KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

Investigation of AI Grammar Checkers on Grammar Learning and Students' Perception in L2 Writing Context

Hea-Suk Kim (Seoul Women's University)

Eunhye Song (Soongsil University)

Hea-Suk Kim
(1st author)
Associate Professor,
Department of General Education,
Seoul Women's University
Email: shskim@swu.ac.kr

Eunhye Song
(corresponding author)
Associate Professor,
Department of General English Education,
Baird College of Education,
Soongsil University
Email: edujws@hanmail.net

ABSTRACT

Kim, Hea-Suk and Eunhye Song. 2024. Investigation of AI grammar checkers on grammar learning and students' perception in L2 writing context. *Korean Journal of English Language and Linguistics* 24, 531-553.

This study investigates three distinct AI Grammar Checkers (AI GCs) in English writing to assess their impact on grammar enhancement and learners' perceptions. The study included 125 students divided into three experimental groups, each employing SpellCheckPlus, Virtual Writing Tutor, or Grammarly. The first question pertains to how three AI GCs contribute to grammar learning. While all groups demonstrated improved grammar after using AI GCs, there were no significant discrepancies in the post-test between the groups. That is, no significant variations in improvement were observed across the different GCs. The second question explored learners' perceptions of the three AI GCs. Results from a questionnaire revealed that all groups exhibited favorable aspects towards the AI GCs, highlighting benefits in active involvement, error identification, convenience of use, easiness, and others. However, aspects like interests, motivation to write, and comprehensive feedback demonstrated no substantial changes before and after AI GCs use. Regardless of the benefits and drawbacks of the three AI GCs of the study, Grammarly's ability to enhance learning stood out in particular in terms of error identification, writing feedback, and user-friendliness. All groups presented diverse opinions on drawbacks, including accuracy issues, inconvenience, feedback inadequacy, and tool dependency for all three programs.

1. Introduction

As the world becomes global, English has been emphasized as an important tool for communication. In addition, the importance of English writing is being emphasized more in practical and academic aspects. Accordingly, instructor feedback has been considered an essential factor in improving English writing (Hyland 2003, Hyland and Hyland 2006, Kang 2008, Lindblom-ylanne et al. 2006). However, in most English writing classes at most universities in Korea, there could be educational limitations as it could be a burden for a professor to give detailed feedback on essays' content, organization, and grammatical aspects to many students due to time limitations and efforts.

Accordingly, within the revision phase of process writing, peer feedback and self-feedback were implemented in conjunction with the feedback provided by the instructor. Researchers and educators are increasingly harnessing technology, particularly AI GCs, to enhance the quality of L2 writing (Ghufron and Rosyida 2018, Im 2021, O'Neill and Russell 2019, Park and Yang 2020).

Previous studies (Link et al. 2014, Min 2020) have suggested that it was more effective for teachers to give content and structure-oriented feedback, whereas using online grammar checkers for simple linguistic and grammatical errors was more effective. It has been suggested that AI GCs can help reduce the burden on instructors to provide real-time feedback, especially on linguistic aspects of many students' writing (Park and Yang 2020). In addition, using grammar checkers is beneficial for students in self-directed grammar learning (Park 2019).

The following explains the previous studies on AI GCs. First, among many AI GCs, Grammarly was applied the most for learners' English writing feedback to investigate grammar error analysis (Agustin and Wulandari 2022, Im 2021, Moon 2011) and perception (Wang 2023). Second, the AI GCs' feedback was compared and analyzed with the teacher feedback (Ghufron and Rosyida 2018, Park and Yang 2020), self-directed (Long 2022, Song and Kim 2021), or self-directed and peer-feedback (Min 2020). Finally, the effectiveness of error analysis of different AI GCs was compared and analyzed (Badi et al. 2020, John and Woll 2020, Ranali 2018, Yang 2018, Zhang and Hyland 2018).

