International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

The Usability of ChatGPT: A Language Learning Tool for Improving Grammar and Vocabulary in L2 Writing

Donna Mae M. Sanchez

Student, Faculty of Advanced and International Studies, Davao Oriental State University, Davao Oriental, Philippines

Abstract

The study investigates the usability of ChatGPT as a language learning tool and the extent to which it can improve grammar and vocabulary in L2 writing among grade 9 students. The researcher adapted the questionnaire of Lund (2001) to determine the usability evaluation of ChatGPT in terms of usefulness, ease of use, ease of learning, and satisfaction. Pre- and post- written outputs were also collected to identify the effectiveness of the tool in improving grammar and vocabulary. Moreover, quasi-experimental design was used in the study wherein respondents were divided into control and experimental groups. The latter used ChatGPT for four weeks whereas, the former continued with the traditional language learning tools. Statistical mean, standard deviation, paired T-test, independent T-test, and ANOVA were also used in analyzing and interpreting the data gathered. The results reveal a high usability evaluation rate of ChatGPT. A significant difference is also observed in the experimental group's grammatical accuracy and average words per sentence but not on total words and unique words. Furthermore, results reveal that male students and frequent users of ChatGPT have more vocabulary diversity. Overall, the experimental group outperformed those in the control group in all parameters. Thus, the results suggest that ChatGPT is effective in improving grammar but has limited impact in vocabulary development. This entails additional vocabulary instructional strategies to maximize the effectiveness of the tool.

Keywords: AI-assisted learning, ChatGPT usability, grammar improvement, second language writing, vocabulary improvement

1. Introduction

Several studies considered writing as the most indispensable language skill that plays a crucial role in both personal and professional successes (Gautam, 2019; Mustafa et al., 2019; Ingale, 2019; Urbano et al., 2021). As important as it may be, writing is still regarded as one of the most challenging skill to develop (Ventayen, 2020).

This difficulty is attributed to lack of proficiency in grammar and vocabulary (Dwi Putri et al., 2022; Labicane & Oliva, 2022). In a study conducted in Kosovo, language teachers perceive difficulties in writing to stem from limited vocabulary and grammar proficiency (Jashari & Fojkar, 2019). Similarly, in the Philippines, only 6 percent demonstrated writing proficiency expected of their level as reflected in the 2019 Southeast Asia Primary Learning Metrics (SEA-PLM) (DepEd, 2021).

This calls for a need to constantly enhance methods in improving writing skill. The presence of advanced technologies such as artificial intelligence provides opportunities to effectively improve second language writing (Tan, 2023). For instance, research conducted in Malaysia highlights the effectiveness of an online grammar checker in improving essay writing (Jayavalan & Razali, 2020). Learners who are exposed to the tool enhanced their grammatical knowledge such as subject-verb agreement, sentence structure, spelling, and punctuation. Further, Filipino university students agree that artificial intelligence tools are usable in detecting mistakes in writing, grammar, sentence structure, and vocabulary (Ventayen, 2020). These tools are significant in creating informative and accurate papers.

A new tool that has recently gained prominence is ChatGPT. This utilizes deep learning techniques to generate human-like responses (Magruder et al., 2023). Research indicates that this tool has the potential to revolutionize language learning (Kohnke et al. 2023). ChatGPT is said to have the potential to enhance language skills (Kim et al., 2023; Koraishi, 2023). However, this has not been the focus of most existing researches. Most of the existing studies emphasize the ethical implications of ChatGPT but not its impact to writing proficiency (Jiao et al., 2023; Feng et al., 2023; Yan, 2023) Similarly, this tool might have reached the students locally but, there are no researches that yet explored ChatGPT in the locality. This further indicates that no researches have yet discussed the extensiveness of ChatGPT use among the learners in Mati City.

Thus, the study is conducted to gain insights on the use of ChatGPT as a potential tool to improve grammar and vocabulary in L2 Writing. Moreover, it seeks to understand the holistic user experience by evaluating its usability based on the perspectives of the users. This research also addresses a local need. In a school in Mati City, only 40.83% of junior high school students have attained the writing proficiency expected of their grade level. This clearly emphasizes the urgency to explore effective and innovative tools to enhance writing skill. Needless to say, the study is conducted to address this gap and serve as a springboard for future researches to recognize the use of ChatGPT and artificial intelligence tools in language learning.

2. Research Objectives

The objectives of the study are the following:

1. To determine the demographic profile of the respondents using ChatGPT in terms of gender and ChatGPT usage.
2. To evaluate the usability of ChatGPT as perceived by the respondents using ChatGPT in terms of usefulness, ease of use, ease of learning, and satisfaction.
3. To determine the extent to which the control and experimental group improve grammar and vocabulary in L2 writing in terms of grammatical accuracy, total words, unique words, and average words per sentence.
4. To determine the significant difference in the grammar and vocabulary improvement between the control and experimental group in terms of grammatical accuracy, total words, unique words, and average words per sentence.
5. To determine the significant difference in the grammar and vocabulary improvement of the experimental group in terms of gender and ChatGPT usage.

3. Research Methodology

3.1. Research Design. This study used quantitative method particularly quasi-experimental design because the researcher considered it as the most adaptable research design for school settings. It enabled

the assessment of the extent of improvement in grammar and vocabulary without requiring random assignment of the students to experimental and control groups. This way, the researcher was able to use existing groups or classes without disrupting the normal activities of the school.

