Analysis of Linguistic Errors in ESL Writing Using an AI-Driven Assistant Tool

Suhail Shafea, Ph.D., Assistant Professor, Department of Languages and Translation, College of Arts and Humanism, Taibah University, Saudi Arabia

Abdulrahman Olwi, Ph.D., Associate Professor, Department of Curriculum and Instruction, College of Education, Taibah

***Abstract*—**

This study investigates the potential of Gemini Advanced as an AI-driven tool for analyzing L2 errors and generating feedback. Thirty native Arabic-speaking undergraduate students majoring in English participated in this research, engaging in peer feedback prior to receiving the feedback generated by the AI tool. A mixed-methods approach was employed to compare the quantity and variety of errors identified by human and AI evaluators, as well as to analyze the quality and consistency of the feedback provided. Findings indicate that AI excels in identifying and correcting grammatical and spelling errors, surpassing human evaluators in terms of efficiency. However, AI exhibits limitations in addressing pragmatic and contextual nuances and tends to overcorrect. A qualitative analysis of the feedback reveals inconsistencies and occasional inaccuracies in AI's assessment of student writing. The study concludes by advocating for a balanced approach to integrating AI in L2 writing tasks, highlighting the necessity for human oversight and critical evaluation of AI-generated feedback to maximize its potential advantages for language learning.

***Index Terms*—artificial intelligence (AI), second language (L2), English language, learning, teaching.**

الملخص—

تبحث هذه الدراسة في إمكانات برنامج *جيميناي أدفانسد* كأداة مدعَّمة بالذكاء الاصطناعي لتحليل أخطاء اللغة الثانية وتوليد الملاحظات. شارك في هذا البحث ثلاثون طالبًا جامعيًا من الناطقين باللغة العربية المتخصصين في اللغة الإنجليزية، حيث شاركوا في ملاحظات الأقران قبل تلقي الملاحظات الناتجة عن أداة الذكاء الاصطناعي. تم استخدام نهج الأساليب المختلطة للمقارنة كميا ونوعيا بين الأخطاء التي حددها المقيِّمون البشريون والذكاء الاصطناعي، وكذلك لتحليل جودة واتساق الملاحظات المقدمة. تشير النتائج إلى أن الذكاء الاصطناعي يتفوق في تحديد الأخطاء النحوية والإملائية وتصحيحها، متجاوزًا المقيِّمين البشريين من حيث الكفاءة. ومع ذلك، يُظهر الذكاء الاصطناعي قيودًا في معالجة الفروق الدقيقة التداولية والسياقية ويميل إلى الإفراط في التصحيح. يكشف التحليل النوعي للملاحظات عن تناقضات وعدم دقة عرضية في تقييم الذكاء الاصطناعي لكتابة الطلاب. تدعو نتائج الدراسة إلى اتباع نهج متوازن لدمج الذكاء الاصطناعي في مهام الكتابة باللغة الثانية، مع تأكيد ضرورة الإشراف البشري والتقييم الناقد للملاحظات التي يُوَلِّدها الذكاء الاصطناعي لتعظيم الأثر في تعلم اللغة.

***الكلمات المفتاحية—*الذكاء الاصطناعي، اللغة الثانية، اللغة الإنجليزية، التعلم، التدريس.**

I. INTRODUCTION

Previous research on second language acquisition demonstrated that second language (L2) learners often encounter numerous errors throughout their learning process (MacNeil et al., 2022). Recognizing and clearly articulating the linguistic errors committed by L2 learners holds considerable importance, as it enhances the comprehension of their proficiency concerning their target language

(Lightbown & Spada, 2021). Traditionally, L2 learners have depended on either their instructors or their more proficient peers to evaluate their writing, identify weaknesses, and propose a developmental plan to enhance their L2 knowledge. Nevertheless, the rapid advancement of evolving AI-driven tools capable of aiding L2 learners has prompted both educators and learners to incorporate these tools into the processes of L2 instruction and acquisition. Still, there persists an uncertainty on how consistent and reliable AI-driven tools are when analyzing L2 learners' errors. This empirical study aims to contribute to this growing area of research by investigating the ways in which L2 learners may employ Gemini as an AI-driven tool for L2 error analysis, with a particular emphasis on the tool's effectiveness and limitations in this context.

II. Literature Review

A growing body of literature exhibits the effectiveness of AI-driven tools in supplementing second language learning/teaching. In fact, AI is regarded as an innovative tool that could drastically transform English language instruction (Iqbal et al., 2022). Teachers should consider utilizing it to create immersive learning experiences that adhere to distinct cultural norms and ethical standards (Kohnke et al., 2023). AI tools are around-the-clock assistants, and the technologies embedded in these modern tools have proven effective (Abdelghani et al., 2023). According to Schmidt-Fajlik (2023), AI-driven tools are a crucial advantage that has been incorporated into English language teaching and learning

as they provide students with extraordinary learning assistance at any time of day, flexibly and conveniently. This accessible support allows English language learners to practice independently, at their own pace (Bonsu & Baffour-Koduah, 2023). It could also imitate human conversations and communicate with students in a natural-like way (Yan, 2023), making them feel as if they were talking to a true native speaker. Moreover, Ouyang and others found that AI tools can perform better than the human eye and detect complicated errors that might otherwise be missed.

