ISSN 1798-4769
Journal of Language Teaching and Research, Vol. 14, No. 5, pp. 1395-1404, September 2023
DOI: https://doi.org/10.17507/jltr.1405.28

Investigating EFL Students' Writing Skills Through Artificial Intelligence: Wordtune Application as a Tool

Fawaz Al Mahmud
English Language and Translation Department, Faculty of Sciences and Arts, Jeddah University, Khulais, Saudi Arabia

***Abstract*—Artificial intelligence (AI) powered writing technologies such as Wordtune and Grammarly are increasingly penetrating the L2 writing domain. Despite the growing significance of such digital tools, few studies have explored how AI-driven applications impact Saudi English as a Foreign Language (EFL) writing. The current study aimed to fill this gap by investigating whether and to what extent Wordtune facilitates Saudi students' writing. The participants were divided into two male and two female groups using a mixed-method design. For both male and female samples, one group was identified as the control group, whereas the other one was identified as the experimental group. Quantitative data were collected using pretests and post-tests and analyzed using SPSS. Qualitative data were derived from students' writing samples and assessed by two human raters. The results showed that using Wordtune, students in the experimental group surpassed those in the control group. Quantitative findings demonstrated that using Wordtune, the experimental group upgraded their writing and outperformed the control group in the final writing exam. Qualitative findings indicate that using Wordtune, the experimental group made modest writing gains at the lexical and syntactic levels. While lexical gains included more concrete nouns, vivid adjectives, and precise verbs, sentence-structure gains included the increasing presence of complex phrases and complex and compound sentences. Finally, the impact of Wordtune on writing quality was the same for both male and female participants.**

***Index Terms*—artificial intelligence, writing skills, Wordtune, EFL students, English teaching**

I. Introduction

Effective writing is valuable in academic and professional environments (Lahuerta, 2017; McDonough & Crawford, 2018; Rosário, 2019). Through this process, writers showcase and synthesize knowledge, advance and evaluate arguments, and clarify and shape opinions (Zhou & Hiver, 2022). A well-written paper or proposal engages with, enlightens, and persuades the reader much more than a poorly written piece on the same subject. For English as a Foreign Language (EFL) learners, writing well is a prime predictor of academic prowess at the college level (Geiser & Studley, 2001). However, written academic English presents a challenge for students (Hamzaoui, 2021; Ruscetti et al., 2018) and more so for EFL learners (Alsied & Ibrahim, 2017; Hanauer et al., 2019). Academic writing barriers for English as an Additional Language (EAL) learners are many and varied: grammar, lexis, and mechanics of writing (Phuong, 2021; Singh, 2017), information structure (Flowerdew, 2019; Komba, 2015; Mahammoda, 2016), and genre conventions (Bitchener & Basturkmen, 2006; Finn, 2018).

Research has shown that digital writing tools powered by Artificial Intelligence (AI) can raise the level of learners' written English (Nobles & Paganucci, 2015; Dickson, 2017; Azah, 2019; Karyuatri, 2018; Coenen et al., 2021). Using AI, we refer to a sub-discipline within computer science focused on creating systems that can imitate the intelligence of human beings. The term "Artificial Intelligence" was first used by John McCarthy in 1956 to designate basic types of thinking machines (McCarthy et al., 2006). Since then, AI has been used to achieve goals in various fields, including education (Lin & Mubarok, 2021; Moussalli & Cardoso, 2020).

An example of AI in education is the advent of AI-powered digital writing assistants. These online tools are used worldwide to enhance writing quality (Zheng & Warschauer, 2017). However, currently available digital writing tools mostly focus on helping writers edit drafts for grammatical correctness (Winans, 2021), and only a few online writing applications assist users across the full spectrum of the writing process. One such application is Wordtune.

Wordtune is an AI-driven software that facilitates assists learners with their writing by allowing them to paraphrase their original text. When a user inputs a sentence or a group of sentences into Wordtune, this digital writing tool rewrites the original sentences by replacing words and modifying the sentence structure (Zhao, 2022).

Given that an increasing number of EFL students are using digital writing assistants, such as online dictionaries, thesaurus, Grammarly, and Wordtune, to enhance their written communication, investigating the degree to which such technologies positively impact the quality of English writing is warranted, especially in the EFL context, where students encounter more substantial challenges with writing than their L1 counterparts (Lin & Morrison, 2021; Hanauer et al., 2019).

Although a sizable body of research focuses on the role of lexicographic and correctness-focused digital writing tools in EFL writing, only a few studies have explored the extent to which a broad-spectrum digital writing tool, such as Wordtune, facilitates EFL writing students. Moreover, few studies have foregrounded the effects of a broad-based digital writing assistant, such as Wordtune, on Saudi EFL students' writing skills. The aim of the current study is to address this gap in the literature.

Quantitative data were collected using pretest and posttest, and qualitative data were derived from the textual analysis of students' writing samples. The following research questions formed the focus of the current study.

1. To what extent does Wordtune enhance student writing?
2. What are the specific improvements in Wordtune-mediated student writing?
3. Are there any statistical differences in students' writing skills based on their sex?