3.2. Research Locale. The study was conducted in the biggest public secondary school in Mati City and the entire province of Davao Oriental. This institution is situated at Mangga St., Barangay Sainz, City of Mati, Davao Oriental. Currently, it holds the largest number of enrollees in the city. The school was the best place to gather data since it has necessary facilities required in the study such as an ICT laboratory. It has desktop devices and strong internet connectivity, which were enough to cater the number of respondents.

3.3. Research Respondents. The respondents of the study consisted of two grade 9 classes or eighty-two (82) students in a secondary school in the Division of the City of Mati. The control and experimental groups each had forty-one (41) students. These classes were chosen based on the following inclusion criteria: students have foundational knowledge in English language; they are actively engaged in learning English as a second language; are willing to commit to the study's activities; and have access to a computer or mobile device. Further, this study employed non-random purposive sampling technique in determining the respondents since it is the most suitable sampling technique to the study's research design.

3.4. Research Instrument. The Usability, Satisfaction, and Ease of Use (USE) Questionnaire of Arnold Lund (2001) is utilized to collect the data on the ChatGPT usability evaluation of the respondents. This questionnaire is one of the standard questionnaires to assess the usability of any technical system (Lund, 2001). It is divided into four components namely: Usefulness, Ease of Use, Ease of Learning, and Satisfaction. USE Questionnaire has a Cronbach's alpha of 0.973 which means that the instrument is reliable or shows consistency of results. As stated, a variable is said to be reliable if it gives a Cronbach's alpha of more than 0.70 (Fadhilah et al., 2022). In terms of determining the extent to which ChatGPT improves grammar and vocabulary, two writing samples were collected from the experimental and control groups: the first one represented their initial grammar and vocabulary skills and the second represented their writing performance after being exposed to ChatGPT (experimental group) and the traditional language learning tool (control group). Additionally, grammatical accuracy, total words used, unique words, and average words per sentence per writing sample were analyzed using ChatGPT itself. Grammatical accuracy and average words per sentence are the parameters for grammar improvement, whereas total words and unique words pertain to vocabulary improvement. Identifying these components helped assess the grammar and vocabulary of the respondents.

3.5. Data Procedures. To gather the data needed, the researcher sought consent and approval from respective authorities before commencing the study. These offices include: the Davao Oriental State University Research Ethics Office; Schools Division Superintendent of the Department of Education-Division of the City of Mati; and the respondent school. Once approved, two grade 9 classes were chosen as experimental and control group. Pre-implementation phase was conducted for both the experimental and control group. Here, a writing task was given to evaluate the initial grammar and vocabulary skills of the respondents. The students wrote a three-paragraph essay about their real-life hero. These outputs were encoded and analyzed for its grammatical accuracy, total words, unique words and average words per sentence using ChatGPT. In the implementation phase, the experimental group used ChatGPT for four weeks to complete learning tasks in grammar, vocabulary, and writing. They also had sessions with the researcher at least once a week in the ICT laboratory. The remaining days of the week were spent for personal interaction with the tool. Meanwhile, the control group followed the traditional language learning

tools without exposure to ChatGPT. Lastly, the post-implementation phase included a post-written assessment for both the experimental and control groups. The students were involved in a writing procedure by addressing a similar topic namely: a three-paragraph essay describing their thoughts on the qualities of a real-life hero. These outputs were analyzed using the same procedure as the pre-written outputs. Similarly, the demographic profile of the experimental group was determined and ChatGPT usability evaluation was also answered through the USE Questionnaire (Lund, 2001). The answered questionnaires and written outputs were subjected to statistical treatment, analysis, and interpretation.

3.5. Data Analysis. ChatGPT was used to assess the quality and characteristics of the pre-and post- written outputs of the respondents. ChatGPT performed a detailed analysis focusing on grammatical accuracy, total words used, unique words, and average words per sentence. Further these statistical tools were used to analyze the data: frequencies/percentages, mean, standard deviation, paired t-test, independent t-test, and ANOVA (Analysis of Variance).

4. Results and Discussions

4.1. Demographic Profile of the Respondents. Displayed in Table 1 is the demographic characteristics of the respondents using ChatGPT in terms of gender and frequency of ChatGPT Usage. Out of 41 respondents, 80.49% were female and 19.51% were male. This implies that majority of the respondents were female. This uneven distribution of respondents in terms of gender was inevitable since the researcher used the quasi-experimental design. The latter required a pre-existing group. In reference to the study, the researcher chose classes that met the inclusion criteria and these sections had fewer male students than female.

	Category	Frequency(n=41)	Percentage (100%)
Gender	Male	8	19.51%
	Female	33	80.49%
Frequency of ChatGPT Usage	Once a day or more	11	26.83%
	Almost Daily	8	19.51%
	Twice a week	14	34.15%
	Once a week	5	12.20%
	Every other week	3	7.32%

Table 1: Demographic Profile of the Respondents in Terms of Gender and ChatGPT Usage

Further, the frequency of ChatGPT usage indicates that 34.15% of the respondents used ChatGPT twice a week, 26.83% for once a day or more, and 19.51% for almost daily. The high percentage of respondents using ChatGPT at least twice a week or more suggests that they had sufficient exposure to the tool. This further implies that ChatGPT was accepted as a learning tool by the respondents that can guide and assist them in their various learning tasks. Ahmed (2024), in his survey of 100 participants, stated that ChatGPT was widely accepted as a learning tool. Students also reported enhanced engagement, individualized learning, and better comprehension.