A recent study by Punar Özçelik and Ekşi (2024) showed that leveraging AI tools in English language learning personalizes the learning experience and adds considerable momentum that extends the benefits and produces learning solutions. Also, AI tools can be used to improve learners' critical thinking by integrating resources into English language learning activities (Amaro et al., 2023). Omar and others (2024) similarly discussed enhancing engagement in English language learners by utilizing tailored AI-generated content that considers individual differences.

Another study by Fitria (2023) revealed that AI can provide robust feedback by identifying students' strengths and areas for improvement. Hong (2023) noted that AI-powered simulation software can analyze students' responses and offer valuable recommendations. Ausat and others (2023) found that learners seek AI assistance for proofreading their assignments, which aids in spelling and grammar checks, ensuring polished writing styles. Furthermore, they utilize AI to search

for and double-check clauses, phrases, and modifiers, enabling them to deconstruct sentences into familiar basic components. This provides invaluable insights that clarify structural syntax and show the potential of AI to deal with linguistic complexity and ensure language coherence.

English teachers can use AI to analyze student learning outcomes and adjust lesson plans. Analyzing performance in essays and written assignments is time-consuming, but AI provides valuable insights into performance patterns and common errors, offering personalized feedback based on set criteria. This ensures consistency and fairness, enhancing reliability and validity. Teachers can then create prompts and exercises encouraging students to express complex ideas. Additionally, AI can help implement innovative deep learning models (Baskara & Mukarto, 2023) and aid students in mastering English vocabulary and grammar through improved communication skills (Yan, 2023).

Focusing on the use of AI for language error analysis, AI-driven tools have rapidly developed and improved as content analysis and error detection are increasingly recognized (Kasneci et al., 2023). These tools are equipped with semantic and content analysis capabilities to examine the context in depth, identify grammatical errors, and ensure that the language scope is well understood (Dai et al., 2023). Bhat et al. (2022) claim that these technological assets have transformed practical methodologies by detecting and correcting language errors. These tools offer solutions to L2 learners (Murtaza et al., 2022) since they can identify common language learning errors and mistakes, resolve

writing issues, and suggest structure improvement to ensure the entire meaning of writing is fully understood (Mayer et al., 2022). Previous research shows that AI-driven error analysis can address language discrepancies and offer solutions that enhance overall grammatical accuracy (Ouyang et al., 2022), improve the clarity and consistency of writing (Wu et al., 2023; Wei et al., 2022), and enable learners to convey their writing coherently and free of errors (Santos 2023).

For example, almost all AI tools can locate instances where words are falsely capitalized through maintaining text consistency and readability (Schwichow et al., 2022). Thus, learners can be assisted with capitalization mistakes, which are considered among the most frequent errors committed by L2 learners (Marmo, 2022; Kranz et al., 2022). Similarly, AI tools could assist in correcting the frequent errors in subject-verb agreement (Zhai & Nehm, 2023) caused by subject-verb mismatch (Moore et al., 2022), along with spelling errors (Liu et al., 2023). Khosravi and others (2023) stated that spacing discrepancies are other writing issues that AI tools can identify and modify to ensure content quality.

Some researchers (e.g., Liu et al., 2023) have also asserted that AI tools could identify instances where words are misused or incorrectly chosen, thereby ensuring the quality of the written text, as even minor word choices can affect the meaning and tone of the content. Moreover, AI tools can deal with syntactical errors involving wrong arrangement and run-on sentences that can lead to confusion and ambiguity (Ji et al., 2022). A run-on sentence is another common problem that occurs when two independent clauses are joined inappropriately

without adequate punctuation (Molenaar, 2022). AI tools effectively identify and fix such sentences (Alfredo et al., 2024), ensuring that the content flows smoothly and logically (Dai et al., 2023).

Conversely, a conservative approach to utilizing AI has emerged. Kohnke et al. (2023), Hong (2023), and Baskara and Mukarto (2023) examined the varying levels of engagement that each learner may exhibit. Learners' reactions and responses differ significantly when encountering such an advanced tool. Some may fully embrace it, while others might approach it cautiously, evaluating its effectiveness compared to traditional learning methods. According to Lo (2023), some learners may encounter considerable challenges when using AI, as it can expose them to unpredictable situations that lead to confusion and frustration. Consequently, incorporating AI in acquiring the English language elicits diverse attitudes, thus necessitating further research to thoroughly investigate its efficacy before drawing overarching conclusions.