This study will make a meaningful contribution to the field of technology-enabled writing by foregrounding the efficacy of Wordtune in EFL writing, with specific reference to Saudi students. A deeper understanding of how Wordtune aids EFL writers will contribute to designing instructional strategies for Saudi EFL writers. This research is needed and timely for two major reasons. First, it determined the role of Wordtune in the writing of a specific subset of the EFL student community, namely, the Saudi EFL learners. This identifies for the stakeholders not only the aspects of EFL writing that Wordtune can change for the better but also the linguistic areas the students struggle with. Second, the conclusions drawn from this study point to potential areas for further studies on the role of AI-driven writing tools in EFL writing.

The remainder of this article is structured thus: first, under the section 'Literature Review' I discuss the relevant past research on the utility of digital writing assistants for school and college learners and the position of Wordtune in the environment of EFL writing. Next, under the 'Methodology' section, the participants, research tools and instruments of data analysis are reported. Then, the findings are reported and discussed. Afterwards, recommendations are made for stakeholders and the limitations of this research are noted. The paper concludes by reiterating the core findings of the current study.

II. Literature Review

Digital writing is an umbrella term for various related activities such as writing on Facebook and other social networking sites, blog writing, writing with a word processor, and other forms of online communication (Zhao, 2022). Digital writing assistants are online applications that help users with various aspects of writing. Digital writing tools include lexicographic applications, such as online dictionaries and thesauruses, grammar-focused tools such as Grammarly, and writing-process-focused tools, such as QuillBot and Wordtune.

A. Digital Writing Assistants and Academic Writing

Prior research has investigated the impact of digital writing assistants on students' writing quality (AbdAlgane & Othman, 2023; Choo et al., 2017; Gayed et al., 2022). Moore et al. (2016) studied whether and to what extent digital writing tools improved the writing quality of college EFL students in Canada. Their findings revealed that while digital writing software could enhance writing skills, it alone was not sufficient to make a marked difference. Their qualitative findings underscored the significance and primacy of strategy instruction in face-to-face classrooms. In another study, Perry (2021) analyzed the extant literature on digital writing assistants and concluded that, in the L2 context, these online tools were most effective when embedded in judiciously designed writing programs. While the studies cited above addressed the question of whether digital writing tools facilitated L2 writing, these works did not discuss whether writing enhancements remained with participants in the long run. Hamouma and Menezla (2019) studied 80 EFL students and found that students proficient in the use of digital writing assistants were more likely to develop written English than those who were not well-acquainted with the technical features of digital writing tools. In another investigation, Purcell et al. (2013) surveyed 2,462 teachers to ascertain how digital writing tools impacted students' texts. They discovered that the recent digital writing platforms, such as Google Docs, with their advanced features, could improve students' writing.

The following key inferences may be drawn from the above survey of the role of digital writing tools in the context of academic written English: first, by and large, digital writing tools help students improve their written English, and second, digital writing assistants work best for the learners with an advanced understanding of digital writing technology.

B. Wordtune as a Digital Writing Assistant

Driven by AI, Wordtune is an online writing tool that allows learners to paraphrase their writing. It does so by replacing the words in the original text with their synonyms and varying the sentence structure at the phrasal and clausal levels. Based on the natural language processing technique, Wordtune manipulates a large database of written text to rewrite sentences. This digital writing tool uses AI to determine patterns learned from a writing database to produce multiple rewrites of the original sentences (Zhao, 2022). Wordtune does not lift expressions verbatim from already available online material in producing a paraphrase.

Wordtune-produced paraphrases help the learner in two important ways: one, the learners begin to see what they intend to say, and two, as pointed out by Barrot (2020), when learners notice the qualitative difference between their text and the usually better paraphrases produced by Wordtune, they internalize different linguistic options available to shape a thought into a sentence, thereby enhancing their written English writing through self-regulated learning.

Wordtune serves users with different levels of writing proficiency. For instance, Wordtune provides beginners with the ability to translate phrases from foreign languages into English. In this manner, a user can even have an English paraphrase of the original text containing a few non-English phrases. Wordtune can help more skilled L2 learners polish their prose by providing multiple rewrites. By comparing and considering these rewrites, intermediate users can choose the correct words and clear sentence structures for their prose (Fitria, 2021). Additionally, Wordtune provides users with the option of compressing or expanding original sentences. Users can also compress or expand the original text using Wordtune (Zhang et al., 2022).

III. Methodology

A. Research Design

This study employed a mixed-methods research design (Creswell, 2013). Quantitative data were collected first. To that end, pretest and posttest were administered to the control and the experimental groups. Next, qualitative data were collected through a textual analysis of some participants' writing samples to determine the impact of Wordtune on their writing quality, substantiating thereby the quantitative findings.

B. Participants and Setting

This study was conducted in two higher secondary schools, a boys' school and a girls' school, in Jeddah, Saudi Arabia. Using purposeful sampling, two classes were selected from each school for data collection after the participants had consented to volunteer for the study. From the boys' school, two classes, comprising 38-39 students each were randomly selected. While one class constituted the control group, the other one constituted the experimental group. Similarly, from the girls' school, two classes, comprising 39-40 students each were randomly selected. Here too, one class constituted the control group, and the other one constituted the experimental group. The researcher chose a government school instead of a private one because he was interested in students with varying English proficiency levels. Most private school students have high levels of English proficiency. The number of student participants in each class was high because the number of students in government schools is usually high.