4.2. Usability Evaluation of ChatGPT as Perceived by the Respondents using ChatGPT in terms of Usefulness, Ease of Use, Ease of Learning and Satisfaction.

Table 2 shows the usability evaluation of ChatGPT in terms of usefulness. Based on the data, respondents generally agree that ChatGPT is useful. ChatGPT is seen as a beneficial tool that provides assistance in academic or personal activities. This can be attributed to the interactivity and responsiveness of the tool. As such, Menon and Shilpa (2023) indicated that these are key-features as to why ChatGPT is perceived as valuable and enjoyable to use. Since ChatGPT provides real-time feedback, it can increase users' intention to use the technology. It can also be seen that the statement about granting ChatGPT control over daily activities received comparatively lower ratings. This suggests that while ChatGPT is useful, it still has some limitations in term of control. ChatGPT is used more as an assistive tool rather than a decision-making authority. This reduced reliance is attributed to some misleading and incorrect information provided by the tool.

Abernathy (2024) supports this in his narrative review in which he emphasized that ChatGPT is often perceived as supportive rather than authoritative. It best complements traditional methods since it is always accessible and provides immediate feedback. However, users of ChatGPT also reported that it generates inaccurate and misleading information. This is the reason why some users do not fully trust it for critical decisions.

Table 2: Usability Evaluation of ChatGPT as Perceived by the Respondents using ChatGPT in terms of Usefulness

Statements	Mean	SD	Descriptive Interpretation
1. It helps me be more effective.	3.88	0.68	Agree
2. It helps me be more productive.	3.85	0.73	Agree
3. I consider it useful.	4.27	0.81	Strongly Agree
4. It gives me more control over the activities in my life.	3.32	0.96	Fairly Agree
5. It makes the things I want to do easier to accomplish.	4.00	0.74	Agree
6. It saves me time when I use it.	4.10	0.80	Agree
7. It meets my needs.	3.63	0.94	Agree
8. It does everything I would expect it to do.	3.51	0.75	Agree
Overall	**3.82**	**0.54**	**High**

Table 3 presents the usability evaluation of ChatGPT in terms of ease of use. Here, ChatGPT is perceived to be generally user-friendly. Based on the data, the statement "It is easy to use" got the highest rating among the statements. This can be attributed to the tool's design and interface that effectively cater the needs of the users.

Table 3. Usability Evaluation of ChatGPT as Perceived by the Respondents using ChatGPT in terms of Ease of Use

Statements	Mean	SD	Descriptive Interpretation
1. It is easy to use.	4.32	0.72	Strongly Agree

2. It is simple to use.	4.37	0.70	Agree
3. It is user-friendly.	4.22	0.79	Strongly Agree
4. It requires the fewest steps possible to accomplish what I want to do with it.	3.83	0.89	Agree
5. It is flexible.	3.51	0.75	Agree
6. Using ChatGPT is effortless.	3.37	1.02	Fairly Agree
7. I can use ChatGPT without written instructions.	2.80	1.19	Fairly Agree
8. I do not notice any inconsistencies as I use it.	3.00	0.92	Fairly Agree
9. Occasional and regular users would like it.	3.85	0.96	Agree
10. I can recover from mistakes quickly and easily as I use it.	4.17	0.86	Agree
11. I can use it successfully every time.	3.88	0.78	Agree
Overall	**3.76**	**0.54**	**High**

Meanwhile, although ChatGPT is perceived to be generally easy to use, others still encounter minor difficulties and inconsistencies. This is evident in the statement that received lower mean score which is "I can use ChatGPT without written instructions". This can be explained by lack of familiarity. Since the students were new users of ChatGPT, some may have initially struggled in using it. Further, some of the students only used ChatGPT during the sessions they had with the researcher which only lasted an hour. This implies less personal interaction. Pang et.al. (2024) stated that less personal interaction with the tool can lead to less familiarity. Users who engage frequently with ChatGPT tend to develop a better understanding of its functions.

Moreover, it can be seen in table 4 that ChatGPT is positively rated in terms of ease of learning. As seen in the table, the statement "Learning how to use it is easy" got the highest mean score. This implies a high level of usability. This can be due to ChatGPT's interface which closely resembles daily communication. Gupta (2024) emphasized that ChatGPT uses advanced natural language processing (NLP) techniques to generate human-like text. This facilitates interactions that resemble human conversation. Meanwhile, the statement "I quickly became skillful with it" has had the lowest mean score. This means that while most respondents found ChatGPT easy to learn, some still required sufficient time to develop proficiency in it.

Table 4: Usability Evaluation of ChatGPT as Perceived by the Respondents using ChatGPT in terms of Ease of Learning

Statements	**Mean**	**SD**	**Descriptive Inter-pretation**
1. I learned to use it quickly.	4.27	0.78	Strongly Agree
2. I easily remember how to use it.	4.22	0.85	Strongly Agree
3. Learning how to use it is easy.	4.34	0.73	Strongly Agree
4. I quickly became skillful with it.	3.78	0.76	Agree
Overall	**4.15**	**0.62**	**High**

In reference to the current study, the implementation phase only took place for four weeks and it may not be enough to allow few respondents to skillfully master the tool. Pang et al (2024) stated that mastering ChatGPT becomes difficult due to lack of prior exposure to technology and less interaction with it. They

further emphasized that frequent interaction to technology is hindered by financial constraints. In lieu of the current study, some of the respondents have limited access to technology such as computers or mobile phones. These students most likely use the phones of their parents or other family members since these are too expensive for them. Stable internet connection is also a problem as this would require additional funds. Moreover, although the school may have an available computer laboratory, it cannot accommodate all the students at the same time. These financial reasons hindered some of the students to constantly use ChatGPT.

