Information and Knowledge Management
ISSN 2224-5758 (Paper) ISSN 2224-896X (Online)
Vol.15, No.1, 2025

Effect of Artificial Intelligence Tools on Male and Female IELTS Test Taker's Writing Performance: The Case of Wordtune on Lexical Resource and Grammatical Range and Accuracy

Mehdi Ghaedrahmat (Corresponding author)
Assistant professor, Amin Institute of Higher Education, Fooladshahr, Isfahan

Afsaneh Alsadat Movahedpour
Islamic Azad University, Yazd Branch, Yazd, Iran

Shakila Hajian
MA student of TEFL, Shahid Ashrafi Esfahani University, Isfahan, Iran

Abstract

The present research evaluated the effects of Wordtune as an Artificial Intelligence (AI) assistance on Iranian IELTS candidates' lexical resources and grammatical range and accuracy in writing. Utilizing a quasi-experimental design, the research scrutinized quantitative data. A number of 80 male and female learners from IELTS center located in Isfahan, Iran, with intermediate language proficiency, participated in the study and were split equally to an experimental group with AI-supported instruction and a control group receiving traditional instruction exclusively. All participants underwent the same vocabulary and grammar pre-test and post-test measures with the two groups being evaluated simultaneously. The results of an ANCOVA indicated that students' lexical resources increased significantly due to regular practice with AI. However, such practice did not significantly impact grammar learning. In addition, no significant gender differences were found in the impact of AI on grammar and vocabulary resources. The study emphasized the potential of AI in education irrespective of gender differences and suggested using it alongside traditional instruction for optimal results, especially for grammar. Additionally, it proposed further enhancement of AI technologies in order for them to provide more sophisticated, grammar based feedback.

Keywords: Artificial Intelligence, Grammatical Range and Accuracy, IELTS Writing Performance, Lexical Resource, Test Improvement

DOI: 10.7176/IKM/15-1-01

Publication date: March 31st 2025

1. Introduction

The IELTS (International English Language Testing System), is a globally acknowledged standardized examination that assesses one's competency in the English language. The IELTS test is essential for assessing an individual's proficiency in English communication, which is important for studying, working, and migrating to English-speaking countries. Test takers often have difficulties in attaining their target band score, particularly in the writing component, IELTS test takers often face significant challenges in the writing module, primarily due to the demanding nature of the tasks that assess both grammatical range and lexical resources. Many struggles with effectively structuring their essays to meet the coherence and cohesion criteria, while others find it difficult to employ a wide range of vocabulary and grammatical structures without errors. These difficulties can be exacerbated by time constraints and the pressure to perform well, which can affect their overall fluency and ability to present clear, well-argued responses. (Ahmadian & Momeni, 2023).

The IELTS (International English Language Testing System) test is issued to non-native English speakers who are willing to study or work in an English dominated environment. The exam consisted of four main parts: speaking, writing, reading, and listening. Most people do not receive the target band score because they do not prepare extensively or understand the register. The level of vocabulary received during schooling does not

prepare individuals for the amount of vocabulary they will need to utilize in their writing skills. Others cannot meet the criteria regarding coherence and cohesion in essay writing task. A seasoned writer cannot compose an essay in an hour, yet these students are benchmarked against seasoned professionals. Those students are forced to give up certain basic tenets of fluent speech. (Ahmadian & Momeni, 2023).

Writing, which is responsible for 25% of the total score, is an important section in the IELTS exam. It assesses whether an individual is capable of organizing ideas, providing arguments, and utilizing a variety of vocabulary and grammar. The assessment takes into consideration two major components, which are termed as writing skills: lexical resource and grammatical variety and correctness (Farnia et al., 2024).

One variable which was investigated in the present research was lexical resources in the IELTS writing. Lexical resource is known as the test taker's proficiency in using a wide variety of terminology appropriately (Rafieyan et al., 2024). It includes employing phrases and words which are related to the issue under argument, including showing the capability to use language to communicate thoughts effectively (Clark & Yu, 2024).

For the IELTS exam, a strong lexical resource is of great significance, because it reflects one's ability in expressing him/herself clearly and effectively in the English language. Moreover, it can show the general language ability of test takers and can affect their scores in the productive language skills of the exam (Souzandehfar, 2024).

In order to reflect a high level of vocabulary proficiency on the IELTS exam, it is necessary to get engaged in practices of various vocabulary items in both speaking and writing tasks, and confirming its suitable use (Sari & Alyousef, 2024).

Developing a strong vocabulary is necessary for getting a good score in the IELTS exam. According to Nguyen (2024), practicing might enhance one's vocabulary knowledge and enables them to use it effectively during a test.

One of the challenges in meeting the requirements of lexical resources of IELTS writing is to use a wide range of vocabulary (Kashef et al, 2024). Test-takers must have a wide range of vocabulary and avoid using the same words recurrently in their work. This may be problematic for many test takers, especially English as second language learners.

Moreover, grammatical accuracy and grammatical range assesses the test taker's ability in employing different grammatical structures effectively. These factors noted by Pratama and Munandar (2024) affect the band score for the writing component of the IELTS test.

Grammar plays a substantial role in the IELTS writing exam, because it assesses a test takers' ability in employing accurate grammatical structures and terminology in written communication. The grammar in the IELTS exam is scored by the diversity of grammatical structures which have been used, in addition to developing complex sentences and using a various tenses (Alanezi & Alenezi, 2024).