Previous studies have identified various challenges and limitations that may hinder language acquisition when using AI as a pedagogical tool. One such challenge is unlocking new avenues for linguistic growth and cultural exchange (Baskara & Mukarto, 2023). As learners become increasingly familiar with innovative technologies, artificial intelligence is gaining popularity. Furthermore, AI's inherently conversational nature enhances interactive experiences and simulates authentic dialogues while reducing the risk of harsh judgments (Kohnke et al., 2023). Additionally, both teachers and students have the potential

to misuse this technology. For example, the personalized feedback of educators fosters a level of empathy between them and their students that is difficult to replicate (Hong, 2023). Although artificial intelligence adeptly manages language nuances, cultural insights, and complex structures, its efficacy remains dependent on advancements, particularly in emotional intelligence (Hariri, 2023).

Clearly, AI tools require a more comprehensive contextual understanding (Fitria, 2023). They generate coherent responses quickly but struggle with linguistic nuances essential for effective communication (Baskara & Mukarto, 2023). This limitation impairs their ability to provide meaningful interpretations (Schmidt-Fajlik, 2023) and explanatory feedback (Yan, 2023). AI also finds it challenging to offer personalized guidance and correct errors, potentially confusing language development (Bonsu & Baffour-Koduah, 2023).

Another limitation is that AI relies entirely on existing data and patterns, which may inadvertently generate errors (Lo, 2023) and negatively impact student learning. Moreover, language learning is a highly individualized process, meaning that learners gain significant benefits from constructive criticism, which remains questionable in AI responses (Iqbal et al., 2022). These limitations could hinder students' debating skills and interpersonal communication (Sarrion, 2023). Furthermore, inadequate internet access limits AI use in language learning (Rudolph et al., 2023). Concerns about data privacy and algorithmic bias persist, even among those fully accessing AI (Fitria, 2023). Protecting learner data is critical over AI educational applications (Baskara & Mukarto, 2023). If

effectively addressed, the challenges posed by AI could lead to a dynamic and inclusive language learning environment (Ausat et al., 2023).

Research Questions

1. To what extent can Artificial Intelligence (AI)-driven tools effectively conduct error analysis on written texts and provide comprehensive feedback to language learners?
2. What are the qualitative and quantitative differences between peer-generated feedback and AI-generated feedback in the context of linguistic error analysis?

III. METHODOLOGY

Google's Gemini Advanced AI service was selected for this research as an AI-driven tool. A total of 42 undergraduate English language students, all native Arabic speakers, were invited to participate in this empirical study; only 30 students agreed to take part; they were asked to write a text, provide feedback to their classmates, receive feedback generated by Gemini, and then compare the peer-review feedback with the Gemini-generated feedback. The effect of integrating peer error analysis with AI-generated feedback on students' language development was examined. Participants engaged in peer error analysis prior to receiving AI-generated feedback, fostering cooperative learning and critical thinking. This method was overwhelmingly perceived as beneficial, promoting active learning and enhancing awareness of language weaknesses. The process

involved language learners applying their existing language skills to discuss each other's work before receiving AI-generated feedback. This multi-faceted approach encouraged students to:

1. Actively apply their language knowledge. Peer error analysis required language learners to apply their understanding of different language aspects such as grammar, vocabulary, and syntax within specific contexts.
2. Develop critical thinking skills. Evaluating peer work promoted a deeper understanding of language use.
3. Increase awareness of errors. The process reinforced attention to potential errors by encouraging self-reflection and future improvement.
4. Promote learner autonomy. Language learners got a chance to actively go beyond just getting feedback. Rather, they took greater responsibility for language development.

IV. Findings and Discussion

This research explores the potential application of Gemini Advanced as an AI-driven assistant tool for L2 error analysis and examines its effectiveness and limitations in this area. The errors identified and corrected by the participants were compared to those identified and corrected by Gemini Advanced. The comparison encompassed both numerical and content analysis of peer feedback alongside feedback generated by the AI tool.

Table 1

ERRORS IDENTIFIED AND CORRECTED BY THE PARTICIPANTS COMPARED TO THOSE IDENTIFIED AND CORRECTED BY GEMINI ADVANCED

Type of Error		Evaluator	
		Peer	AI
Punctuation	Identified	2	12
	Corrected	2	31
Capitalization	Identified	7	36
	Corrected	6	48
Missing words	Identified	2	4
	Corrected	2	4
Word choices	Identified	6	20
	Corrected	6	34
Spelling	Identified	21	107
	Corrected	20	117
Redundancy	Identified	2	12
	Corrected	2	13
Subject-verb agreement	Identified	7	14
	Corrected	7	21
Run-on	Identified	2	4
	Corrected	2	4
Fragmented	Identified	0	2
	Corrected	0	3
Comma splices	Identified	1	1
	Corrected	1	3
Preposition	Identified	0	0
	Corrected	0	1
Awkward phrase	Identified	4	72
	Corrected	4	70
Pragmatics	Identified	0	0
	Corrected	0	0
Other	Identified	5	2
	Corrected	5	5
Total	Identified	59	286
	Corrected	57	354