C. Research Instruments

Quantitative data were collected through a pre-test and a post-test designed to assess participants' performance on English writing tasks before and after traditional and Wordtune-enabled teaching. The pre-test consisted of paragraph prompts asking students to write short, informative paragraphs. The post-test also included paragraph prompts asking students to compose short, informative paragraphs. Qualitative data were derived from the human ratings of some participants' writing samples to ascertain whether the participants benefitted from using Wordtune as a digital writing tool. Students submitted their writing samples six times over 24 weeks. They completed assignments and writing tests. When completing assignments and taking tests, the students were not allowed to use Wordtune. Instead, they used Wordtune when they practiced writing in class or at home. Two professors from outside the s institution'author compared and evaluated the writing samples for lexical, syntactic, and textual gains.

D. Data Analysis

The data were analyzed using both quantitative and qualitative methods. Quantitative data, constituted by students' scores on the pretest and posttest, were examined using SPSS. Means and standard deviations of the students' writing samples were calculated to obtain the general characteristics of the score distribution. Quantitative data were subjected to paired- and independent-sample t-tests. While paired-sample t-tests were performed to test the effectiveness of the teaching methods, an independent sample t-test was run to ascertain whether there was a difference in writing scores between the experimental and control groups.

The qualitative data, comprising the students' writing samples, were assessed by two human raters, two professors of English from outside Jeddah, with five years of experience teaching English at a Saudi university. The writing samples were examined based on the assessment rubric used to evaluate vocabulary writing tasks, which rates written texts based on their lexical resourcefulness, grammatical accuracy and range, cohesion and coherence, and task completion.

E. Procedure

This study, spanning 24 weeks, took place at two high schools, a boys' school and a girls' school, in Jeddah, Saudi Arabia, during the first semester of 2022/2023. Two classes were selected from each school for data collection. Two classes from the boys' school comprising 38-39 students each were randomly selected. One class constituted the control group and the other one constituted the experimental group. Similarly, from the girls’ school, two classes, comprising 39-40 students each were randomly selected. Here too, one class constituted the control group and the other one constituted the experimental group.

For 24 weeks, each group at each school spent four hours per week learning writing skills from their textbooks. While the control groups received instructions through traditional teaching, the experimental groups received instructions using the Wordtune application. Each group at each school was administered a pre-test at the commencement of the semester. Afterward, at the close of the second semester, each group at each school was given a post-test. The results were recorded and statistically analyzed using SPSS.

In addition to administering the pretest and posttest, the researcher also collected students' writing samples at the beginning and end of the first semester and then at the close of the second semester (six times with six samples). These writing samples were subsequently analyzed qualitatively to determine whether and to what extent the use of Wordtune improved the participants' writing. Moreover, it provided a clear understanding of students' writing improvements and highlighted their writing problems and weaknesses.

IV. Results

A. Phase 1: Quantitative Research

(a). Male Students

Table 1 shows the means and standard deviations for the sample of 38 male students, the control group, and the sample groups, including their marks on writing skills tests before and after receiving instruction in traditional mode.

To test the effectiveness of instruction in traditional mode on students' writing skills, paired-sample t-tests were performed. Specifically, the researcher wanted to test the research questions whether the means of pre-handling the instruction in traditional mode (M = 58.4211, SD = 17.08901) and post-handling the instruction in traditional mode (M = 60.9474, SD = 17.80062) were equal for writing skills. Table 2 contains the results of paired-sample t-test. It reveals that, on average, learners did better after receiving instruction in traditional mode. This improvement was statistically significant [t (37) = 2.314, p =.026 < .05]. The 95% confidence interval of the difference between the means ranged from .31446 to 4.73817, and this difference was significant between the means of the sample.

Table 1
T-Test of Writing Skills Test for Male Students (Control)

	N	M	SD
Pre-control	38	58.4211	17.08901
Post-control	38	60.9474	17.80062

Table 2
T-Test of Writing Skills Test for Control Male Students

	Paired Differences				T	df	sig
			95% CI				
	Mean	SD	Upper	Lower			
Pair 1	2.52632	0.6531	4.73817	.31446	2.314	37	.026

Table 3 shows means and standard deviations for the sample of 39 male students constituting the experimental group, including their scores on the writing skills tests before and after using Wordtune.

To determine whether and to what extent Wordtune enhanced students' writing skills, paired samples t-tests were performed, testing the research query whether the means of writing skills before (M = 58.4615, SD = 15.65054) and after (M = 76.1538, SD = 17.52673) using the Wordtune application were equal. Table 4 shows the findings of the paired-sample t-test. It reveals that, on average, learners performed better with Wordtune. The level of improvement was statistically significant [t (38) = 16.317, p =.000 < .01]. The 95% confidence interval of the difference between means ranged from [15.49731 to 19.88730], which is significant between the means of the samples.

Table 3
T-Test of Writing Skills for Male Students (Experimental)

	N	M	SD
Pre-experimental	39	58.4615	15.65054
Post-experimental	39	76.1538	17.52673

Table 4
T-Test of Writing Skills for Male Students (Experimental)

	Paired Differences				T	df	sig
			95% CI				
	Mean	SD	Upper	Lower			
Pair 1	17.69231	6.77128	19.88730	15.49731	16.317	38	.000

To analyze the results of the final examination, an independent sample t-test was run. Table 5 presents descriptive statistics for each male group. Results reveal that the experimental group outscored (M = 76.1538, SD =17.52673) the control group (M = 60.9474, SD = 17.80062). An independent samples t-test was run to ascertain whether there was a

difference in writing scores between the experimental and control groups. The findings in Table 6 specify a significant difference between the experimental and control groups [t (75) 3.78, p =.000 < .01]. The 95% confidence interval (CI) of the difference between the means ranged from 7.18634 to 23.22662, indicating a significant difference between the means of the sample.