Grammar plays an important role in communicating ideas (Ghaemi & Khorsand, 2024). Grammatical problems may lead to misunderstanding and block the comprehension of the writer's message (Horak & Mcculloch, 2024). Additionally, using a wide range of grammatical structures shows a developed linguistic knowledge and may lead to improvement of the whole work (Kashef et al, 2024).

One common problem about grammar is the use of complex sentences as well as punctuation. Test-takers need to be proficient about grammatical principles and need to apply them effectively in their written work. This might provide can be challenging for test takers who are not familiar with the complications of English grammar.

To address these challenges, test-takers need to practice regularly and they need to get familiar with the features of the IELTS writing test (Kashef et al, 2024). Artificial Intelligence (AI) can be considered as a possible solution in order to solve difficulties related to writing, especially in terms of grammar and vocabulary. AI has been very demanding across several domains, and its inclusion into the field of education has attracted much attention in recent years. By using AI as an assistant, both lexical resources and grammatical knowledge can be boosted among test takers (Tetty, 2024).

According to Fathi et al, (2024), AI has become a useful instrument in the field of education in general and language acquisition in particular. AI can give precise feedback, identify weak points, and suggest suitable plans for development. AI can be employed for IELTS to help test takers improve their writing proficiency, particularly in terms of lexical resource and grammatical accuracy (Li & Chan, 2024).

Using AI technology, IELTS candidates can get feedback on their writing tasks specifically in grammar and word selection. AI can investigate writing patterns and provide suggestions to help test takers improve their vocabulary and grammar proficiency. The use of AI tools may have a significant effect on the test taker's performance, which can result in better scores in the writing module of the IELTS test (Sun, 2023).

Employing AI to improve the writing performance of IELTS candidates is of significance because of many reasons. The IELTS test is known as a high-stakes test that assesses a test taker's proficiency in using English in various situations. Accordingly, it is essential for candidates to have a high level of writing proficiency. Using AI tools candidates can receive immediate feedback on their writing, which allows them to correct lexical or grammatical errors. Moreover, AI tools can provide suggestions for alternative words or phrases to enhance their

vocabulary. This can finally improve their lexical resource. AI-powered tools can detect common grammatical errors, and provide recommendations on how to correct them. IELTS candidates can check and improve the clarity and cohesiveness of their writing in the IELTS using AI technology to improve their grammatical accuracy.

According to what was said, the primary objective of this study was to explore the impact of using AI tools on enhancing the writing performance of IELTS test takers, focusing on improvements in lexical resources and grammatical range and accuracy. By integrating AI tools and technologies, the research aims to identify measurable improvements in test scores and writing quality, thereby addressing common challenges faced by examinees in these areas. Overall, the primary objective of this research was to investigate the effect of implementing artificial intelligence to improve IELTS test taker's writing performance.

2. Literature Review

The use of AI in language education has been an important field of interest for researchers in recent years (e.g., Sutrisno, 2023). Using AI can enrich learners' experiences by providing personalized process253VC M Mes (Altakhaineh & Melo-Pfeifer, 2024.; Fathi, et al. 2022). Specialists in the field of education have employed technology to make virtual learning environment that actively engage learners and enhance language acquisition (e.g., Fathi, et al., 2024; Nguyen, 2024).

AI is considered as a promising tool in language teaching and learning to improve the process of learning and teaching (Effatpanah, et al., 2024). AI tools help interpret and reply to human searches, similar to the ways which are dependent on human intelligence to provide relevant information (Fenton-Smith & Humphreys, 2017).). As an example, ChatGPT, which is an AI-powered tool, provides the information that they seek for (Fitria, 2023). AI has caused changes in different areas of life, including education and language learning (Roquet Pugès et al., 2023).

Writing needs different cognitive processes as well as their systematic activation in our cognitive framework (Olive, 2003). The role of mental processes in organizing and controlling of L2 production has been the center of attention since last decades (e.g., Kellogg, 1996). This helps us understand the contribution of these processes in the development of L2 writing skills. AI can help language teachers teach different language skills and sub-skills, one of the most complicated ones being writing proficiency.

Writing includes different cognitive mechanisms which are responsible for managing various levels of representation (Wang, et al, 2023). For example, during writing, writers have the chance to plan and to consider different issues before beginning the main writing process. The learners have time for reflection; however, this process still needs working memory to do several processes simultaneously (Olive, 2003). AI can help in handling such processes.

So far a number of studies have been carried out investigating the role of AI on writing proficiency of L2 learners (Alharbi, 2023; Cardon et al., 2023; Shi & Aryadoust, 2024; Song & Song, 2023; Yang et al., 2023; Zou et al., 2023). In this section some of the most recent ones are reviewed. In his work Sun (2024) investigated the use of generative AI in assessing ESL writing focusing on IELTS writing tasks, and proved the positive role of AI in this regard.

In a related research, Oskoui, et al. (2024) investigated the strategies that L2 writers used to do academic IELTS essay writing task. The participants included 11 L2 learners. The results showed that although the participants used various strategies to make high-quality writing tasks, none was completely dependent on the AI outputs without post-editing. They believed that the AI was a writing assistant, and they used some revised sections of AI responses and included their own words.

In addition, several studies have investigated the use of ChatGPT to address grammatical mistakes (e.g., Winans 2023) and the connection between ChatGPT and Grammarly (e.g., Yoandita & Yenni Hasnah 2024). These studies have shown positive effects of ChatGPT as an AI tool on grammatical accuracy of the L2 writing.