The analysis comparing peer feedback and AI feedback highlighted notable discrepancies in detecting and correcting errors. Table 1 illustrates that AI consistently identified a greater number of errors than peer reviewers across most

categories, except for the category pertaining to pragmatic errors, where neither identified any. Peers identified a total of 59 errors and successfully corrected 57, while AI identified 286 errors and corrected 354. The most pronounced difference was observed in the category of spelling errors, with peers identifying 21 errors and correcting 20, whereas the AI detected 107 errors and corrected 117. Another significant disparity was discerned in awkward phrases, where peers identified and corrected four errors, in contrast to AI, which detected 72 errors and corrected 70. Notably, AI exhibited significant patterns of overcorrection across several categories, including punctuation (12 identified/31 corrected), capitalization (36 identified/48 corrected), word choice (20 identified/34 corrected), redundancy (12 identified/13 corrected), subject-verb agreement (14 identified/21 corrected), fragment sentences (2 identified/3 corrected), comma splices (1 identified/3 corrected), preposition (0 identified/1 corrected), and in the "Other" category (2 identified/5 corrected).

TABLE 2
PAIRED T-TEST RESULTS OF THE STATISTICAL COMPARISON OF ERROR IDENTIFICATION AND CORRECTION

Test Aspect	T-Statistic	P-Value	Statistical Significance ($P < 0.05$)	Conclusion
Error Identification	-2.23	0.044	Yes	Significant difference: AI identified more errors than peers.
Error Correction	-2.71	0.018	Yes	Significant difference: AI corrected more errors than peers.

Regarding the statistical comparison of error identification and correction, the paired t-test results provided insights into the effectiveness of peer correction and AI correction for identifying and correcting language errors. As shown in Table 2, the test assessed whether there is a statistically significant difference in performance between peer and AI evaluations. The t-statistic of -2.23 with a p-value of 0.044 indicates a statistically significant difference between peers and AI, showing that the AI identified more errors than the peers. For error correction, the t-statistic of -2.71 with a p-value of 0.018 suggests a significant difference, with the AI correcting more errors than the peers. Consequently, both t-tests revealed that the AI evaluator outperforms peer evaluations in identifying and correcting errors, as shown by the p-values being below 0.05.

Accordingly, in agreement with previous research (Abdelghani et al., 2023; Reiss, 2023; Dai et al., 2023; Aust et al., 2023), the AI-driven tool proved to be more efficient at spotting and correcting language errors than the peer evaluators. This disparity may result from several factors. First, L2 learners might lack sufficient linguistic knowledge and can be inconsistent in their attention to detail. They may typically focus their text analysis on more apparent or impactful errors, whereas AI has a better ability to apply grammatical rules consistently across various texts. Second, AI tools are trained on large text corpora, enabling them to recognize subtle linguistic patterns and variations that L2 learners might miss. Third, AI's programming enables simultaneous evaluation across multiple

linguistic dimensions (syntax, semantics, and style) in a quick and consistent manner that L2 learners may find challenging to replicate.

Nonetheless, AI's deficiency in recognizing errors in the pragmatic category and its propensity to overcorrect—evidenced by the number of corrections exceeding the identified errors across various categories—suggests inherent limitations in its contextual comprehension and pragmatic judgment, in which humans typically demonstrate superiority. These findings agree with a few studies (Fitria, 2023; Baskara & Mukarto, 2023; Schmidt-Fajlik, 2023; Algaraady & Mahyoob, 2023) and highlight the significance of human judgment in assessing AI-generated error analysis, ensuring that contextual understanding of language use is not neglected.

In analyzing AI-generated feedback, Table 3 presents an example of how AI identifies, describes, and classifies writing errors and proposes corrections. This feedback possesses the capacity to enhance second language learning significantly. By accurately pinpointing errors and providing individualized suggestions for improvement, AI enables learners to acknowledge their deficiencies and actively participate in the learning process. This cultivates a profound understanding of grammatical structures, vocabulary usage, and stylistic conventions, which facilitates the enhancement of their linguistic knowledge and skills. Furthermore, it can empower L2 learners by assisting them in recognizing their shortcomings, enhancing their awareness, and facilitating the advancement of their linguistic knowledge and skills.

Nonetheless, the presented example, along with other instances within the current dataset, reveals multiple occurrences in which the AI tool inadequately addressed the errors made by L2 learners and failed to deliver a comprehensive analysis. This observation aligns with Lo (2023) and suggests that the effectiveness of AI-generated feedback may be subject to certain limitations. Several such occurrences will now be presented and discussed.

TABLE 3
EXAMPLE 1

"The first thing i want to do after graduating is travel abroad and enjoying my vacation . Then i want to take the IELTS exam to pursue my master's degree somewhere abroad And if things go as planned i get a nice job , And hopefully start a family and live happily afterwards."			
Original Text	Error Description	Error Classification	Proposed Correction
"The first thing i want to do after graduating is travel abroad and enjoying my vacation ."	Incorrect capitalization, awkward phrasing, and missing word.	Mechanics (Capitalization) Vocabulary	"The first thing I want to do after graduating is travel abroad and enjoy a vacation."
"Then i want to take the IELTS exam to pursue my master's degree somewhere abroad"	Incorrect capitalization.	Mechanics (Capitalization)	"Then I want to take the IELTS exam to pursue my master's degree somewhere abroad..."
"And if things go as planned i get a nice job ,"	Incorrect capitalization.	Mechanics (Capitalization)	"And if things go as planned, I'll get a nice job..."
"And hopefully start a family and live happily afterwards ."	Slightly wordy or cliché ending.	Vocabulary	"And hopefully start a family and find happiness."