TABLE 5
GROUP STATISTICS OF USING THE WORDTUNE APPLICATION FOR MALE STUDENTS

	n	M	SD
Experimental	39	76.1538	17.52673
Control	38	60.9474	17.80062

TABLE 6
INDEPENDENT T TEST OF USING THE WORDTUNE APPLICATION FOR MALE STUDENTS

	t-test for equality of means							
							95% CI	
	F	Sig	t	df	P-value	Mean difference	Lower	Upper
Equal variances assumed	.134	.715	3.777	75	.000	15.20648	7.18634	23.22662

(b). Female Students

Table 7 shows the means and standard deviations of the sample of 40 female students and the control sample, including their scores on writing skills before and after receiving teaching through the traditional method.

To test the impact of this classical teaching method on students' writing skills, pretest and posttest were used to test the research questions, that the means of pre-(M =55.1250, SD =17.70439) and post (M = 61.1250, SD =18.48379) were equal in writing skills. Table 8 shows the results of the paired-sample t-test. The results of the paired-samples t-test revealed that, on average, students performed better after receiving instruction using the traditional method. This improvement was statistically significant [t (39) = 10.014, p = .000 < .01]. The 95% confidence interval of the difference between means ranged from 4.78812 to 7.21188], and this difference is significant between the means of the sample.

TABLE 7
T-TEST OF WRITING SKILLS TEST FOR FEMALE STUDENTS

	N	M	SD
Pre-control	40	55.1250	17.70439
Post-control	40	61.1250	18.48379

TABLE 8
T-TEST OF WRITING SKILLS FOR FEMALE STUDENTS (CONTROL)

	Paired Differences						
			95% CI				
	Mean	SD	Upper	Lower	T	df	sig
Pair 1	6.00000	3.78932	7.21188	4.78812	10.014	39	.000

Table 9 shows the means and standard deviations for the sample of 39 female students and the experimental sample, including their writing skill scores, before and after using Wordtune.

To test the effectiveness of using the Wordtune application on students' writing skills, paired sample t-tests were performed to test whether the means of pre-using the Wordtune application (M = 54.1026, SD = 18.56255) and post-using the Wordtune application (M = 72.0513, SD= 18.01746) were equal in writing skills. Table 10 shows the findings of the paired-sample t-test, demonstrating that, on average, learners performed better when using Wordtune. This improvement was statistically significant [t (38) = 19.269, p = .000 < .01]. The 95% confidence interval of the difference between means ranged from 16.06301 to 19.83442, and this difference is significant between the means of the sample.

TABLE 9
T-TEST OF WRITING SKILLS FOR FEMALE STUDENTS (EXPERIMENTAL)

	N	M	SD
Pre-experimental	39	54.1026	18.56255
Post-experimental	39	72.0513	18.01746

TABLE 10
PAIRED SAMPLES TEST OF WRITING SKILLS FOR EXPERIMENTAL MALE STUDENTS

	Paired Differences						
			95% CI				
	Mean	SD	Upper	Lower	T	df	sig
Pair 1	17.94872	5.81716	19.83442	16.06301	19.269	38	.000

An independent-samples t-test was run to ascertain whether there was a difference in writing scores between the experimental and control groups. The results showed a significant difference between the experimental (M = 72.0513, SD = 18.01746) and control (M = 61.1250, SD = 13583.47) [t (77) = 2.660, p = .01 < .05] groups. The 95% confidence interval between the means ranged from [2.74609 to 19.10647], suggesting a significant difference between the means of the sample. Therefore, the research question whether there is any difference between the sample means was rejected. Further details are presented in Tables 11 and 12.

TABLE 11
GROUP STATISTICS OF USING THE WORDTUNE APPLICATION FOR FEMALE STUDENTS

	n	M	SD
Experimental	39	72.0513	18.01746
Control	40	61.1250	18.48379

TABLE 12
INDEPENDENT T TEST OF USING THE WORDTUNE APPLICATION FOR MALE STUDENTS

	t-test for equality of means							
							95% CI	
	F	Sig	t	df	P-value	Mean difference	Lower	Upper
Equal variances assumed	.255	.615	2.660	77	.01	10.92628	2.74609	19.10647

An independent-samples t-test was run to ascertain whether the use of Wordtune impacted the writing skills of male and female students differently. The results revealed that this difference was insignificant between male students (M = 76.1538, SD = 17.52673) and female students (M = 72.0513, SD = 18.01746), [t (76) = 1.019, p = 311> .05], as shown in Tables 13 and 14. The 95% confidence interval of the difference between the means ranged from [3.91386 to 12.11899]. This shows no difference between the means of the sample.