In another research, Hz et al. (2023) showed that the development of AI tools has enhanced the motivation among L2 writers. The researchers stated that AI writing tools can help learners compose and improve text in Indonesia. On the other hand, Hz et al. (2023) stated that the accessibility of these tools can lead to a dependence on AI in writing, which can lead to worries about the possible descent of writing skills.

Moreover, Godwin-Jones (2022), reviewed the results of related studies about AI-based writing tools. The findings showed that these technologies are useful for both students and teachers. Teachers help learners to use technology effectively. Moreover, Yan (2023) showed a significant improvement in L2 learners' writing performance as a result of ChatGPT. A review of the existing literature shows a gap in using AI tools on writing skills of Iranian IELTS candidates. Accordingly, the following research questions were posed:

RQ1: To what extent do female test takers using Wordtune as an AI tool for lexical and grammatical practice achieve higher IELTS writing scores compared to those who do not?

RQ2: To what extent do male test takers using Wordtune as an AI tools for lexical and grammatical practice achieve higher IELTS writing scores compared to those who do not?

3. Methodology

3.1 Research Design

This study employed a quasi-experimental design to analyze quantitative data, focusing on the impact of AI-assisted instruction on lexical resources, grammatical range, and accuracy. The design was selected to address the practical and logistical constraints of real-world educational settings, such as non-randomized participant selection and varying levels of pre-existing knowledge. It allowed for controlled comparisons between two groups: an experimental group receiving AI-supported instruction and a control group following traditional, instructor-led methods without AI intervention.

3.2 Participants

The participants of the research included a number of 80 male and female Iranian IELTS test takers. Their age ranged from 20 to 35, and they were chosen from the IELTS candidates registering in the IELTS center which is a famous IELTS training center in Isfahan, Iran. The participants' mother tongue was Farsi and English was their second language. In order to choose the participants, a placement procedure was carried out prior to the main treatment. A standardized mock IELTS test was employed to establish the participants' level of general proficiency. The test included four language skills. Those candidates who scored within the intermediate range were chosen for the study. Then, participants were randomly assigned to either an experimental group (N=40) in which instruction was integrated with Word tune as an AI-assisted tool. The control group (N=40) followed traditional teaching methods with feedback from teachers. Moreover, the groups were classified by gender to investigate potential gender-related differences in the effectiveness of AI-assisted instruction.

3.3 Materials and Instruments

To effectively collect and analyze data, a range of tools and resources were employed, each carefully selected to ensure the study's validity and reliability. These instruments allowed for the systematic evaluation of the participants' lexical resources, grammatical range, and accuracy before and after the intervention.

3.3.1 IELTS Writing Exam

The pre-test and post-test consisted of carefully designed writing tasks modeled after actual IELTS Writing Task 2. The reason why this task was selected was its alignment with the research objectives, assessing participants' lexical resources and grammatical proficiency in essay writing. The test was used as both pre-test and post-test to ensure that any differences could be ascribed to the treatment rather than changes in task difficulty

3.3.2 Wordtune, AI-Powered Writing Assistant

According to Zhao (2023) Wordtune is a digital writing assistant that uses AI technology to offer alternative sentence structures and synonym replacements for highlighted text, all while preserving the original meaning. As defined by Suman et al. (2023), Wordtune is a digital writing assistance built specifically for EFL authors. Using this tool to help students build or translate ideas into English can also enhance the quality of their writing. The program also helps students avoid getting stuck on difficult English words or idioms and maintain motivation while writing. Based on Fitria (2024), the following functionalities are available in Wordtune: 1. Word choice: Wordtune can offer better word suggestions based on writing style, coherence, and clarity, among other factors. 2. Grammar: Wordtune may fix errors in spelling, punctuation, and sentence construction. 3. Structural: Wordtune provides sentence themes, phrase layout, and paragraph division. 4. Style: Wordtune can offer advice on how to make writing more engaging, reduce clichés, and increase language proficiency. Wordtune is available online as an editor that runs on the web or as an extension for browsers. The AI-powered assistant allowed participants to refine their work iteratively, fostering active learning and autonomy in the experimental group.

Longman Academic Writing Series 3 (Hogue 2013) was employed as the course book in the present research. The first part of this book provides information on how to develop various types of academic paragraphs. Various paragraph types, including process, definition, narrative, cause /effect, as well as comparison/contrast paragraphs, are taught. The second part provides information about the basic concepts of essay writing. The last part includes information about, appendices, and an index which make the text an easy-to-use reference tool.

3.4 Procedures

The data collection process was organized into three distinct phases to ensure a systematic and comprehensive evaluation of the instructional interventions. Below is a detailed clarification of how the experimental group received the treatment:

3.4.1 Phase 1: Pre-Test

At the beginning, all participants sat for a pre-test to have a benchmark for their lexical and grammatical proficiency. The pre-test evaluated writing an essay prompt, the ability to develop logical arguments and the ability to employ advanced grammatical structures. The test provided data to compare the experimental and

control groups and identify specific areas for improvement.

3.4.2 Phase 2: Treatment

The treatment lasted eight weeks, during which the experimental group was exposed to AI-assisted instruction and the control group following a traditional method of teaching.

3.4.3 Experimental Group

AI Tool Integration:

Participants were exposed to an AI-powered writing tool, namely, Wordtune, to get feedback on their writing.
Wordtune provided suggestions for enhancing lexical enrichment of the writing such as synonyms. In addition, Wordtune provided suggestions on grammatical accuracy such as sentence structure and punctuation.