Several inconsistencies in AI's feedback are evident in Table 4. First, an incorrect error identification occurred, as the AI flagged the word "going" as improperly capitalized, despite its correctness. This inaccuracy in error identification has the potential to mislead students. Second, there was a misdiagnosis of the type of error, as the AI classified "Will" as incorrectly capitalized when the actual error pertained to either a misspelling or an inappropriate word choice. Such misclassification may lead to confusion among learners. Third, the absence of corrective feedback was notable; although the second phrase was labeled as "awkward," the AI failed to offer a revised version, thereby leaving learners without the necessary guidance for improvement. Fourth, the identification of errors proved inadequate, as the AI addressed the missing article "the" without pointing it out as an error. While this oversight may facilitate quick fixes, it may not be conducive to effective learning, as students might not comprehend the rationale behind the correction. Last, an erroneous correction was introduced, with the AI proposing the non-existent word "graduationuation" in its suggested correction. Given these inconsistencies, learners of a second language may not derive benefits from this feedback; rather, they could experience confusion and a sense of being misled.

TABLE 4
EXAMPLE 2

"Job hunting is first thing I am going to do after graduation. You might ask why not take a break? Will first, I want to become financialy stibilaized, so I can no longer worry about mony."			
Original Text	Error Description	Error Classification	Proposed Correction
"Job hunting is first thing I am going to do after graduation."	Incorrect capitalization ("going")	Mechanics (Capitalization)	"Job hunting is the first thing I am going to do after graduationuation."
"You might ask why not take a break? Will first,..."	Incorrect capitalization ("Will"), awkward phrasing	Mechanics Vocabulary	"You might ask why not take a break? Well, first, ..."
"... I want to become financialy stibilaized, so I can no longer worry about mony."	Misspellings of "financially" and "stabilized," incorrect usage of "mony"	Spelling Vocabulary	"I want to become financially stable, so I can no longer worry about money."

Table 5 shows how AI's capabilities are limited in addressing metalinguistic factors. In the example, the student began their writing with, "Every one has his own plans after attending university. For me ...." This excerpt reveals two points for discussion. First, there was a false positive error detection, as the AI flagged "one" as incorrectly capitalized, even though it was written in lowercase. This indicates a flaw in AI's ability to identify capitalization errors accurately. Second, there was an issue distinguishing between "Every one" and "Everyone," as AI appears to consider "Everyone" to be the correct form. To elaborate, the AI seems to fail to consider that the use of "Every one" in this example is meant to

emphasize each individual, not a group. This is evident from using the pronoun "his" to refer to individuals as well as starting the next sentence with "For me….".

Another point to highlight is the importance of analyzing split ideas. The student divided a single thought into two separate sentences, which could be improved for clarity and conciseness. He wrote, "Recently, I wanted to become a Police officer. But I have been inspired by a person to be a professor." The AI pointed out the awkwardness of the second sentence but did not offer a direct correction. Instead, it merely suggested changing "But" to "However." Additionally, the AI could have provided a more comprehensive suggestion, such as combining the two sentences or rephrasing them to express the student's intention better. This analysis seems accurate; however, the AI's response demonstrates a limited ability to address the underlying issue of the split idea. Simply changing "But" to "However" does not resolve the lack of cohesion between the two sentences.

Here are some possible explanations to understand why AI might have provided such a limited response:

1. Sentence-Level Focus: AI may primarily be designed to analyze and correct individual sentences rather than larger chunks of text. This could hinder its ability to identify and address issues that span multiple sentences.
2. Lack of Contextual Understanding: AI might not fully grasp the student's intended meaning or the relationship between the two sentences. This could make it difficult to suggest a correction that accurately reflects the student's thoughts.

3. Limited Scope of Suggestions: AI's suggestions might be constrained by their programming or training data. It could be trained to provide only certain types of corrections, such as word substitutions or minor grammatical adjustments.

To improve AI's ability to manage such cases, it could be beneficial to enhance contextual understanding, expand the scope of suggestions, and provide more targeted feedback. By addressing these areas, AI could become a more valuable tool for students and writers looking to enhance their expression and communication skills.

TABLE 5
EXAMPLE 3

"Every one has his own plans after attending university. For me, I want to…" "For me, I want to prepare for taking the IELts exam to apply for academic Posts. Moreover, I really desire to continue my higher studies, and become a great Professor. recently, I wanted to become a Police officer. But, I have been inspired by a person to be a professor."			
Original Text	Error Description	Error Classification	Proposed Correction
"Every one has his own plans after attending university."	Incorrect capitalization ("one"). Also, slightly formal phrasing.	Mechanics Vocabulary	"Everyone has their own plans after attending university."
"Moreover, I really desire to continue my higher studies, and become a great Professor."	Incorrect capitalization of "professor."	Mechanics (Capitalization)	"Moreover, I really desire to continue my higher studies and become a great professor."