TABLE 13
GROUP STATISTICS OF USING THE WORDTUNE APPLICATION FOR MALE AND FEMALE STUDENTS

	n	M	SD
Male	39	76.1538	17.52673
Female	40	72.0513	18.01746

TABLE 14
T-TEST OF USING THE WORDTUNE APPLICATION FOR MALE AND FEMALE STUDENTS

	t-test for equality of means							
							95% CI	
	F	Sig	t	df	P-value	Mean difference	Lower	Upper
Equal variances assumed	.496	.483	1.019	76	.311	4.10256	3.91386	12.11899

B. Phase Two: Qualitative Research

(a). Lexical Gains

A benefit accrued from the experimental groups' inclusion of Wordtune in their writing process was a noticeable improvement in their productive vocabulary. The evaluation of the participants' writing samples revealed that the paragraphs written after practicing writing with Wordtune used lexical resources, especially nouns, adjectives, and verbs, more precisely and appropriately than the writing samples produced without any prior practice with Wordtune. The analysis of writing samples revealed that progressing from the initial samples to later samples, students' successive texts contained more concrete and contextually appropriate nouns, more vivid adjectives, and more precise verbs than the nouns, adjectives, and verbs used in the initial writing samples. A few excerpts from the written samples substantiate this point:

Excerpt A: It is noticeable that the vocabulary of the sentences is very weak. This was based on the writing sample of a male participant, Student 10, in the first cycle of writing samples.

Rayyan is a good guy. He is help his friend. One time, he go 50 kilometers and give a wallet to a friend. The friend missed his wallet. And also one time Rayyan when he come from his office, he helped his friend. Because his friend was sick. Rayyan is a very good man. He is in his office right time. And also he does things what his company says.

Excerpt B: It is noticeable that the level of vocabulary level has improved. This was based on a writing sample from a male participant in the last cycle of writing samples.

Rayyan is a good guy. He is an amazing person because he helps his friends. Once, he drove 50 kilometers to give a friend a wallet. His friend before had missed that wallet. One time, after he came from his office, Rayyan also helped his friend who was not well. Rayyan is a sincere man also. He is in his office room at the right time. And also he does things according to company instructions.

Excerpt A is glaringly weak in lexical resourcefulness: it contains a vague adjective (good), an informal noun (guy); and imprecise verbs (went, does). In contrast, Excerpt B is at least a touch better in terms of lexical resourcefulness: a sharp adjective (amazing); a formal noun (person), and a better verb (drove).

(b). Syntactic Gains

In addition to vocabulary gains, gains in sentence structure were also observed. Analyses of writing samples demonstrated that texts composed after repeated writing practice with Wordtune exhibited greater syntactic maturity (phrasal complexity; greater number of conjoined and embedded clauses in the sentence) than texts written in the initial cycles of writing samples. The following excerpts illustrate this finding:

Excerpt A: It is noticeable that the structure of the sentences is weak. This was taken from the writing sample of a female participant, Student 7, in the second cycle of writing samples.

I sat in plane before five years. And also I think about it. I was afraid turbulence came. And also I was very very nervous. But the plane landed on airport. And also I was happy. I do not forget the flight.

Excerpt B: It is noticeable that the sentence structure has improved. This was based on the writing sample of a male participant, Student 3, in the last cycle of writing samples.

I first sat in a plane five years ago. Even today I think about that experience. I become afraid because during the flight there was turbulence. In turbulence, I was very nervous and also I was very afraid. Finally our plane landed safely on the ground and then I was really happy. Even today, I remember that scary experience.

Comparing the two excerpts for syntactic maturity, one can easily notice the prevalence of simple sentences in Excerpt A to the extent that the excerpt appears as a collection of choppy sentences. Except for the third sentence, Excerpt A is composed entirely of simple sentences accounting for choppiness. Excerpt B, however, is a modest improvement over Excerpt A. Excerpt B at least manages to have more phrasal structures ("even today," "in turbulence," "during the flight") and compound and complex sentences ("I become afraid because during the flight there was turbulence," "Finally our plane landed safely on the ground and then I was really happy.")

V. Discussion

This study aimed to determine whether and to what extent the Wordtune application enhanced participants' academic writing. To this end, the data were analyzed statistically (quantitative analysis in SPSS) and holistically (assessment of writing samples by human raters). I now return to the research questions posed in the current study, unifying the findings of this inquiry and placing them within the discussion of the relevant literature.

The first question posed in this research was whether and to what extent the Wordtune application improved the writing skills of the participants. Findings from this study demonstrated that the use of Wordtune had a positive impact on participants' writing. This result was established using both the paired-sample t-test and the independent sample t-test. A paired-sample t-test was run to test the research question whether the means of pre- and post-use of the Wordtune application were equal for writing skills. The results demonstrate that, on average, the participating students performed better when using the Wordtune application and this improvement was statistically significant. In addition to the paired-sample t-test, an independent sample t-test was run to ascertain the differences in the final exam writing scores between the control group and the experimental group. The results revealed that the experimental group made significant gains in writing compared with the control group. These findings are consistent with the following inquiries into the role of digital writing assistants in enhancing students' writing skills: Barrot (2020), Rad et al. (2023), Choo and Li (2017). For example, Rad et al.'s (2023) investigation of AI's role in fostering students' writing skills demonstrated that the experimental group, which practiced a Wordtune-assisted writing process, was able to significantly improve their writing skills compared to the control group, which did not use Wordtune. Likewise, the findings of this study support those of Barrot's (2020) study on the impact of grammar on students' writing accuracy and clarity. Additionally, the treatment group comprehensively outperformed the control group when the writing post-test outcomes were compared. Qualitatively, Barrot (2020) concluded that participants were able to internalize points of grammar through metalinguistic clarifications attached to Grammarly's feedback.