Guided Practice Sessions:

Participants engaged in structured writing tasks (e.g., essays, summaries) with AI feedback integrated into the process.
They revised their work iteratively based on AI suggestions, focusing on error correction and vocabulary enhancement.

Interactive Exercises:

Participants completed interactive exercises designed to reinforce learning, such as vocabulary-building games and grammar quizzes, all supported by AI.
They were encouraged to reflect on their progress and identify recurring errors using AI-generated analytics.

Session Structure:

Each session lasted 90 min and included:
Theoretical Instruction: Brief explanations of key concepts (e.g., grammar rules, lexical strategies).
Guided Practice: Writing tasks with AI feedback.
Feedback Review: Discussion of AI-generated feedback and strategies for improvement.

3.4.4 Control Group

Traditional Curriculum:

Participants were exposed to the curriculum by an instructor while receiving:
Grammatical Exercises: Practices about specific grammatical structures.
Vocabulary Drills: Vocabulary and phrase list to memorize.
Textbook-Based Writing Exercise: Essay writing tasks and summary tasks with no AI assistance.
Instructor Feedback:
Participants received feedback from the instructor during class discussions and homework evaluations.
Feedback was provided manually, focusing on general areas for improvement rather than real-time, detailed corrections.

Phase 3: Post-Test

At the end of the treatment period, all participants completed the same test as the pre-test to measure improvements in their lexical and grammatical proficiency. The post-test results were analyzed to determine:
The effectiveness of AI-assisted instruction in enhancing writing skills.
The comparative performance of the experimental and control groups.
Gender-based differences about the instructional methods.
The experimental group was exposed to Wordtune, while the control group was exposed to traditional teaching methods.

3.4.5 Scoring Rubrics

Explicit scoring rubrics were utilized to objectively evaluate participants' writing performance. These rubrics, aligned with IELTS Writing band descriptors, included clear, quantifiable criteria for lexical resource, grammatical range and accuracy, coherence and cohesion, and task achievement.

3. Results

1.1 Research Design

To examine the research variables from the respondents' perspectives, the following results were obtained, representing key statistical parameters. Below the Table1 shows the mean values for grammar learning and lexical resources for the experimental and control groups:

Table 1. *Mean Values for Grammar Learning*

Group	Variable	Mean	Std. Deviation	Minimum	Maximum
Control	Pre-Grammar	3.0938	0.5569	3.0	5.0
Experimental	Pre-Grammar	3.0938	0.4265	3.0	5.0
Control	Post-Grammar	3.0963	0.4143	3.5	5.0
Experimental	Post-Grammar	4.0088	0.3904	3.5	4.5
Control	Pre-Lexical	4.0438	0.5212	3.5	5.0
Experimental	Pre-Lexical	4.0325	0.5006	3.5	5.0
Control	Post-Lexical	4.0425	0.4319	3.5	5.0
Experimental	Post-Lexical	5.0475	0.4797	4.5	6.0

The post-test means for grammar learning in the experimental group was 4.088, compared to 3.963 for the control group. The post-test mean for lexical resources in the experimental group was 5.475, whereas for the control group it was 4.425.

Inferential Statistics and Hypothesis Testing

After describing the responses from the statistical sample, this section addresses the hypotheses and statistical tests employed in the research. The field data analysis aims to statistically verify the hypotheses. Covariance analysis was used for inferential statistics in this study.

Analysis of covariance (ANCOVA) is a statistical method that enables examination of the effect of an independent variable on a dependent variable while controlling for the impact of other variables. Prior to conducting ANCOVA, its assumptions are tested.

Testing Data Normality Using Skewness and Kurtosis

First, skewness and kurtosis values were examined.

Table 2. *Skewness and Kurtosis Values for The Distribution of Variables*

Group	Variable	N	Skewness	Kurtosis
Control	Pre-Grammar	40	-0.211	-0.153
Experimental	Pre-Grammar	40	-0.795	1.794
Control	Post-Grammar	40	0.714	0.220
Experimental	Post-Grammar	40	-0.323	-1.267
Control	Pre-Lexical	40	-0.454	-0.978
Experimental	Pre-Lexical	40	0.128	-1.182
Control	Post-Lexical	40	-0.200	-0.710
Experimental	Post-Lexical	40	-0.446	-0.826

The observed skewness values for the variables were within the range of (-2, 2), which showed that the distribution of variables were symmetrical and met the normality criterion. Likewise, the kurtosis values also were in the range of (-2, 2) which suggested a normal distribution. Accordingly, all variables were normally distributed.

Homogeneity of Regression Slopes

In ANOVA, it is assumed that the variables should exhibit linearity across all data. Additionally, it is necessary that regression lines be homogeneous across groups. If the regression lines are not homogeneous, ANOVA may not be an appropriate analysis. Homogeneity of regression slopes is a critical consideration in ANOVA.

In this study, post-test scores for grammar learning and lexical resources were treated as dependent variables, with their respective pre-test scores as covariates. Homogeneity of slopes will be satisfied if there is equality between covariates (pre-tests) and dependent variables (post-tests) across all factor levels (experimental

and control groups). A non-significant interaction between the dependent and covariate variables indicates homogeneity.