“recently, I wanted to become a Police officer.”	Incorrect capitalization of “police.”	Mechanics (Capitalization)	“Recently, I wanted to become a police officer.”
“But, I have been inspired by a person to be a professor.”	No error, but slight awkwardness in phrasing.	Vocabulary	“But, I was inspired by someone to become a professor.”

Table 6 shows an example of one participant who exhibited inaccuracies in his writing, specifically stating, “…skill writing to…” and “…skills of language.” In both cases, AI described the errors as “awkward phrasing,” although these two errors could be deliberately linked to the student's first language. More specifically, the participant's errors can be attributed to a direct translation from Arabic to English rather than labeling them as “awkward phrasing.” This indicates that the AI failed to recognize the potential impact of the learner's first language (Arabic) and culture on language assessment.

TABLE 6
EXAMPLE 4

“First of all, I would like to do after graduation is to improve my skills on different Places. I would like to improvment my skill writing to Pass the ILETS examination, I must do some prectse on writing and other skills of Language.”			
Original Text	Error Description	Error Classification	Proposed Correction
“First of all, I would like to do after graduation is to improve my skills on different Places.”	Incorrect phrasing, possible capitalization issue.	Vocabulary Mechanics (Capitalization)	“First of all, after graduation I would like to improve my skills in various areas.”

"I would like to improvment my skill writing to Pass the ILETS examination..."	Incorrect spelling of "improvement," awkward phrasing.	Spelling Vocabulary	"I would like to improve my writing skills to pass the IELTS examination..."
"...I must do some prectse on writing and other skills of Language."	Incorrect spelling of "practice," awkward phrasing.	Spelling Vocabulary	"...I must practice my writing and other language skills."

Table 7 displays the analysis of a text written by a student who expressed his ideas throughout it without using full stops to separate them into sentences. The AI divided the text into smaller segments and identified and corrected most of the errors. However, the analysis and correction focused mainly on surface errors, with some remaining unrecognized and uncorrected. Specifically, the AI overlooked the phrase "before that i have to travle to landon though and relax a bit," which may lead the student to remain unaware of these errors and assume that his original phrase is correct. This finding aligns with Lo (2023) and indicates that relying solely on AI for students with limited language knowledge and skills may provide insufficient feedback to learners, thereby hindering language development. It also emphasizes the teacher's critical role, suggesting that AI should be used as a learning or assistive tool.

TABLE 7
EXAMPLE 5

Table 7			
"My plan after graduation is to persue my hobby of making videos, I hope to get millions of views and become rich and be lazy for the rest of my life , *In case that dosint work i will be persuing higher studies and work hard so i dont become homeless, then i will look for a job to pay the pills before that i have to travle to landon though and relax a bit And enjoy theater."*			
Original Text	Error Description	Error Classification	Proposed Correction
"My plan after graduation is to persue my hobby of making videos..."	Misspelling of "pursue" (1)	Spelling	"My plan after graduation is to pursue my hobby of making videos..."
"I hope to get millions of views and become rich and be lazy for the rest of my life ,"	Awkward phrasing and a comma splice (run-on sentence). (2)	Vocabulary Grammar Punctuation	"I hope to get millions of views, become rich, and then relax for the rest of my life."
"In case that dosint work..."	Misspelling of "doesn't" (3)	Spelling	"In case that doesn't work..."
"... i will be persuing higher studies and work hard..."	Misspelling of "pursuing"(4)	Spelling	"...I will be pursuing higher studies and work hard..."
"...so i dont become homeless, then i will look for a job to pay the pills..."	Incorrect spelling of "bills," (5) awkward phrasing (6)	Spelling Vocabulary	"...so I don't become homeless. Then, I will look for a job to pay my bills..."
"...And enjoy theater."	Awkward punctuation before "And". (7) Lacking strong connection to preceding sentences. (8)	Punctuation Grammar	"...and enjoy some theater."

V. CONCLUSION

This study investigates the capabilities of Gemini Advanced as an AI-driven tool for analyzing L2 errors and providing detailed feedback to second language learners. The findings of this research indicate that, although Gemini Advanced demonstrates superior capabilities in identifying and correcting technical aspects of language such as spelling and grammatical errors, outperforming human evaluators in both efficiency and scope, it exhibits limitations in capturing pragmatic and contextual insights. The AI's tendency to overcorrect, coupled with occasional inconsistencies in its feedback, highlights the essential need for teachers' oversight when implementing this technology in language learning contexts.