Table 5. Usability Evaluation of ChatGPT as Perceived by the Respondents using ChatGPT in terms of Satisfaction

Statements	Mean	SD	Descriptive Interpretation
1. I am satisfied with it.	4.22	0.69	Strongly Agree
2. I would recommend it to a friend.	4.15	0.94	Agree
3. It is fun to use.	3.83	0.97	Agree
4. It works the way I want it to work.	3.98	0.91	Agree
5. It is wonderful.	4.21	0.78	Strongly Agree
6. I feel I need to have it.	3.93	0.88	Agree
7. It is pleasant to use.	3.98	0.79	Agree
Overall	**4.04**	**0.64**	**High**

Lastly, table 5 indicates the usability evaluation of ChatGPT in terms of satisfaction. It can be seen in the data that the respondents were generally satisfied with their experience in ChatGPT. The respondents find the tool effective and valuable. Satisfaction can be attributed to how it meets the students' needs, whether for quick information or other relevant tasks. Yassin & Bashir (2024) highlighted this in their study in which a significant majority of the 200 students surveyed expressed a high level of satisfaction with the tool. ChatGPT meets educational needs effectively and this convenience is a major factor contributing to the students' satisfaction.

Meanwhile, the statement "It is fun to use" got the lowest rating among the statements. Although the mean score indicates that respondents generally agree that using ChatGPT is fun, their agreement is not as strong compared to other statements. This means that most of the students view ChatGPT as a practical tool rather than a source of entertainment. This further suggests that the respondents have primarily engaged with ChatGPT for productivity rather than recreational purposes. This aligns to the study of Postigo-Zumaran et al. (2024) wherein ChatGPT is positively valued for its ability to improve productivity and facilitate understanding of theoretical concepts.

4.3. Comparison of Pre-test and Post-Test Results in Grammar and Vocabulary for the Control and Experimental Group. Presented in table 6 is the comparison of pre-test and post-test results in terms of grammatical accuracy, total words, unique words, and average words per sentence for the control group or the students who did not use ChatGPT using a paired t-test. The findings reveal significant differences in all assessed areas. The mean scores for each area decreased from pre-test to post-test. This suggests that students in the control group demonstrated lower grammatical accuracy, used fewer total and unique words, and wrote shorter sentences in the post-test. This further implies that without ChatGPT, students have struggled to improve or sustain their writing performance over time.

This can be associated with limited interaction and immediate feedback as ChatGPT can provide. Since the control group did not use ChatGPT in accomplishing language tasks, they have missed valuable opportunities for real-time feedback and support. In the study of Al-Durayhim (2023), the control group also performed less in the post-test than the pre-test. He emphasized that absence of immediate feedback to grammar mistakes caused this decrease in the students' performance. Similarly, Shaikh et al. (2023) stated that one of the advantages of ChatGPT is its availability to support students all the time which a human partner cannot do so easily. As further indicated by Wahyuni (2022), ChatGPT provides direct feedback to a learner way faster than teachers or peers. This is what students in the control group missed before the conduct of the post-test.

Table 6: Comparison of Pre-Test and Post-Test Results in Grammar and Vocabulary using Paired T-Test for the Control Group

Areas	**Tests**	**Mean**	***t-value***	***p-value***	**Remarks**
Grammatical Accuracy	Pre-test	1.68	2.933	.005	There is a significant difference
	Post-test	1.32			
Total Words	Pre-test	2.68	3.000	.004	There is a significant difference
	Post-test	2.29			
Unique Words	Pre-test	2.61	3.480	.001	There is a significant difference
	Post-test	2.17			
Average Words per Sentence	Pre-test	1.78	3.583	.001	There is a significant difference
	Post-test	1.37			

Another possible reason of this decrease is the interference from other subjects or activities in school and less motivation during the conduct of the post-test. Over the course of the month, students were most likely exposed to other subjects, assignments, and learning tasks. This may have caused them to have less motivation with the study. Particularly, Al-Durayhim (2023) indicates that this could happen by how the pre-test may have temporarily boosted motivation, but if students are preoccupied with other tasks which they consider to be more important than the said study, they may lose interest.

Moreover, table 7 compares the pre-test and post-test results for the experimental group or the students who used ChatGPT using a paired t-test. The results show significant improvement in grammatical accuracy and average words per sentence. This means that exposure to ChatGPT improved understanding of grammar rules. Particularly, ChatGPT served as a source of immediate feedback which helps the learners recognize and correct their mistakes. This also entails that students have actively used ChatGPT to review and learn grammar concepts.