Table 3. *Results of the Homogeneity of Regression Slopes Test for Research Variables Across Both Groups:*

Variable	Source of Variance	Stage: Pre-Test — Post-Test	F	Sig.
Grammar Learning	Group * Pre-Test		0.357	0.552
Lexical Resources	Interaction		0.071	0.791

According to Table 3, the level of significance for the interaction terms (Group * Pre-Test) is higher than 0.05. Furthermore, the F-values for the regression slope homogeneity test were not significant. This confirms that the homogeneity of regression slopes assumption is met for grammar learning and lexical resources, allowing for ANCOVA analysis on these variables.

As shown in Table 3, the significance level for the interaction terms (Group * Pre-Test) is greater than 0.05, and the F-values for interaction are non-significant for the regression slope homogeneity across variables. This indicates that the homogeneity of regression slopes is maintained for both grammar learning and lexical resources.

Homogeneity of Variances

Levene's test was applied to examine homogeneity of variances across different samples, which tests the equality of the dependent variable variances for groups determined by categorical factors.

Table 4. *Results of Levene's Test for Variance Homogeneity Across Groups*

Variable	F	df1	df2	Sig.
Grammar Learning	1.193	1	78	0.278
Lexical Resources	0.187	1	78	0.666

As reported in Table 4, Levene's test indicates that the assumption of equal variances is met for both grammar learning and lexical resources variables. Consequently, ANCOVA can be reliably used to test the research hypotheses.

Hypothesis Testing

Given the fulfilled assumptions, the research hypotheses were tested using ANCOVA.

Table 5. *Results for the Interaction Between Groups Using Univariate ANOVA for the Grammar Learning Variable*

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power
Pre-Grammar	5.610	1	5.610	61.467	0.000	0.444	1.000
Group	0.313	1	0.313	3.424	0.068	0.043	0.447
Error	7.028	77	0.091				
Total	1309.000	80					

In Table 5, the F-ratio for the group factor, with degrees of freedom of 1 and 80, is smaller than the critical F value at the 95% confidence level ($\alpha = 0.05$). Additionally, the significance level is greater than 0.05. This indicates that, when controlling for pre-test scores, there is no significant difference in grammar learning between the experimental and control groups. Therefore, the research hypothesis is not supported, meaning that regular AI-assisted practice sessions did not have a significant impact on grammar learning, despite an increase in the mean scores.

To test Hypothesis 2, univariate analysis of covariance (ANOVA) was used.

Table 6. *Results for the Interaction between Groups Using Univariate ANCOVA for the Lexical Resources Variable*

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power
Pre-Lexical	2.003	1	2.003	10.826	0.002	0.123	0.901
Group	23.267	1	23.267	125.753	0.0001	0.620	1.000
Error	14.247	77	0.185				
Total	1998.500	80					

In Table 6, the F-ratio for the group factor, with degrees of freedom of 1 and 80, is greater than the critical F value at the 95% confidence level ($\alpha = 0.05$), and the significance level is less than 0.05. This indicates a significant difference in lexical resources between the experimental and control groups when controlling for pre-test scores. Therefore, Hypothesis 2 is supported, meaning that regular AI-assisted practice sessions significantly increase the level of lexical resources.

Hypothesis 3: The effect of regular AI-assisted practice sessions on grammar does not differ significantly between male and female participants. Given the normal distribution of the grammar variable, an independent samples t-test was used to test this hypothesis.

Table 7. *The Mean Scores for The Grammar Variable by Gender*

Group	Gender	N	Mean	Std. Deviation
Control	Female	20	4.000	0.3627
	Male	20	3.925	0.4667
Experimental	Female	20	4.175	0.3354
	Male	20	4.000	0.4292

Since the t-test determines if there is a significant gender difference in the impact of AI-assisted sessions on grammar learning, further analysis of the results will confirm if the hypothesis is supported or rejected based on the obtained p-value.

Table 8. *Independent Samples T-Test Results for Grammar Variable between ale and Female Participants*

Group	Levene's Test for Equality of Variances	T-Test for Equality of Means
	F	Sig.
Control	1.001	0.323
Experimental	0.851	0.362

As shown in Table 8, the significance level for Levene's test for the grammar variable is greater than 0.05, indicating equal variances between groups. The results of the T-test show that the mean grammar scores between male and female participants in the control and experimental groups are not significantly different. Thus, the research hypothesis is rejected, implying that the impact of regular AI-assisted practice sessions on grammar does not significantly differ between males and females, even though females have a slightly higher mean grammar learning score.

Hypothesis 4: The effect of regular AI-assisted practice sessions on lexical resources does not differ significantly between male and female participants. Given the normal distribution of the lexical resource's variable, an independent samples T-test was used to test this hypothesis.

Table 9. *Mean Scores for Lexical Resources by Gender*

Group	Gender	N	Mean	Std. Deviation
Control	Female	20	4.450	0.4261
	Male	20	4.400	0.4472
Experimental	Female	20	5.550	0.3940
	Male	20	5.400	0.5525

Table 10. *Independent Samples T-Test Results for Lexical Resources Variable*

Group	Levene's Test for Equality of Variances	T-Test for Equality of Means
	F	Sig.
Control	**0.386**	**0.538**
Experimental	3.889	0.056

As shown in Table 10, the significance level of Levene's test for the lexical resource's variable is greater than 0.05, indicating that variances are equal across groups. The T-test results show that the mean lexical resources scores between male and female participants in the control and experimental groups are not significantly different. Therefore, the research hypothesis is rejected, meaning that the impact of regular AI-assisted practice sessions on lexical resources does not differ significantly between males and females, although females have a slightly higher mean score for lexical resources.