The implications of these findings hold considerable significance for second language (L2) pedagogy. AI tools, such as Gemini Advanced, have the capacity to empower L2 learners by offering detailed and individualized feedback, thereby facilitating greater autonomy in the learning process. However, the limitations acknowledged in this study necessitate a cautious approach. It is imperative for language learners to recognize that feedback generated by AI tools requires thorough evaluation and validation and should not be accepted without criticism. Language instructors must also maintain their roles as central figures in guiding students in interpreting and applying AI-generated feedback, ensuring that it complements rather than substitutes human interaction and instruction.

As AI technology continues to evolve, there exist numerous avenues for investigation and exploration that remain essential for future research. Additional studies are needed to verify the efficacy of AI tools for L2 error analysis and to identify methods that may lead to more contextually sensitive feedback. Research focused on students' awareness of AI tools and the necessity of critically assessing their outputs is equally important. Moreover, further exploration of the pedagogical applications of AI tools, including their integration into varied L2 learning environments and their influence on learner motivation and autonomy, is also essential.

Ultimately, this study advocates for a balanced approach, where the integration of human insight with the capabilities of AI holds the key to unlocking new levels of language proficiency and fostering a more inclusive learning experience that is both engaging and accessible to a diverse range of learners.

REFERENCES

Abdelghani, R., Wang, Y. H., Yuan, X., Wang, T., Lucas, P., Sauzéon, H., & Oudeyer, P. Y. (2023). Gpt-3-driven pedagogical agents to train children's curious question-asking skills. *International Journal of Artificial Intelligence in Education*, 1-36.

Alfredo, R., Echeverria, V., Jin, Y., Yan, L., Swiecki, Z., Gasevic, D., & Martinez-Maldonado, R. (2024). Human-Centred Learning Analytics and AI in Education: a Systematic Literature Review. Computers and Education: Artificial Intelligence. 6. 100215. 10.1016/j.caeai.2024.100215.

Algaraady, J. & Mahyoob, M. (2023). ChatGPT's Capabilities in Spotting and Analyzing Writing Errors Experienced by EFL Learners. Arab World English Journal (AWEJ) Special Issue on CALL Number 9. Pp.3- 17. DOI: https://dx.doi.org/10.24093/awej/call9.1

Amaro, I., Della Greca, A., Francese, R., Tortora, G., & Tucci, C. (2023). AI Unreliable Answers: A Case Study on AI. In H. Degen, & S. Ntoa (Eds.), Artificial Intelligence in HCI. HCII 2023. Lecture Notes in Computer Science (pp. 23-40). Springer, Cham. https://doi.org/10.1007/978-3-031-35894-4_2

Ausat, A. M. A., Massang, B., Efendi, M., Nofirman, & Riady, Y. (2023). Can AI Replace the Role of the Teacher in the Classroom: A Fundamental Analysis. Journal on Education, 5(4), 16100-16106. https://jonedu.org/index.php/joe

Baskara, F. X. R. & Mukarto, F. X. (2023). Exploring the Implications of AI for Language Learning in Higher Education. Indonesian Journal of English Language Teaching and Applied Linguistics, 7(2), 343- 358. https://ijeltal.org/index.php/ijeltal/article/view/1387

Bhat, S., Nguyen, H. A., Moore, S., Stamper, J., Sakr, M., & Nyberg, E. (2022). Towards automated generation and evaluation of questions in educational domains. In *Proceedings of the 15th international conference on educational data mining* (Vol. 701).

Bonsu, E. M., & Baffour-Koduah, D. (2023). From the consumers' side: Determining students' perception and intention to use AI in Ghanaian higher education. Journal of Education Society & Multiculturalism, 4(1), 1-29. https://doi.org/10.21203/rs.3.rs-2686760/v1

Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y. S., Gašević, D., & Chen, G. (2023). Can large language models provide feedback to students? A case study on AI. In *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)* (pp. 323-325). IEEE.

Fitria, T. N. (2023). Artificial intelligence (AI) technology in OpenAI AI application: A review of AI in writing English essay. Journal of English

Language Teaching, 12(1), 44-57. https://doi.org/10.15294/elt.v12i1.64069

Hariri, W. (2023). Unlocking the potential of AI: A comprehensive exploration of its applications, advantages, limitations, and future directions in natural language processing. Technology, 15(2), 1-23. http://arxiv.org/abs/2304.02017

Hong, W. C. H. (2023). The impact of AI on foreign language teaching and learning: opportunities in education and research. Journal of Educational Technology and Innovation (JETI), 5(1), 37-45. https://jeti.thewsu.org/index.php/cieti/article/view/103

Iqbal, N., Ahmed, H., & Azhar, K. A. (2022). Exploring Teachers' Attitudes Towards Using AI. Global Journal for Management and Administrative Sciences, 3(4), 97-111. https://doi.org/10.46568/gjmas.v3i4.163

Ji, H., Han, I., & Ko, Y. (2022). A systematic review of conversational AI in language education: focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, *55*(1), 48–63. https://doi.org/10.1080/15391523.2022.2142873

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., & Kasneci, G. (2023). ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. 103. 102274. 10.1016/j.lindif.2023.102274.