Table 7: Comparison of Pre-Test and Post-Test Results in Grammar and Vocabulary for Experimental Group using Paired T-Test

Areas	**Tests**	**Mean**	***t-value***	***p-value***	**Remarks**
Grammatical Accuracy	Pre-test	2.17	-4.092	<.001	There is a significant difference
	Post-test	2.71			
Total Words	Pre-test	3.02	1.432	.160	There is NO significant difference
	Post-test	2.88			

Unique Words	Pre-test	2.90	-.573	.570	There is NO significant difference
	Post-test	2.95			
Average Words per Sentence	Pre-test	2.54	-2.96	.046	There is a significant difference
	Post-test	2.85			

As such, the study of Lalira et al. (2024) revealed a significant improvement in grammar scores with an average increase of 15%. They further emphasized that this significant improvement is attributed to the ability of ChatGPT to provide real-time feedback in grammar. Chhabriya et al. (2024) further concluded that ChatGPT-assisted students showed more improvement in grammar proficiency than those exposed with the traditional approach. This implies that ChatGPT provides more effective feedback in grammar. Niyozov et al. (2023) also conducted a comprehensive evaluation of the efficacy of ChatGPT in second language writing. The results of their study demonstrated a significant increase in grammar accuracy which improved by 19%.

It can also be seen in the data that there is a significant difference in average words per sentence. This means that ChatGPT encouraged more complex sentence construction. This is made possible by how ChatGPT generate detailed responses. Baskara (2023) stated that ChatGPT offers guidance in structuring sentences and paragraphs. Similarly, Avalur et al. (2023) revealed that ChatGPT's model training on large datasets allows it to produce more comprehensive answers. This increases the average words per sentence of users in L2 writing. With this, as students interact with the tool, they may have gained confidence in writing longer sentences and adapted the complexity of sentence structures provided by ChatGPT.

Meanwhile, no significant differences were found in total words ($t = 1.432$, $p = .160$) and unique words ($t = -0.573$, $p = .570$), suggesting stable vocabulary usage. While writing became more structured and complex, the respondents did not necessarily write more words. This means that vocabulary diversity did not significantly change. The results infer further that the students prioritized accuracy over quantity. They focused more on producing grammatically correct sentences rather than lengthy responses. Similarly, Giray (2024) emphasized in his research that ChatGPT indeed enhances the quality of students' writing. It promotes conciseness since this ChatGPT assists in refining ideas and providing concrete examples.

Compared to the control group's decrease in all parameters, these results suggest that ChatGPT-assisted learning enhances grammar. However, its effect on overall vocabulary usage is limited. This means that ChatGPT enhances writing structure but there is still a need for additional vocabulary-focused instructional strategies.

4.4. Comparison of Grammar and Vocabulary Assessment Results Between Control and Experimental Group using Independent T-Test. Presented in table 8 is the comparison of grammar and vocabulary assessment results in terms of grammatical accuracy, total words, unique words, and average words per sentence between the control and experimental groups using an independent t-test. The findings show a significant difference in all areas. The experimental group outperformed the control group in grammar and vocabulary.

Table 8: Comparison of Grammar and Vocabulary Assessment Results between Control and Experimental Group using Independent T-Test

Areas	Group	Mean	t-value	p-value	Remarks
Grammatical Accuracy	Control	1.50	-9.953	<.001	There is a significant difference
	Experimental	2.44			
Total Words	Control	2.49	-4.532	<.001	There is a significant difference
	Experimental	2.95			
Unique Words	Control	2.39	-6.541	<.001	There is a significant difference
	Experimental	2.93			
Average Words per Sentence	Control	1.57	-13.272	<.001	There is a significant difference
	Experimental	2.70			

Students in the experimental group showed substantially better grammatical accuracy than the control group. There is also a high significant difference in the experimental group in terms of average words per sentence at a mean of 2.70 over the control group at a mean of 1.57. This can be explained by how ChatGPT handles complex inputs which in turn allows students to create more detailed sentences.

These results indicate that ChatGPT-assisted learning can significantly enhance students' writing proficiency as compared to traditional learning. The tool can aid the students in producing less grammatical errors and more complex sentences.

4.5. Analysis of Variance on the Grammar and Vocabulary Assessment Results on the Experimental Group across Gender and ChatGPT Usage. It can be seen in table 9 that there are no significant differences in grammatical accuracy, total words, and average words per sentence across gender. This indicates that both male and female students performed similarly in these areas. Masoudi (2024) indicates that gender does not significantly play a role in grammar and vocabulary improvement since ChatGPT provides personalized feedback to all learners regardless of gender.

Meanwhile, there is a significant difference in unique words wherein male respondents used a more diverse vocabulary than its counterpart. This implies that the former prioritizes expressiveness in language. This leads to the use of a wider variety of words. On one hand, the latter prefers accuracy which could result to a more controlled and concise vocabulary.

Additionally, Borhan (2023) stated that male students are more experimental in their language use than female students. Siregar et al. (2023) further emphasized that male students display higher intensity in using ChatGPT for learning purposes than female learners. The latter excels more in academic writing and demonstrates superior organization of sentences which reflects their emphasis on clarity. Kronberg (2024) also explained that this can be associated to how male learners use ChatGPT for basic use and integrate it into their current habits.

In general, the findings suggest that ChatGPT's impact on grammar and vocabulary improvement is consistent among genders.

Table 9: Analysis of Variance on the Grammar and Vocabulary Assessment Results on the Experimental Group across Gender

Areas	Gender	Mean	F	p-value	Remarks
Grammatical Accuracy	Male	2.44	.001	.993	There is NO significant difference
	Female	2.44			
Total Words	Male	2.94	.006	.939	There is NO significant difference
	Female	2.95			
Unique Words	Male	3.13	3.382	.041	There is a significant difference
	Female	2.58			
Average Words per Sentence	Male	2.56	.878	.354	There is NO significant difference
	Female	2.73			

Further, table 10 shows the analysis of variance on the grammar and vocabulary assessment results for the experimental group across different frequencies of ChatGPT usage. It can be seen in the data that no significant differences are observed in grammatical accuracy, total words, and average words per sentence. This means that the frequency with which the students use ChatGPT does not significantly affect their grammar accuracy, the length of their responses, and the complexity of their sentence structures.