The second question posed in this research is as follows: What, if any, specific improvements resulted from Wordtune-mediated student writing? The qualitative findings indicated that the participants made modest lexical and

syntactic writing gains by incorporating Wordtune into their writing processes. As far as lexical gains were concerned, a benefit accrued from the experimental groups' use of Wordtune as a digital writing assistant was the improvement in their productive vocabulary. The writing samples produced after Wordtune-enabled practice showed better lexical resourcefulness. They used nouns, adjectives, and verbs with greater precision and appropriateness than lexical resources in writing samples produced without any prior practice with Wordtune. The writing samples produced in the later stages of Wordtune intervention contained more concrete and contextually appropriate nouns, vivid adjectives, and precise verbs than the nouns, adjectives, and verbs used in the initial Wordtune intervention writing samples. These findings largely support previous research on AI's role in enhancing students' academic writing (Aljohani, 2021; Zhao, 2022).

Concerning writing gains at the syntactic level, analyses of writing samples produced in the initial and later stages of Wordtune intervention showed that texts written after repeated writing practice with Wordtune had at least a touch-better sentence structure (phrasal complexity; conjoined and embedded clauses in the sentence) compared with texts written within the initial writing practice with the Wordtune application. To substantiate this, while initial Wordtune writing samples suffered from choppiness owing to the prevalence of bare-bone simple sentences, later Wordtune writing samples were less choppy, as the sentences had at least some degree of phrasal and clausal complexity evident in the presence of compound and complex sentences in the writing samples. This result is consistent with those of other studies exploring the role of AI-based writing tools in supporting students (Rad et al., 2023; Coenen et al., 2021). For instance, Rad et al. (2023) demonstrated that over time, using Wordtune as a digital writing assistant enabled students to compose sentences that combined related ideas of equal or unequal status.

The third research question was the following: Are there any statistical differences in students' writing skills based on gender? To that end, an independent-samples t-test was performed to ascertain whether practicing writing with Wordtune impacted the writing skills of male and female students differently. The results revealed that Wordtune fostered the writing of male and female participants in largely the same way. Quantitatively, for both male and female participants, the post-Wordtune writing test scores were higher than the pre-Wordtune writing test scores. Qualitatively, both male and female participants benefited from writing gains at the vocabulary and sentence structure levels.

VI. Research Implication and Recommendation

The results of this research have key implications for technology-enabled EFL writing. As it offers instant feedback on written texts, incorporating Wordtune into teaching and learning of writing will provide students with the opportunity to learn writing profitably by identifying writing issues and addressing them satisfactorily. Moreover, because Wordtune enables self-directed and self-paced learning, its use in and out of the classroom will likely lead to better performance by cultivating favorable writing habits and practices. Another important implication is that Wordtune usage leads to an increased level of student engagement with writing. Working with Wordtune, they begin to see the interaction between their input sentences and Wordtune rewrites as acts of discovery, prompting them to think about and work more on their own writing practices and products.

Anchored in the findings of the current study, the following recommendations are made to stakeholders. First, instructors should consider integrating AI-powered digital writing tools, such as Wordtune, into their writing classrooms. By explaining to students the merits of Wordtune rewrites over the original input sentences, instructors would get students to think and learn about the factors that distinguish strong sentences from weak sentences, and this consciousness-raising will lead to better writing. Second, students should be encouraged to use Wordtune rigorously in their out-of-class writing practice. When they see multiple and better rewrites of their input sentences, students naturally attempt to understand what made the rewrites better than the original; in this process, they cultivate good writing habits and adopt effective writing techniques. Finally, instructors should inform students about the limitations of Wordtune because an uncritical reliance on a digital writing tool might put students on the path of becoming blind followers of writing technology, which would be counterproductive to their writing.

Since the results and conclusions of the current study are rooted in a rather moderately-sized experimental sample, it would be valuable if researchers worked with larger samples in the future to explore the extent of the facilitative potential of Wordtune or other digital writing tools in the L2 environment. Such studies would help determine the scope of Wordtune and other tools for L2 writing. Additionally, other studies could vary the research period to determine if and to what degree the length of Wordtune-enabled writing practices influences writing development. The field of AI-assisted writing in the L2 domain will benefit from studies that compare the relative worth of different digital writing assistants. Finally, future studies should compare instructors' and students' perceptions of the effectiveness of Wordtune or similar digital writing assistants in EFL learners' writing development to help stakeholders consider these critical digital applications from two different vantage points.

VII. Conclusion

This current study examined the impact of Wordtune on the writing skills of Saudi high school students. Overall, the findings revealed that using Wordtune fostered participants' writing. On average, the students' post-Wordtune writing performance was better than their pre-Wordtune writing performance, and this improvement was statistically significant.

Quantitatively, the Wordtune intervention led to higher writing test scores in the experimental group than in the control group. Qualitatively, it was found that the students using Wordtune gradually gained lexical resourcefulness and sentence structure. With rigorous practice with the Wordtune application, the experimental group participants could produce writings that were better than their initial Wordtune sample texts in terms of lexical resourcefulness and sentence variety. Another key finding is that Wordtune fostered the writing of both male and female participants in largely the same manner. Quantitatively, for both male and female participants, the post-Wordtune writing test scores were higher than the pre-Wordtune writing test scores. Qualitatively, both male and female participants made similar writing gains at the vocabulary and sentence structure levels.