Most of the studies related to AI GCs have proven their effectiveness using Grammarly, but the feedback function of the free version is limited compared to the paid version. Existing studies using other AI GCs, such as SpellCheckPlus (Yang 2010 2018), Virtual Writing Tutor (Badi et al. 2020, John and Woll 2020), or Pigai (Wang 2023, Zhang and Hyland 2018) have also demonstrated the usefulness of AI GCs. In this way, studies have been conducted to investigate the effectiveness of grammar error analysis using grammar checkers. Nevertheless, the evidence bases for using various grammar checkers is limited. It is difficult to find comparative studies on the effectiveness of the three most widely used AI grammar checkers. Therefore, this study aims to conduct a comparative analysis of three educational tools, each with its own unique features.

Accordingly, this paper aims to compare and analyze three AI GCs that have proven individually useful among the most commonly used AI GCs in L2 writing classes. Therefore, in this study, we would like to compare and analyze whether there is an AI grammar checker that is more helpful in improving English grammar learning in an EFL environment and investigate how learners perceive it. Through this, we would like to derive an effective utilization plan for the AI grammar checker to improve grammar in L2 writing classes. The research questions are below:

1. To what extent does the utilization of three distinct AI GCs (SpellCheckPlus, Virtual Writing Tutor, Grammarly) contribute to enhancing students' grammatical proficiency?
2. How do learners perceive using three AI GCs (SpellCheckPlus, Virtual Writing Tutor, Grammarly) during the English writing course?

2. Literature Review

2.1 Enhancing Writing Skills through Feedback in Process Writing

English writing teaching methods can be classified into product- and process-based writing. In particular, process-oriented writing emphasizes the writing process (Hedgecock 2005). Tompkins (2008) explains the five stages of the writing process as pre-writing, drafting, revising, editing, and publishing. In particular, instructors' feedback on the revision stage is an essential factor in improving writing in process-oriented writing teaching methods (Hyland 2003, Hyland and Hyland 2006, Kang 2008, Lindblom-ylanne et al. 2006). In addition, process-based writing emphasizes not only the content but also the linguistic accuracy in each stage of the writing process (Hedgecock 2005). In particular, EFL learners need grammatical accuracy and feedback on grammatical errors in the text written to accurately convey the meaning of the text (Ferris 2003, Hyland and Hyland 2006). Ferris and Roberts (2001) emphasized that the group that received grammar-related feedback in the English composition class showed better self-correction skills in the writing class.

In this way, corrective feedback regarding language and the content presented by the instructor for the student's writing is important (Van Beuningen et al. 2012). Above all, previous studies have presented results that favor teacher feedback (Kang 2008, Lindblom-ylanne et al. 2006). However, it is not easy for one instructor to provide real-time feedback to many students in many classes simultaneously in English writing classes of university. In addition, it is a process that takes considerable time and effort to present detailed feedback on content and linguistic aspects (Junqueira and Payant 2015).

Accordingly, various feedback methods, such as self-feedback and peer feedback, were attempted in English education to reduce the work burden on instructors. Participating in student-centered feedback activities has been shown to maximize learning effects and improve learners' autonomy rather than just receiving instructor feedback passively (Lee et al. 2019, Mansoor 2009, Park and Kim 2016). Previous studies have shown that peer feedback effectively improved English writing and increased responsibility and confidence in writing (Kim 2013, Kim and Kim 2013, Mansoor 2009, Park and Kim 2016). In addition, studies have suggested that self-feedback helps improve meta-cognitive skills in English writing (Park and Kim 2016) and is more effective in grammar or linguistic aspects than in content aspects (Chung 2018, Joo 2020, Lindblom-ylanne et al. 2006).

2.2 AI-based Grammar Checkers on Grammar Learning in L2 Writing

While attempting various feedback methods such as self-feedback and peer feedback in English writing education, studies have recently been conducted to investigate the effectiveness of feedback using AI GCs along with developing computer and AI technologies. These AI GCs immediately and automatically present corrective feedback on grammatical and punctuation errors in real-time for written English sentences or essays. Previous studies have presented many types of AI GCs or reviewed their effectiveness, such as Grammarly, SpellCheckPlus, Virtual Writing Tutor, ClearEdits, Ginger, ProWritingAid, and GrammarBase (Al-Ahdal 2020, Dale 2016, Sahu et al. 2020). Chun et al. (2021) argued that as there are various AI GCs, choosing a program can increase the effectiveness of education for an appropriate purpose.