Table 10: Analysis of Variance on the Grammar and Vocabulary Assessment Results on the Experimental Group across Frequency of Usage

Areas	Frequency of Usage	Mean	F	p-value	Remarks
Grammatical Accuracy	Once a day or more	2.27	.917	.464	There is NO significant difference
	Almost Daily	2.50			
	Twice a week	2.61			
	Once a week	2.20			
	Every other week	2.50			
Total Words	Once a day or more	2.91	.241	.913	There is NO significant difference
	Almost Daily	2.88			
	Twice a week	3.07			
	Once a week	2.90			
	Every other week	2.83			
Unique Words	Once a day or more	3.25	3.195	.049	There is significant difference
	Almost Daily	2.98			
	Twice a week	2.72			
	Once a week	2.60			
	Every other week	2.45			
Average Words per Sentence	Once a day or more	2.68	.153	.961	There is NO significant difference
	Almost Daily	2.75			
	Twice a week	2.68			
	Once a week	2.60			
	Every other week	2.83			

This could be attributed to superficial interaction with ChatGPT. Some of the students may not have used the tool for in-depth language practice during their personal interaction with it. This happens if ChatGPT was mostly used to generate quick answers without critically analyzing and learning from the feedback. As stated by Al-Garaady and Mahyoob (2023), this kind of interaction can only ask the tool to identify surface-level errors with no depth for comprehensive understanding.

On one hand, a significant difference is seen in unique words with students who used ChatGPT once a day or more having the highest vocabulary diversity. Meanwhile, those using it every other week had the lowest. This suggests that more frequent ChatGPT use may enhance vocabulary, though overall writing proficiency remains consistent regardless of how often students use it. This could mean that frequent users used ChatGPT not as source of answers but as a way to learn the language. As mentioned by Khzouz et al. (2024), the more frequently they interact with it this way, the more opportunities they have for such learning. This can help users learn and apply new vocabulary.

5. Conclusions

Derived from the findings are the following conclusions:

1. Junior high school students do use ChatGPT in language learning. The tool is useful in accomplishing language learning tasks because it provides immediate feedback to students' queries at all times.
2. ChatGPT is a satisfying tool that caters the needs of learners. However, mastery of the tool can be hindered by financial reasons. These include limited access to personal gadgets; unstable personal internet connection; and insufficiency of computer laboratory in schools. Due to this, students cannot regularly use ChatGPT which then resulted to less familiarity.
3. ChatGPT is an effective tool in improving grammar, particularly grammatical accuracy and complexity of sentence structure. Meanwhile, the tool has less impact in improving total words and unique words. Thus, ChatGPT helps refine sentence structures and correct grammatical errors but its impact on vocabulary development is limited.
4. ChatGPT-assisted learning enhances grammatical accuracy and sentence structure more than the traditional language instruction but would still require additional vocabulary-focused instructional strategies.
5. Male students have greater vocabulary diversity than female learners. The former prioritizes expressiveness while the latter prefers accuracy and clarity.

6. Recommendations

The following recommendations are derived from the conclusions given previously.

1. Students should be encouraged to use ChatGPT for self-directed learning and actively work on vocabulary enrichment through other sources.
2. Curriculum experts, policy makers, and the Department of Education should also consider regulating AI-assisted tools like ChatGPT in schools. This can be done by investing in additional computer units and stable internet connection.
3. Educators should consider integrating ChatGPT into language instruction. Teachers can utilize ChatGPT as a writing aid to help students refine their grammatical and sentence construction skills while encouraging independent learning.
4. School heads are encouraged to conduct an orientation for teachers on the proper use of ChatGPT in classroom instruction. This will guarantee the appropriate integration of ChatGPT.

5. Future researchers should investigate AI-driven strategies for vocabulary development and explore how ChatGPT and similar tools can be further optimized. Possible suggested areas include: organization of ideas; writing cohesion or coherence; and orthography. This in-depth investigation can further provide a more comprehensive view of the effectiveness of ChatGPT in language learning.