ACKNOWLEDGEMENTS

This work was funded by the University of Jeddah, Jeddah City, Saudi Arabia, under grant No. (UJ-23-SHR-1). The author is grateful for the University of Jeddah's technical and financial support and would like to extend his heartiest thanks to the instructors and students who willingly participated in the study.

REFERENCES

[1] AbdAlgane, M., & Othman, K. A. J. (2023). Utilizing artificial intelligence technologies in Saudi EFL tertiary level classrooms. *Journal of Intercultural Communication*, *23*(1), 92–99. https://doi.org/10.36923/jicc.v23i1.124

[2] Aljohani, R. A. (2021). Teachers and students' perceptions on the impact of artificial intelligence on English language learning in Saudi Arabia. *Journal of Applied Linguistics and Language Research*, *8*(1), 36–47. Retrieved January 26, 2023, from https://www.jallr.com/index.php/JALLR/article/download/1156/pdf1156

[3] Alsied, S. M., & Ibrahim, N. W. (2017). Exploring challenges encountered by EFL Libyan learners in research teaching and writing. *IAFOR Journal of Language Learning*, *3*(2), 143–158. https://doi.org/10.22492/ijll.3.2.06

[4] Azah, D. N. (2019). *The effectiveness of Grammarly checker toward student's writing quality of the fourth semester of English Department at IAIN Tulungagung* [Unpublished undergraduate thesis]. IAIN Tulungagung.

[5] Barrot, J. S. (2022). Integrating technology into ESL/EFL writing through Grammarly. *RELC Journal*. *53*(3), 764–768. https://doi.org/10.1177/0033688220966632

[6] Bitchener, J., & Basturkmen, H. (2006). Perceptions of the difficulties of postgraduate L2 thesis students writing the discussion section. *Journal of English for Academic Purposes*, *5*(1), 4–18. https://doi.org/10.1016/j.jeap.2005.10.002

[7] Choo, Y. B., & Li, K. L. (2017). Digital writing in English language writing instruction. *ARIEL- An International Research Journal of English Language and Literature*, *28*, 1–16. Retrieved February 15, 2023, from https://sujo.usindh.edu.pk/index.php/ARIEL/article/view/19/16

[8] Coenen, A., Davis, L., Ippolito, D., Reif, E., & Yuan, A. (2021). *Wordcraft: A human-AI collaborative editor for story writing*. arXiv. https://doi.org/10.48550/arXiv.2107.07430

[9] Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.

[10] Dickson, B. (2017, November 20). How artificial intelligence is shaping the future of education. *PC Magazine*. Retrieved from https://www.pcmag.com/news/how-artificial-intelligence-is-shaping-the-future-of-education

[11] Finn, H. B. (2018). Articulating struggle: ESL students' perceived obstacles to success in a community college writing class. *Journal of Second Language Writing*, *42*, 101–106. https://doi.org/10.1016/j.jslw.2018.09.001

[12] Fitria, T. N. (2021). Grammarly as AI-powered English writing assistant: Students' alternative for writing English. *Metathesis Journal of English Language, Literature, and Teaching*, *5*(1), 65–78. https://doi.org/10.31002/metathesis.v5i1.3519

[13] Flowerdew, J. (2019). The linguistic disadvantage of scholars who write in English as an additional language: Myth or reality. *Language Teaching*, *52*(2), 249–260. https://doi.org/10.1017/S0261444819000041

[14] Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, *3*, 100055. https://doi.org/10.1016/j.caeai.2022.100055

[15] Geiser, S., & Studley, R. (2001). *Relative contribution of high school grades, SAT I and SAT II scores in predicting success at UC: Preliminary findings* [Unpublished manuscript]. University of California.

[16] Hamouma, C., & Menezla, N. (2019). The impact of digital literacy proficiency on EFL students' academic writing performance: A case study of Algerian third year EFL students. *International Journal of Digital Literacy and Digital Competence (IJDLDC)*, *10*(4), 40–55. https://doi.org/10.4018/IJDLDC.2019100103

[17] Hamzaoui, C. (2021). Scrutinizing Algerian EFL students' challenges in research teaching and writing. *Hungarian Educational Research Journal*, *11*(4), 478–488. https://doi.org/10.1556/063.2021.00047

[18] Hanauer, D. I., Sheridan, C. L., & Englander, K. (2019). Linguistic injustice in the writing of research articles in English as a second language: Data from Taiwanese and Mexican researchers. *Written Communication*, *36*(1), 136–154. https://doi.org/10.1177/0741088318804821

[19] Karyuatry, L. (2018). Grammarly as a tool to improve students' writing quality: Free online-proofreader across the boundaries. *JSSH (Jurnal Sains Sosial dan Humaniora)*, *2*(1), 83–89. https://doi.org/10.30595/jssh.v2i1.2297

[20] Komba, S. C. (2015). Challenges of writing theses and dissertations among postgraduate students in Tanzanian higher learning institutions. *International Journal of Research Studies in Education*, *5*(3), 71–80. https://doi.org/10.5861/ijrse.2015.1280

[21] Lahuerta, A. (2017). Analysis of accuracy in the writing of EFL Students enrolled on CLIL and Non-CLIL programmes: The impact of grade and gender. *The Language Learning Journal*, *48*(2), 121–132. https://doi.org/10.1080/09571736.2017.1303745