5. Discussion

The objective of the present study was to investigate the impact of an AI-assisted tool, namely, Wordtune on the grammar and lexical resources of IELTS candidates, focusing on the potential effect of male and female participants. The results showed that while Wordtune as an AI-assisted tool could significantly improve lexical resources, it did not show a significant positive effect on grammar learning. Additionally, the findings showed no significant gender-based differences were found in the impact of AI practice on either lexical resources or grammar.

These results of the present research about the positive effects of Wordtune on lexical knowledge of IELTS candidates align with previous studies on AI in language learning, such as those by Chen et al. (2024) and Sutrisno (2023), which revealed the efficacy of AI-assisted tools in improving language skills, specifically writing. However, these studies note that AI tools are more effective in fields which need analytical feedback, such as vocabulary and sentence structure, rather than grammar which may rely on deeper understanding. This inconsistency can be explained by AI's limitations in providing detailed feedback on grammatical structures, which can be explained more effectively by traditional instruction. There are several educational theories that support the use of AI in teaching vocabulary and specifically vocabulary in writing. The first theory that supports the use of AI is Constructivist Theory. According to this theory, which was proposed by Jean Piaget in 1930s, learners construct their understanding of the world from the perspective of their experiences and reflecting on them. AI tools can provide personalized learning experiences and are consistent with the personal needs of learners and help learners build their vocabulary through appealing activities. The next theory which supports the results of the present research is the Sociocultural Theory by Lev Vygotsky in 1960s which emphasizes the importance of social interaction in learning. AI can improve collaborative learning environments in which learners can interact with AI teachers or peers to enhance their vocabulary and writing skills.

Furthermore, the results of the present research are supported by Behaviorist Theory by B.F. Skinner in 1930s and 1940s which claimed that learning is because of conditioning. AI-assisted teaching can provide learners with immediate feedback and reinforcement which helps them learn new writing skills through practice and repetition. Moreover, Cognitive Load Theory supports the results of the present research. Drawn from this theory, developed by John Sweller in late 1980s, when cognitive load is managed properly, learning happens more effectively. AI tools can reduce cognitive load by reducing complex writing tasks into manageable chunks and through providing practices.

Another theory which supports the results of the present study is Differentiated Instruction approach

advocated by Carol Ann Tomlinson in early 2000s, which focuses tailoring instruction to meet different needs of students. AI tools can analyze the performance of learners and they can provide personalized vocabulary and writing lessons which help individual learning styles and levels. Furthermore, Multimodal Learning approach developed by Richard Mayer in the late 1990s and early 2000s supports the results of the present research. This approach focuses on the advantages of presenting information through multiple modes including visual, auditory, and kinesthetic. AI tools can provide multimodal vocabulary learning experiences through text, images, audio, and interactive exercises which are demanding for different learning preferences and enhance memory retention.

Finally, Spaced Repetition and Retrieval Practice theories lend support to the findings of the present research. These theories focus on distributed practice and active recall for long-term retention of information. AI tools can plan vocabulary practice based on spaced repetition algorithms and make quizzes that enhance retrieval practice, help learners effectively acquire and maintain new vocabulary.

The growth in lexical resources which was observed in the study is in line with findings of the study by Hsu et al. (2023), which reported significant vocabulary improvement among EFL learners who were exposed to AI-assisted instruction. In the context of this study, participants who received AI support in lexical resources demonstrated higher scores than those who did not, this shows the value of AI tools in expanding vocabulary range. This supports Carpio Cañada et al. (2015) and Ebadi and Amini (2022), who revealed that AI-assisted tools can improve learners' motivation significantly in targeted language skills by providing them with timely, personalized, and varied feedback.

The lack of gender differences in the effect of AI-assisted sessions on both grammar and lexical resources echoes the findings of the studies by Ali et al. (2023) and Kim (2019), which showed that AI-assisted learning and teaching did not favor one gender over another.

One important aspect of this study is the contradiction between AI effectiveness for lexical development and grammar learning. Unlike lexical resource acquisition, grammatical knowledge did not significantly benefit from AI assistance, a finding which lends support to the studies by Chen et al. (2024) and Hz et al. (2023) which noted similar limitations in AI's ability to provide comprehensive grammatical feedback. Considering the fact that grammar often includes understanding complex syntactic and semantic rules, it may be more problematic for AI tools to completely provide the explanations which are usually provided by human instructors.

In addition, while the AI tool employed in this study may have been effective in providing immediate feedback, it may not have had the depth needed to help learners mastering complex grammatical structures. These findings suggest that although AI can support specific aspects of language learning effectively, it also has some limitations in dealing with language rules. Therefore, reinforcing AI tools with explanations from teachers can offer a balanced approach to language learning which allows learners to benefit from AI's positive points in vocabulary and immediate feedback while also obtaining detailed guidance on grammar from instructors.

6. Conclusion

This study backs the existing body of research on the role of AI in language learning, especially considering its efficiency in enhancing vocabulary and lexical resources. The findings emphasize the capability of AI assisted learning and teaching in vocabulary development and its restriction in grammar instruction. The lack of significant gender differences displays AI as a gender-neutral tool, and this shows applicability of AI in different classroom settings.

This result supports the idea that AI-assisted learning provides a gender-neutral platform, and offers equal support to male and female learners. These findings have important implications for educators, since they focus on AI tools as a potential addition to language learning environments, which benefits diverse learners identically.

