apply/implement the theory.)

B. Contextual Inappropriacy

This relates to an unsuitability in the tone and nuance of academic writing.

Example 3 (Corpus B): "***It is imperative to promulgate the findings to the masses***." (Analysis: Imperative and promulgate create a tone that is too formal and emotional for a neutral academic essay. Furthermore, masses is colloquial and non-academic, creating a clash of register.)

Example 4 (Corpus B): Instead of "***The main problem is...***", AI suggested "***The predominant dilemma resides in...***" (Analysis: Replacing simple words with a complex structure (phrasal verb resides in + rare word predominant) is unnecessary, reducing clarity and making the text verbose.)

C. Lack of Nuance

In several instances, AI used AV that did not accurately convey the student's precise intent.

Example 5 (Corpus B): Student intended to say "***important vocabulary***," AI suggested "***The vocabulary is paramount***." (Analysis: Paramount (of supreme importance) conveys an absolute level of importance, which may be too strong compared to the student's original intention (important/crucial), leading to a loss of the writer's control over the argument.)

5. DISCUSSION

The quantitative and qualitative findings of this study have clearly depicted the two-sided impact of AI assistance on the use of Academic Vocabulary (AV). Quantitatively, the significant increase in the frequency and complexity of AV in AI-assisted essays (Corpus B) demonstrates AI's effectiveness as a vocabulary amplifier, leveraging large academic datasets to superficially replicate formal language characteristics. However, this numerical superiority was not accompanied by a corresponding increase in vocabulary diversity, suggesting AI tends to repeat a specific subset of complex vocabulary. This finding leads to the conclusion of a form of "Pseudo-Academic Lexical Proficiency", where the language appears fluent in form but lacks depth in authentic academic thought. The qualitative analysis reinforces this point by identifying serious errors in pragmatics, including collocation errors and contextual inappropriacy. These errors occurred because students skipped the crucial cognitive process of pragmatic evaluation 116and uncritically accepted the AI output. This results in a loss of control over the personal voice and diminishes the ability to perform the cognitive load necessary for deep learning. Consequently, the core contribution of this paper is the call for a shift in the focus of Applied Linguistics pedagogy: rather than measuring language output, the emphasis must be placed on the cognitive quality behind that output. Instructors must design learning activities that teach students to critically review, audit, and refine AI output, focusing specifically on collocation norms and nuance. Teaching must redefine the writer's role in the AI era as the ultimate critic, editor, and pragmatic evaluator of machine output, ensuring AI becomes an enhancement to academic competence, not a replacement for the learning process.

6. CONCLUSION

The quantitative and qualitative findings of this study clearly highlight the two-sided impact of AI support on the use of Academic Vocabulary (AV). Quantitatively, the significant increase in the frequency and complexity of AV in LLM-assisted writing demonstrates that AI effectively functions as a vocabulary amplifier. However, this quantitative superiority is not matched by a corresponding increase in vocabulary diversity, suggesting a tendency for repetition. This finding supports the conclusion of "Pseudo-Academic Lexical Proficiency". Qualitatively, the sustained identification of severe errors in pragmatics, notably collocation errors and contextual inappropriacy, indicates that students bypass the critical cognitive process of pragmatic evaluation when accepting AI output, leading to a loss of control over the personal voice and diminishing their capacity for deep learning. The core contribution of this paper is the appeal for Applied Linguistics to shift its focus from measuring language output to the cognitive quality driving it. We advocate for teaching students to critically review and refine AI output, particularly focusing on collocation norms and nuance. The writer's role must be redefined as the ultimate critic and pragmatic evaluator, ensuring AI serves as an enhancement rather than a replacement for their academic competence.

REFERENCES

[1]. Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–236.
[2]. Hyland, K. (2008). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Corpus Linguistics, 13*(1), 41–64.
[3]. Hyland, K. (2019). *Teaching and researching writing* (3rd ed.). Routledge.
[4]. Molnar, A. (2023). ChatGPT in the classroom: Friend or foe? *Journal of Applied Linguistics and Technology, 5*(1), 45-60.
[5]. Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students: Essential tasks and skills* (3rd ed.). University of Michigan Press