[22] Lin, C. J., & Mubarok, H. (2021). Learning analytics for investigating the mind map-guided AI chatbot approach in an EFL flipped speaking classroom. *Educational Technology & Society*, *24*(4), 16–35. https://doi.org/10.30191/ETS.202110_24(4).0002

[23] Lin, L. H. F., & Morrison, B. (2021). Challenges in academic writing: Perspectives of engineering faculty and L2 postgraduate research students. *English for Specific Purposes*, *63*, 59–70. https://doi.org/10.1016/j.esp.2021.03.004
[24] Mahammoda, S. A. (2016). Factors affecting the quality of undergraduate research work in Bahir Dar University, Ethiopia. *International Journal of Innovative Research and Development*, *5*(12), 23–27. Retrieved from https://www.internationaljournalcorner.com/index.php/ijird_ojs/article/view/136671/95794
[25] McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth Summer Research Project on artificial intelligence, August 31, 1955. *AI Magazine*, *27*(4), 12–14. https://doi.org/10.1609/aimag.v27i4.1904
[26] McDonough, K., & Crawford, W. J. (2018). Identifying effective writing tasks for use in EFL write-to-learn language contexts. *The Language Learning Journal*, *48*(4), 469–480. https://doi.org/10.1080/09571736.2018.1465990
[27] Moore, K. A., Rutherford, C., & Crawford, K. A. (2016). Supporting postsecondary English language learners' writing proficiency using technological tools. *Journal of International Students*, *6*(4), 857–872. https://doi.org/10.32674/jis.v6i4.321
[28] Moussalli, S., & Cardoso, W. (2020). Intelligent personal assistants: Can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning*, *33*(8), 865–890. https://doi.org/10.1080/09588221.2019.1595664
[29] Nobles, S., & Paganucci, L. (2015). Do digital writing tools deliver? Student perceptions of writing quality using digital tools and online writing environments. *Computers and Composition*, *38*, 16–31. https://doi.org/10.1016/j.compcom.2015.09.001
[30] Perry, F. (2021). The use of embedded digital tools to develop English language proficiency in higher education. *Journal of Academic Language and Learning*, *15*(1), 1–12. Retrieved January 15, 2023, from https://journal.aall.org.au/index.php/jall/article/view/699
[31] Phuong, W. T. N. (2021). Difficulties in studying writing of English-majored sophomores at a university in Vietnam. *European Journal of Education Studies*, *8*(10), 313–330. https://doi.org/10.46827/ejes.v8i10.3962
[32] Purcell, K., Buchanan, J., & Friedrich, L. (2013). *The impact of digital tools on student writing and how writing is taught in schools*. Pew Research Center. Retrieved February 7, 2023, from https://www.pewresearch.org/internet/wp-content/uploads/sites/9/media/Files/Reports/2013/PIP_NWP-Writing-and-Tech.pdf
[33] Rad, H. S., Alipour, R., & Jafarpour, A. (2023). Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: A case of Wordtune application. *Interactive Learning Environments*, 1–21. https://doi.org/10.1080/10494820.2023.2208170
[34] Rosário, P., Högemann, J., Núñez, J. C., Vallejo, G., Cunha, J., Rodríguez, C., & Fuentes, S. (2019). The impact of three types of writing intervention on students' writing quality. *PLoS One*, *14*(7), e0218099. https://doi.org/10.1371/journal.pone.0218099
[35] Ruscetti, T., Krueger, K., & Sabatier, C. (2018). Improving quantitative writing one sentence at a time. *PLoS One*, *13*(9), e0203109. https://doi.org/10.1371/journal.pone.0203109
[36] Singh, M. K. M. (2017). International EFL/ESL master students' adaptation strategies for academic writing practices at tertiary level. *Journal of International Students*, *7*(3), 620–643. https://doi.org/10.5281/zenodo.570025
[37] Winans, M. D. (2021). Grammarly's tone detector: Helping students write pragmatically appropriate texts. *RELC Journal*, *52*(2), 348–352. https://doi.org/10.1177/00336882211010506
[38] Zhang, R., & Zou, D. (2022). Types, features, and effectiveness of technologies in collaborative writing for second language learning. *Computer Assisted Language Learning*, *35*(9), 2391–2422. https://doi.org/10.1080/09588221.2021.1880441
[39] Zhao, X. (2022). Leveraging artificial intelligence (AI) technology for English writing: Introducing Wordtune as a digital writing assistant for EFL writers. *RELC Journal*. https://doi.org/10.1177/00336882221094089
[40] Zheng, B., & Warschauer, M. (2017). Epilogue: Second language writing in the age of computer-mediated communication. *Journal of Second Language Writing*, *36*, 61–67. https://doi.org/10.1016/j.jslw.2017.05.014
[41] Zhou, S. A., & Hiver, P. (2022). The effect of self-regulated writing strategies on students' L2 writing engagement and disengagement behaviors. *System*, *106*, 102768. https://doi.org/10.1016/j.system.2022.102768

Fawaz Al-Mahmud has been an assistant professor in the English Language and Translation Department at the University of Jeddah since 2020. He received his Ph.D. in Education from Memorial University, Canada, where he specialized in Teaching English as a Foreign Language. He has published several articles addressing issues in foreign language teaching and learning (https://orcid.org/0000-0002-1029-1556).