First, existing studies have shown that the use of AI GC, specifically Grammarly, is useful for error analysis of writing and improvement of English grammar accuracy (Austin and Wulandari 2022, Im 2021, Moon 2021, Nazari et al. 2021). Using Grammarly significantly improved English grammar related to writing by Korean college students, which helped them learn English and suggested positive opinions that they wanted to continue using it

(Im 2021). It was also helpful for searching for errors in articles, punctuation, and singular/plural, and helped improve accuracy by searching for about 65% of errors in writing and presenting accurate alternative forms (Austin and Wulandari 2022, Hadiat et al. 2022, Moon 2021). In addition, using Grammarly showed significantly higher behavioral, emotional, and cognitive engagement, self-efficacy for writing, and positive emotions. Therefore, the use of Grammarly was effective in improving confidence in writing (Hadiat et al. 2022) and learning behavior through formative feedback and assessment for non-native students' academic writing (Nazari et al. 2021).

Second, investigations assessed the effectiveness of AI-generated feedback compared to feedback provided by teachers or students. The comparative effectiveness of Grammarly feedback and instructor feedback yielded differentiated results. Specifically, Grammarly feedback exhibited greater efficacy than instructor feedback regarding vocabulary, punctuation, and grammar-related feedback about English composition (Ghufron and Rosyida 2018, Park and Yang 2020). Nonetheless, the feedback from the instructor was more helpful than Grammarly for the feedback related to the contents and organization of English composition (Ghufron and Rosyida 2018). Park and Yang (2020) applied instructor feedback first and Grammarly feedback second. As a result, Grammarly feedback presented the most feedback on conciseness, spelling, and English grammar (76.57%) errors. Even though participants in the study trusted the instructor's feedback, they preferred using AI GC because of quick real-time feedback. In addition, very positive survey results were presented on the usefulness and satisfaction of AI GC feedback. On the other hand, Grammarly feedback and self-feedback differed significantly in the effectiveness of grammar learning (Long 2022, Song and Kim 2021). In the study by Song and Kim (2021), Grammarly feedback showed that Korean college students' grammar improved afterward. On the other hand, there was no difference in English grammar improvement between the group using only Grammarly feedback and the group that combined Grammarly and self-feedback. However, the survey answered that combining Grammarly and self-feedback helped to learn English grammar. Long (2022) found that while Grammarly was useful in identifying specific grammatical errors, there was no statistically significant difference between Japanese college students using Grammarly and those relying on self-correction feedback.

Some studies have focused on learners' perceptions of various feedback methods (AI GCs, teachers, self-feedback). Most studies have shown that each method of feedback on English composition has advantages and disadvantages, so it is preferred to combine teacher feedback or student-centered feedback with AI GCs' feedback (Min 2020, O'Neill and Russell 2019). O'Neill and Russell (2019) reported that the experimental group using both teacher feedback and Grammarly programs was more satisfied than the control group, who only received teacher feedback in English composition. Specifically, the feedback presented in Grammarly was helpful and easy to understand, so self-directed modifications to English composition were possible, and confidence in writing was improved. In addition, Min (2020) applied step-by-step feedback, such as self-feedback for step one, peer feedback for step two, and Grammarly feedback for step three. She reported that each feedback has pros and cons. In addition, she also suggested that if students use only feedback from the AI GCs, they may rely on AI and be passive learners. Therefore, combining AI GCs and student-centered feedback, such as self or peer feedback, would be more effective. According to Zhang et al. (2020), participants used Grammarly more often because it was easy to use, whereas they preferred face-to-face tutors' feedback for various needs related to writing. Therefore, combining Grammarly and face-to-face tutor feedback could be more helpful for writing. Zhang and Hyland (2018) found that Chinese university students immersed in the writing and modification stages prefer AI Pigai feedback to professor feedback because they can receive immediate feedback. On the other hand, a study by Wang (2023) found that Chinese students with upper and lower levels of English prefer to receive both AI Pigai and instructor feedback.