References

1. Abernathy, P. C. (2024). *AI Tools as Supplementary Support in Language Acquisition*.https://doi.org/10.20944/preprints202407.2213.v1
2. Ahmed, R. (2024). Exploring ChatGPT Usage in Higher Education: Patterns, Perceptions, and Ethical Implications Among University Students. *Journal of Digital Learning and Distance Education*, *3*(6), 1122–1131. https://doi.org/10.56778/jdlde.v3i6.363
3. Al-Durayhim, S. bin A. bin A. (2023). Enhancing Grammar Competency Through WhatsApp: An Experimental Study in Saudi Arabian Education. *Theory and Practice in Language Studies*, *13*(10), https://doi.org/10.17507/tpls.1310.20
4. Al-Garaady, J., & Mahyoob, M. (2023). ChatGPT's capabilities in spotting and analyzing writing errors experienced by EFL learners. *Arab World English Journals, Special Issue on CALL*, (9).
5. Avalur, R., Ks, S., & Sundaram, K., (2023). Review of ChatGpt. *International Research Journal of Modernization in Engineering Technology and Science*. https://doi.org/10.56726/irjmets43072
6. Baskara, F. R. (2023). Integrating ChatGPT into EFL writing instruction: Benefits and challenges. *International Journal of Education and Learning*, *5*(1), 44-55.
7. Bhandari, P. (2023, June 22). *Ethical Considerations in Research | Types and Examples*.https://www.scribbr.com/methodology/research-ethics/
8. Borhan, N. (2023). A Review Study About Gender Differences in Expressive Language: Spoken and Written Language Differences. https://doi.org/10.51503/gpd.1231592
9. Chhabriya, M., Babu, N. N., & Kumari, A. C. (2024). *AI's Corrective Feedback vs. Traditional Recasts: which fosters better grammar accuracy?*https://doi.org/10.1109/idicaiei61867.2024.10842936
10. Davoodi, L., & Baharlouie, R. (2017). The Interaction of Gender with Text Enhancement and Metacognitive Grammar Instruction on Learning and Recall of English Grammar. *International Journal of Foreign Language Teaching and Research*, *5*(18), 113–130.
11. Dörnyei, Z., & Ushioda, E. (2021, April 2). Teaching and Researching Motivation. https://doi.org/10.4324/9781351006743
12. Dwi Putri, A. L., Tambusai, A., & Hasibuan, A. L. (2022, June 30). Students' Difficulties in Acquiring English Writing. *International Journal of Educational Research Excellence (IJERE)*, *1*(1), 19–26. https://doi.org/10.55299/ijere.v1i1.21
13. Fadhilah, N. F., Meuthia, N. R. F., & Ferdawati, N. (2022). Measuring Usability of Academic Information System Using Use Questionnaire: Case Study of Padang State Polytechnic. *Malaysian Journal of Business and Economic/Malaysian Journal of Business and Economics*, *9*(2). https://doi.org/10.51200/mjbe.v9i2.3932
14. Feng, Y., Qiang, J., Li, Y., Yuan, Y., & Zhu, Y. (2023). Sentence simplification via large language models. https://doi.org/10.48550/arXiv.2302.11957
15. Gautam, P. (2019, December 8). Writing Skill: An Instructional Overview. *Journal of NELTA Gandaki*, *2*, 74–90. https://doi.org/10.3126/jong.v2i0.26605
16. Giray, L. (2024). Enhancing Communication with ChatGPT: A Guide for Academic Writers, Teach

ers, and Professionals. *Journal of the Practice of Cardiovascular Sciences, 10*(2), 113–118. https://doi.org/10.4103/jpcs.jpcs_27_24

17. Gupta, M. R. (2024). ChatGPT-A Generative Pre-Trained Transformer. *International Journal of Advanced Research in Science, Communication and Technology*. https://doi.org/10.48175/ijarsct-15087
18. Ingale, S. D. (2019). Techniques in Teaching Writing Skills. *The IUP Journal of English Studies, 12*(1), 16-20.
19. Jashari L., & Fojkar, M. (2019, December 12). Teachers' Perceptions of Developing Writing Skills in the EFL Classroom. *ELOPE: English Language Overseas Perspectives and Enquiries*, *16*(2), https://doi.org/10.4312/elope.16.2.77-90
20. Jayavalan, K., & Razali, A. B. (2020). Effectiveness of Online Grammar Checker to Improve Secondary Students' English Narrative Essay Writing. *International Research Journal of Education and Sciences (IRJES), 2*(1).
21. Jiao, W., Wang, W., Huang, J. T., Wang, X., Shi, S., & Tu, Z. (2023). Is ChatGPT a good translator? Yes! with GPT-4 as the engine. https://doi.org/10.48550/arXiv.2301.08745
22. Khzouz, A., Salaita, O., Al-Issawi, J. M., AlTaher, B. B., & Alkhenizan, H. H. (2024). Exploring ChatGPT Practices and User Experiences in Learning English Skills. *Theory and Practice in Language Studies, 14*(12). https://doi.org/10.17507/tpls.1412.31
23. Kim M. (2024). Unveiling the e-Servicescape of ChatGPT: Exploring User Psychology and Engagement in AI-Powered Chatbot Experiences. *Behavioral sciences (Basel, Switzerland), 14*(7), 558. https://doi.org/10.3390/bs14070558
24. Kim, S., Shim, J., & Shim, J. (2023). A study on the utilization of OpenAI ChatGPT as a second language learning tool. *Journal of Multimedia Information System*.
25. Kohnke, L., Moorhouse, B. L., & Zou, D. (2023, April 3). ChatGPT for Language Teaching and Learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868
26. Koraishi, O. (2023). Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education and Technology, 3*(1).
27. Kronberg, S. (2024, January). The Effect of Gender on ChatGPT Use: A Technology Acceptance Model 3 (TAM3) study. In *Proceedings of Umea's 27th Student Conference in Computing Science USCCS 2024* (p. 57).
28. Labicane, G. E. M., & Oliva, R. M. M. (2022, June 27). Common Errors in Composition Writing by College Students. *Language Literacy: Journal of Linguistics, Literature, and Language Teaching*, *6*(1), 1–6. https://doi.org/10.30743/ll.v6i1.4975
29. Lalira, J. E., Pangemanan, Y. A. T., Scipio, J. E., Lumi, S., Merentek, T. Ch., & Tumuju, V. N. (2024). Evaluating the Impact of AI Tools on Grammar Mastery: A Comparative Study of Learning Outcomes. *VELES (Voices of English Language Education Society), 8*(3). https://doi.org/10.29408/veles.v8i3.27856
30. Lam, C., & Wolfe, J. (2022). An Introduction to Quasi-Experimental Research for Technical and Professional Communication Instructors. *Journal of Business and Technical Communication, 37*(2), 174–193. https://doi.org/10.1177/10506519221143111
31. Liu, J. W. S. (2024). ChatGPT: perspectives from human–computer interaction and psychology. *Frontiers in Artificial Intelligence*, *7*. https://doi.org/10.3389/frai.2024.1418869
32. Lund, A.M. (2001). Measuring the Usability with USE Questionnaire, STC Usability. *SIG Newsletter, 8:2.*