Another implication of the present research is that although AI tools have valuable support for language learning, such tools should not be employed in isolation. Combining AI with traditional teaching methods can be more effective especially in areas where AI is less effective, such as grammar. Through AI integration into a balanced instructional model, educators can provide a comprehensive, technology-enhanced learning experience that enhances AI's positive side while addressing its drawbacks.

Finally, the use of AI in language learning has considerable promise, but considerate integration is necessary. Educators and developers should keep refining AI tools and teaching strategies to make sure that learners take advantage from a comprehensive educational experience. The present research provides a basis for further exploration into the use of AI in language education, encouraging future research that can expand our understanding of AI's role in enhancing language learning outcomes.

Acknowledgments

The authors would like to express their heartfelt appreciation to the contributors and participants whose assistance was vital to this research. They are also grateful to the reviewers and editorial board for their valuable contributions and feedback on the manuscript.

References

Ahmadian, M., & Momeni, G. (2023). The Influence of Lexically-Based Language Teaching on Task Response and Coherence and Cohesion Aspects of Writing Proficiency of Iranian University EFL Learners. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills), 42*(1), 61-90.

Alanezi, M. A., & Alenezi, A. A. (2024). The Impact of the IELTS Writing Test on Postgraduate Students in Kuwait. *International Journal of English Language Teaching*, *12*(2), 42-51. https://doi.org/10.37745/ijelt.13/vol12n24251

Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., and Mohammed, A. A. (2023). Impact of chat GPT on learning motivation: teachers and students' voices. J. English Stud. Arabia Felix 2, 41–49. https://doi.org/10.56540/jesaf.v2i1.51

Altakhaineh, A. R. M., & Melo-Pfeifer, S. (2024). "This topic was inconsiderate of our culture": Jordanian students' perceptions of intercultural clashes in IELTS writing tests. *Applied Linguistics Review*, *15*(1), 263-286.

Carpio Cañada, J., Mateo Sanguino, T. J., Merelo Guervós, J. J., & Rivas Santos, V. M. (2015). Open classroom: enhancing student achievement on artificial intelligence through an international online competition. *J. Comput. Assist. Learn. 31*, 14–31. https://doi.org/10.1111/jcal.12075.

Chekol, A., Shiferie, K., & Teshome, S. (2023). Effects of using adjunct model of content-based instruction on students' technical report writing performance. *Cogent Education*, *10*(1), https://doi.org/10.1080/2331186X.2023.2214427.

Chen, B., Bao, L., Zhang, R., Zhang, J., Liu, F., Wang, S., & Li, M. (2024). A multi-strategy computer-assisted EFL writing learning system with deep learning incorporated and its effects on learning: A writing feedback perspective. *Journal of Educational Computing Research*, *61*(8), 60-102.

Clark, T., & Yu, G. (2024). The pedagogical remit of test preparation: the case of writing acquisition on an IELTS course. *Applied Linguistics Review*, *15*(3), 1009-1045.

Ebadi, S., & Amini, A. (2022). Examining the roles of social presence and human likeness on Iranian EFL learners' motivation using artificial intelligence technology: a case of CSIEC chatbot. *Interact. Learn. Environ. 2022*, 1–19. doi: 10.1080/10494820.2022.2096638.

Effatpanah, F., Baghaei, P., & Karimi, M. N. (2024). A mixed Rasch model analysis of multiple profiles in L2 writing. *Assessing Writing*, *59*, 34-62. https://doi.org/10.1016/j.asw.2023.100803

Lai, Y., Saab, N., & Admiraal, W. (2022). University students' use of mobile technology in self-directed language learning: using the integrative model of behavior prediction. *Comput. Educ. 179*, 104-113. doi: 10.1016/j.compedu.2021.104413

Fitria, T. N. (2024). Using Wordtune as an Ai-Powered Writing Tool How is the Performance. *Review and Rephrase English Writing, 6*(1), 117–149.

Farnia, S., Fatehirad, N., & Shahabi, H. (2024). EFL Learners' Training on IELTS Writing Skills and Their Possible Self Construction. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills)*, *43*(2), 111-132. . https://doi.org/10.22099/tesl.2024.48631.3238

Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, *121*, 103254. https://doi.org/10.1016/j.system.2024.103254

Fenton-Smith, B., & Humphreys, P. (2017). Language specialists' views on the academic language and learning abilities of English as an additional language postgraduate coursework student: towards an adjunct tutorial model. *Higher Education Research & Development*, *36*(2), 280-296. https://doi.org/10.1080/07294360.2016.1208152

Ghaemi, H., & Khorsand, Z. (2024). A Comparative Study of Writing Scores in IELTS and PTE: An Investigation into the Factors Leading to Disparity in Scores. *African Educational Research Journal*, *12*(1), 36-52.

Horak, T., & Mcculloch, S. (2024). Mind the gap (s): exploring university teaching staff's perceptions of IELTS writing versus university writing requirements. *Journal of English Education and Teaching*, *7*(3), 480-498.