Finally, studies have analyzed the feedback function of different AI GCs other than Grammarly, which are

widely used. Yang (2018) primarily used the AI GC, SpellCheckPlus, available in almost all web browsers. The results showed no improvement in grammar errors for Chinese students' English writing in the first task. However, grammar errors were significantly reduced after utilizing the AI SpellCheckPlus in the second and third tasks. In addition, participants showed a relatively positive attitude toward the AI GC. Badi et al. (2020) showed that Omani college students used the AI GC, Virtual Writing Tutor, in academic writing and improved English essay writing in five areas (task attachment, organization, grammar, vocabulary, and punctuation). In addition, a study was conducted to compare and analyze the feedback functions of several AI GCs. Ranalli (2018) emphasized that AI Grammarly provides much more grammatically specific feedback than Criterion in ETS and supports various feedback delivery modes and different platforms associated with writing. John and Woll (2020) used three AI GCs (Virtual Writing Tutor, paid version of Grammarly, and Microsoft Word's grammar-checking function) to investigate function differences in error detection and correction. As a result, Virtual Writing Tutor and the paid version of Grammarly were better at error detection and correction than Microsoft Word's grammar-checking function.

Meanwhile, some studies have suggested some limitations of AI GCs for error analysis in English writing. Park (2019) compared and analyzed the feedback of native speakers and Grammarly paid version regarding students' essays, showing that some parts did not match the feedback. It was also suggested that the Grammarly paid version analyzes feedback mainly focusing on specific errors. Im (2021) reported that as Grammarly feedback was presented only in English, it was difficult for EFL students with low English levels to understand the suggested feedback.

Despite these limitations of AI GCs, the most significant advantage of AI GCs is possible to give feedback simultaneously to many students, whereas it is difficult for instructors to do so (Badi et al. 2020, Park 2019, Park and Yang 2020, Zhang and Hyland 2018). As a result, using AI GCs can reduce the burden of feedback on instructors, especially in the early stage of English writing (Park 2019). Furthermore, AI GCs have been shown to help students improve their ability and confidence to correct errors in English writing and grammar independently rather than relying solely on instructors' feedback because detailed feedback from AI GCs was helpful (Badi et al. 2020, Hadiat et al. 2022, Im 2021, O'Neill and Russell 2019, Park 2019, Park and Yang 2020). As previous studies have shown, the importance of feedback in process-oriented writing cannot be overlooked, and it can be seen that combining correction of teacher/self-correction and AI GCs is more effective. Accordingly, the present study investigates the most effective AI GC in improving English grammar in English writing and the EFL learners' perceptions among three different AI GCs.

3. Methodology

3.1 Participants

The participants of this study are Korean university students taking classes in required English writing classes at universities in Seoul (50 male students and 75 female students). 125 Participants were primarily first-year students randomly assigned to three experimental groups to use three different AI GCs (See Table 1). They studied the same English writing class for one semester (about 15 weeks) and learned from the same instructor. No students had official English test scores, and the students' majors were diverse. Other previous experience information is as follows: studying in the English-speaking countries and using the AI grammar checker (See Table 1).

Table 1. Demographic Information about Three Experimental Groups

Group	AI Tools	Major	Students (%)	Male	Female	Num. of students studying abroad	Num. of students using AI grammar checkers
1	SpellCheckPlus	College of Business and Law	44 (35.2%)	18	26	0	1 (Ginger)
2	Virtual Writing Tutor	College of Natural Science & Humanities	42 (33.6%)	16	26	2 (less than 3 months) (New Zealand, Canada)	0
3	Grammarly	College of Humanities, Business, & Engineering	39 (31.2%)	16	23	1 (less than 3 months)	1 (MS Word Document)

In experimental group 1, 44 students were asked to use the AI-based SpellCheckPlus free version. Out of these students, 18 were male, and 26 were female. Most students belonged to the College of Business and Law. No students within this group had encountered studying abroad. However, one students had previous experience using AI GC, specifically tool like Ginger.