Khosravi, H., Sadiq, S., & Amer-Yahia, S. (2023). Datamanagement of AI-powered education technologies: Challenges and opportunities.Learning Letters, 1, Article 2. https://doi.org/10.59453//XLUD7002

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). AI for Language Teaching and Learning. RELC Journal, 54(23), 1-14. https://doi.org/10.1177/00336882231162868

Kranz, J., Baur, A., & Möller, A. (2023). Learners' challenges in understanding and performing experiments: a systematic review of the literature. *Studies in Science Education*, *59*(2), 321-367.

Lightbown, P. & Spada, N. (2021). How Languages Are Learned. Oxford Handbooks for Teachers. Oxford University Press. 5th Edition.

Liu, K., Chen, Y., Lin, M., Huang, L.D., & Ping Xiang, C. (2023). Developing a VR-based contextualized language learning system to enhance junior high school students' pragmatic competence. *Proceedings of the 5th ACM International Conference on Multimedia in Asia*.

Liu, T., Zhang, Z., & Gao, X. (2023). Pedagogical design in technology-enhanced language education research: A scoping review. *Sustainability*, *15*(7), 6069.

Lo, C. K. (2023). What Is the Impact of AI on Education? A Rapid Review of the Literature. Education Sciences, 13(4), 1-15. https://doi.org/10.3390/educsci13040410

MacNeil, S., Tran, A., Mogil, D., Bernstein, S., Ross, E., & Huang, Z. (2022, August). Generating diverse code explanations using the gpt-3 large language model. In *Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 2* (pp. 37-39).

Marmo, R. (2023). Artificial intelligence in e-learning systems. In *Encyclopedia of data science and machine learning* (pp. 1531-1545). IGI Global.

Mayer, C. W. F., Ludwig, S., & Brandt, S. (2022). Prompt text classifications with transformer models! An exemplary introduction to prompt-based learning with large language models. *Journal of Research on Technology in Education*, *55*(1), 125–141. https://doi.org/10.1080/15391523.2022.2142872

Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education*, 57, 632–645. https://doi.org/10.1111/ejed.12527

Moore, S., Nguyen, H. A., Bier, N., Domadia, T., & Stamper, J. (2022, September). Assessing the quality of student-generated short answer questions using GPT-3. In *European conference on technology enhanced learning* (pp. 243-257). Cham: Springer International Publishing.

Murtaza, M., Ahmed, Y., Shamsi, J. A., Sherwani, F., & Usman, M. (2022). AI-based personalized e-learning systems: Issues, challenges, and solutions. *IEEE Access*, *10*, 81323-81342.

Omar, H., Owida, H., Abuowaida, S., Alshdaifat, N., Alazaidah, R., Abu Elsoud, E., & Batiha, R. (2024). AI: A New AI Tool for English Language Teaching and Learning among Jordanian Students. 10.58262/ks.v12i1.257.

Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher Education: A systematic review of empirical research from 2011 to 2020. Education and Information Technologies, 27(6), 7893–7925. https://doi.org/10.1007/s10639-022-10925-9

Punar Özçelik, N. & Ekşi, G. (2024). Cultivating writing skills: the role of AI as a learning assistant—a case study. Smart Learning Environments. 11. 10.1186/s40561-024-00296-8.

Rudolph, J., Tan, S., & Tan, S. (2023). AI: Bullshit spewer or the end of traditional assessments in higher education? Journal of Applied Learning and Teaching, 6(1), 1-22. https://doi.org/10.37074/jalt.2023.6.1.9

Santos, R. P. D. (2023). Enhancing physics learning with ChatGPT, bing chat, and bard as agents-to-think-with: a comparative case study. *arXiv preprint arXiv:2306.00724*.

Sarrion, E. (2023). What Is AI?. In. (Ed.), Exploring the Power of AI. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-9529-8_1

Schmidt-Fajlik, R. (2023). AI as a Grammar Checker for Japanese English Language Learners: A Comparison with Grammarly and ProWritingAid.

AsiaCALL Online Journal, 14(1), 105-119. https://doi.org/10.54855/acoj.231417

Schwichow, M., Brandenburger, M., & Wilbers, J. (2022). Analysis of experimental design errors in elementary school: how do students identify, interpret, and justify controlled and confounded experiments?. *International Journal of Science Education*, *44*(1), 91-114.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, *35*, 24824-24837.

Wu, X., He, X., Liu, T., Liu, N., & Zhai, X. (2023, June). Matching exemplar as next sentence prediction (mensp): Zero-shot prompt learning for automatic scoring in science education. In *International conference on artificial intelligence in education* (pp. 401-413). Cham: Springer Nature Switzerland.

Yan, D. (2023). Impact of AI on learners in an L2 writing practicum: An exploratory investigation. Education and Information Technologies, 28(11), 1-25. https://doi.org/10.1007/s10639-023-11742-4

Zhai, X., & Nehm, R. H. (2023). AI and formative assessment: The train has left the station. *Journal of Research in Science Teaching*, *60*(6), 1390-1398.