33. Magruder, M. L., Delanois, R. E., Nace, J., & Mont, M. A. (2023, November). ChatGPT and Other Natural Language Processing Artificial Intelligence Models in Adult Reconstruction. *The Journal of Arthroplasty*, *38*(11), 2191–2192. https://doi.org/10.1016/j.arth.2023.06.030
34. Manoli, P., Athanassopoulos, S., Gouvi, M., Lavidas, K., & Komis, V. (2023). The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom. *Advances in Mobile Learning Educational Research*, *3*(2), 818–824. https://doi.org/10.25082/amler.2023.02.009
35. Masoudi, H. (2024). *Effectiveness of ChatGPT in Improving English Writing Proficiency among Non-native English Speakers*. https://doi.org/10.59992/ijesa.2023.v3n4p
36. Menon, D., & Shilpa, K. (2023). "Chatting with ChatGPT": Analyzing the factors influencing users' intention to Use the Open AI's ChatGPT using the UTAUT model. *Heliyon*, https://doi.org/10.1016/j.heliyon.2023.e20962
37. Mustafa, F., Kirana, M., & Bahri Ys, S. (2019, February 8). Errors in EFL writing by junior high students in Indonesia. *International Journal of Research Studies in Language Learning*, *5*(5). https://doi.org/10.5861/ijrsll.2016.1366
38. Niyozov, N., Bijanov, A., Ganiyev, S., & Kurbonova, R. (2023). The pedagogical principles and effectiveness of utilizing ChatGPT for language learning. In *E3S Web of Conferences* (Vol. 461, p. 01093). EDP Sciences.
39. Pang, Q., Zhang, M., Yuen, K. F., & Fang, M. (2024). When the winds of change blow: an empirical investigation of ChatGPT's usage behaviour. *Technology Analysis & Strategic Management*, 1–15. https://doi.org/10.1080/09537325.2024.2394783
40. Postigo-Zumarán, J. E., Nova Revilla, L. J., Arias-Chávez, D., & Cangalaya Sevillano, L. M. (2024). Chat-GPT in Higher Education, a Comparative Analysis within Business Students at Universities in two Peruvian Cities. *International Journal of Religion*, *5*(12), 2060–2071. https://doi.org/10.61707/3m2s5y94
41. Reichardt, C. S., Storage, D., & Abraham, D. (2023). Quasi-Experimental Research. In *Cambridge University Press eBooks* (pp. 292 313). https://doi.org/10.1017/9781009010054.015
42. Sekewael, M., & Anaktototy, K. (2024). *Chat-GPT Feedback: Fostering Growth in English Writing Abilities among Students*. https://doi.org/10.5772/intechopen.1007287
43. Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023, September 21). Assessing the Usability of ChatGPT for Formal English Language Learning. *European Journal of Investigation in Health, Psychology and Education*, *13*(9), 1937–1960. https://doi.org/10.3390/ejihpe13090140
44. Siregar, F. H., Hasmayni, B., & Lubis, A. H. (2023). The Analysis of Chat GPT Usage Impact on Learning Motivation among Scout Students. *International Journal of Research and Review*, *10*(7), 632-638.
45. Tan, X. (2023, July 30). The Impact of ChatGPT on Education and Future Prospects. *Highlights in Science, Engineering and Technology*, *61*, 138–143. https://doi.org/10.54097/hset.v61i.10285
46. Urbano C.M., Gumangan M.A., Gustilo L., & Capacete, M.P. (2021, June 30). Reading and Writing Needs of Senior High School Students: The Case of Filipino Students in the Philippines. *Modern Journal of Studies in English Language Teaching and Literature*, *3*(1), 140–166. https://doi.org/10.56498/31202154
47. Ventayen, R. (2020, December). Graduate Students' Perspective on the Usability of Grammarly® in one ASEAN State University. *The Asian ESP Journal*, *14*(7.2).

48. Wahyuni, D. S. (2022, April 11). Integrated Classroom-Chatbot Experience: An Alternative Solution for English as Foreign Language Learners. *English Language Education and Current Trends (ELECT)*, 63 68. https://doi.org/10.37301/elect.v1i1.36
49. Yan, D. (2023). Impact of ChatGPT on learners in an L2 writing practicum: An exploratory investigation. Education and Information Technologies, 1-25. https://doi.org/10.1007/s10639-023-11742-4
50. Yassin, A., & Bashir, A. (2024). Student Satisfaction with The Use of Chat-GPT as A Learning Resource. *Journal of Education Technology*. https://doi.org/10.58740/vocational.v1i1.247
51. Yuliawati, L. (2021). The Effects of Vocabulary and Grammar Mastery on Students' Writing Skill. *Journal of English Language Learning*, *5*(2).
52. Zhai, X. (2022). ChatGPT User Experience: Implications for Education. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4312418