Within experimental group 2, 42 students used the AI-powered Virtual Writing Tutor free version. Among these students, 16 are male, and 26 are female. This group represents various majors, encompassing disciplines like the College of Natural Science & Humanities. Among the students in this group, two had prior exposure to studying abroad for less than three months.

Experimental group 3 comprised 39 students using the AI-based Grammarly free version. Among these students, there are 16 males and 23 females. Their majors differed based on the College of Humanities, Business, and Engineering. Within this group, one student has experienced studying abroad, and one student has utilized MS Word Document to check grammar.

To confirm that participants across the three groups were homogeneous, a one-way ANOVA test was used to determine if there were any significant differences between the groups based on their TOEIC Mock scores. The findings indicated that Group 1 had a mean of 32.55 (SD = 10.66), Group 2 scored an average of 29.33 (SD = 12.20), and Group 3 had an average score of 26.87 (SD = 9.31). These results suggest there was no significant variance between the groups ($F = 2.88$; $p > 0.5$).

3.2 Teaching Procedures

The research procedures for this study are as follows. The same professor taught a mandatory English course for college liberal arts for first graders. Classes were held once a week for 100 minutes for 15 weeks a semester and conducted through Zoom, a virtual online conference, during pandemic. The contents of the class focused on writing English essays (opinion, argumentative, and cause & effect essays). Participants searched, summarized, paraphrased, and cited data (APA style) related to writing topics in the textbook.

Table 2. Weekly Schedule and Data Collection Procedure

Weeks	Details
W1	1. Course orientation 2. Mock TOEIC reading test for checking English level
W2	1. Pre-English grammar test & Pre-questionnaire 2. Instructions for the use of AI GCs and self-check based on the list
W3-5	1. Lectures for English writing a paragraph and an essay

W6-7	1. Writing an essay (1) (Opinion)
	  1) Write the first draft.
	  2) Self-check feedback
	    - Self-check the first draft based on the list provided by the Professor and correct the mistakes.
	  3) AI GCs' feedback
	    - Next, use a different AI GC (Depending on each experimental group - SpellCheckPlus, Virtual Writing Tutor, or Grammarly)
	    - Download or capture the AI feedback suggestion.
	    - Correct errors, and revise the draft & upload it on LMS.
	  4) Professor feedback
	  - Professor gives feedback to students' writing.
	  - Students revise the final version of the writing.

W8-9	1. Mid-term Exam 2. Writing an essay (2) (Opinion) - the above same process of self-check, AI GC, and Professor's feedback
W10-11	1. Writing an essay (3) (Cause and effect) - the above same process of self-check, AI GC, and Professor's feedback
W12-13	1. Writing an essay (4) (Argumentative) - the above same process of self-check, AI GC, and Professor's feedback 2. Post-English grammar test & post-questionnaire
W14-15	1. Review & Final Exam

Specifically, as shown in Table 2, a pre-simulated TOEIC reading test was conducted in one week to confirm the orientation and English skills between groups. In the second week, preliminary English grammar tests and preliminary surveys were conducted to diagnose the improvement of English grammar in research participants. In addition, specific AI GCs (SpellCheckPlus, Virtual Writing Tutor, Grammarly) were randomly assigned to each group, and the self-feedback list was described. Lectures on writing English paragraphs and essays were given between the third and fifth weeks. However, there was no lecture on grammar during the class. First, the first draft of students' essay writing (1) was uploaded to the school's LMS in the sixth and seventh weeks. Second, students conducted self-feedback and revised the first draft of self-feedback. Third, depending on the group, students used a specific AI GC. The feedback presented by the AI GC was revised and uploaded to the LMS (AI GC's first draft). Finally, the instructor gave feedback on the second draft revised by self- and AI feedback, and students uploaded the final version of essay writing (1) to the LMS. Based on the same process of self-feedback, AI GC feedback, and instructor feedback, the following was conducted:

  1) Midterms and essay writing (2) in Week 8 and Week 9;
  2) Essay writing (3) in Week 10 and Week 11;
  3) Essay writing (4) in Week 12 and Week 13.