Hsu, T., Chang Ch. & Tien-Hsiu J. (2023). Artificial Intelligence image recognition using self-regulation learning strategies: effects on vocabulary acquisition, learning anxiety, and learning behaviors of English language learners. *Interactive Learning Environments 32*(6). 3060–3078 https://doi.org/10.1080/10494820.2023.2165508

Hz, B. I. R., Laiya, R. E., & Sarumaha, M. S. (2023). Navigating English Writing Proficiency Tests in the Era of Artificial Intelligence. *Journal of English Education and Teaching*, *7*(3), 480-498. https://doi.org/10.33369/jeet.7.3.480-498

Kashef, L., Ghafoori, N., Oghani, A. V., & Mehrpouyan, A. (2024). The Effect of Interactionist vs. Interventionist Dynamic Assessment on Writing Accuracy of Young vs. Adult IELTS Candidates. *International Journal of Language Testing*, *14*(1), 1-16. https://doi.org/10.22034/ijlt.2023.390643.1240

Kim, N. Y., Cha, Y., & Kim, H. S. (2019). Future English learning: Chatbots and artificial intelligence. *Multimedia Assisted Lang. Learn. 22*, 32–53.

Li, Q., & Chan, K. K. (2024). Test takers' attitudes of using exam-oriented mobile application as a tool to adapt in a high-stakes speaking test. *Education and Information Technologies*, *29*(1), 219-237. https://doi.org/10.1007/s10639-023-11678-4

Nguyen, T. T. (2024). A Contrastive Study of Theme Types in Academic IELTS Essays Written by English-Majored Students and IELTS Candidates. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *3*(2), 30-42.

Oskoui, K.; Mirzaeian, V. R. & Nafissi, Z. (2024). AI-assisted EAP Testing: A Case of Academic IELTS Writing by Iranian EFL Learners. Journal of English Language Teaching and Learning, 16(34), 307-330. https://doi.org/10.22034/elt.2024.63345.2691

Piaget, J. (1930). *The child's conception of physical causality*. Kegan Paul, Trench, Trubner & Co.

Pratama, M. A., & Munandar, A. (2024). Decoding IELTS writing performance: Grammatical Clauses mastery by Universitas Ahmad Dahlan students. *NOTION: Journal of Linguistics, Literature, & Culture*, *6*(1), 34-56.

Rafieyan, M., Salahshour, F., & Davatgari Asl, H. (2024). The Utility of Lexical Bundle Teaching for Improving ESP Learners' Writing Ability on IELTS. *Journal of English Language Teaching and Learning*, *16*(33), 278-297. https://doi.org/10.22034/ELT.2024.59447.2588

Roquet Pugès, H., Navarro Gil, N., & Nicolás-Conesa, F. (2024). The impact of adjunct instruction on EFL academic writing at university. *Journal of Immersion and Content-Based Language Education*. https://doi.org/10.1075/jicb.22005.roq

Sari, A., & Alyousef, H. (2024). The Use of Engagement Resources in GCSE EFL And ESL Reading Exams: A Practical Application Of Appraisal Theory [In English]. *Journal of Studies in Applied Language (JSAL)*, *7*(1), 1-22.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. Appleton-Century.

Souzandehfar, M. (2024). New Perspectives on IELTS Authenticity: An Evaluation of the Speaking Module. *International Journal of Language Testing*, *14*(1), 34-55. https://doi.org/10.22034/ijlt.2024.48631.3238

Suman, R. S., Moccia, S., Chinnusamy, K., Singh, B., & Regin, R. (Eds.). (2023). *Handbook of research on learning in language classrooms through ICT-based digital technology*. IGI Global. https://doi.org/10.4018/978-1-6684-6682-7

Sun, T. (2023). The Potential Use of Artificial Intelligence in ESL Writing Assessment: A Case Study of IELTS Writing Tasks. *Irish Journal of Technology Enhanced Learning*, *7*(2), 42-51.

Sutrisno, D. B. (2023). Practicing IELTS Writing for L2 Writers with ChatGPT; An Exploratory Self-Study. *EDU Journal-English Department of UMMU Journal*, *3*(2), 41-55.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257-285. https://doi.org/10.1207/s15516709cog1202_4

Tetty, M. (2024, January). Writing Essay Smoothly by Using IELTS Strategies. In *Proceedings of the 5th International Conference on Innovation in Education, Science, and Culture, ICIESC 2023, 24 October 2023, Medan, Indonesia*.

Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms*. Upper Saddle River, NJ: Pearson Education

Ueda, D., Kakinuma, T., Fujita, S., Kamagata, K., Fushimi, Y., Ito, R., ... & Naganawa, S. (2024). Fairness of artificial intelligence in healthcare: review and recommendations. *Japanese Journal of Radiology*, *42*(1), 3-15. https://doi.org/10.1007/s11604-023-01600-8

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, S., Christensen, C., Cui, W., Tong, R., Yarnall, L., Shear, L., & Feng, M. (2023). When adaptive learning is effective learning: comparison of an adaptive learning system to teacher-led instruction. *Interactive Learning Environments*, *31*(2), 793-803. https://doi.org/10.1080/10494820.2020.1808794

Winans, M. D. (2021). Grammarly's Tone Detector: Helping Students Write Pragmatically Appropriate Texts. *RELC Journal, 52*(2), 348–352. https://doi.org/10.1177 /00336 882 211 010506

Yoandita, & Yenni Hasnah. (2024). Quillbot as an Alternative Writing Tool: Examining Its Uses on the Academic Writing Performance of Efl Learners. *Esteem Journal of English Education Study Programme, 7*(2), 401–412. https://doi.org/10.31851/esteem.v7i2.15254

Zhao, X. (2023). Leveraging Artificial Intelligence (AI) Technology for English Writing: Introducing Wordtune as a Digital Writing Assistant for EFL Writers. *RELC Journal, 54*(3), 890–894. https://doi.org/10.1177/00336882231112345