In Week 13, a post-test of English grammar and post-survey were conducted. In weeks 14 and 15, courses were reviewed and included the final exam.

3.3 Instruments

The research data collection tools for this study are as follows. First, at the beginning of the semester, a certified English mock TOEIC reading test was conducted to understand the level of students.

3.3.1 Pre-simulated TOEIC English reading test

At the beginning of the semester, a mock TOEIC English reading test was conducted to investigate the difference in English skills among the three experimental groups. The test was conducted as an online test by accessing the university's LMS during class. Twenty-five questions (four points per question made 100 points) were conducted for about 15 minutes.

3.3.2 Pre- and post-English grammar tests

An English grammar test was conducted at the beginning of the semester through the university's LMS to investigate whether there was a difference in English grammar proficiency before and after using AI-based grammar checkers to be used in this study. A total of 40 test questions (100 points with 2.5 points per question) were given, and the time was 20 minutes. There were 35 questions related to English grammar, and five questions related to capital letters and punctuation were included. On the LMS, the automatic system randomly changed the question number and viewing order, and the answer was not disclosed.

3.3.3 Free version of AI-based grammar checkers

Existing studies have examined several AI-based grammar checkers (Al-Ahdal 2020, Dale 2016, Sahu et al. 2020, Chun et al. 2021). Among them, three free version AI GCs were used in this study to provide feedback on written English text.

The first AI GC is SpellCheckPlus free version, which is very easy to use and a free version available in most web browsers (Yang 2010, 2018). As soon as text is entered without membership, it immediately presents real-time feedback on the wrong parts of Word choice, word form, and spelling (Figure 1).
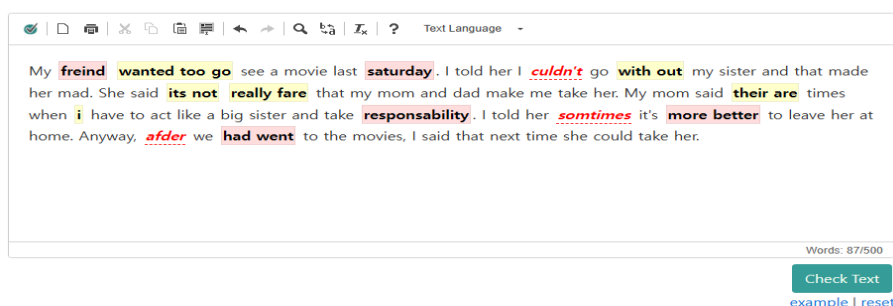

Figure 1. A Screen Shot of the SpellCheckPlus Feedback

The second AI GC is Virtual Writing Tutor (Figure 2). It can be used immediately without membership registration and is originally produced for ESL learners. Various functions are provided free of charge. It presents linguistic feedback such as 'Check Grammar' and 'Check Vocabulary' (Badi et al. 2020) (Figure 3). In addition,

18 very detailed functions related to the content of the writing, such as 'Check Topic Sentence', 'Check Paraphrase', 'Check Cohesion', and 'Check Essay' can be selected, and immediately feedback is presented.
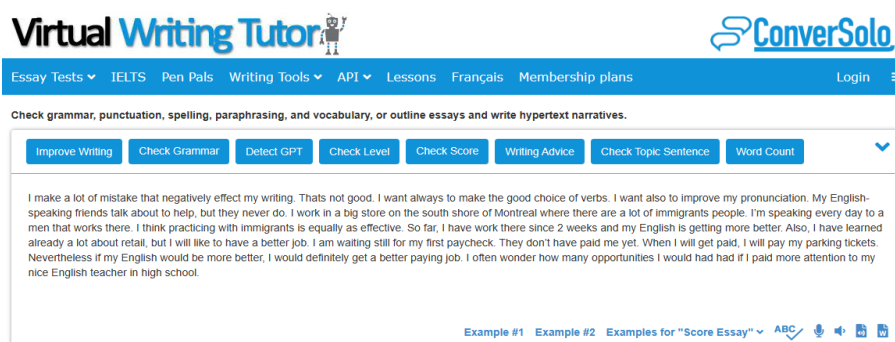

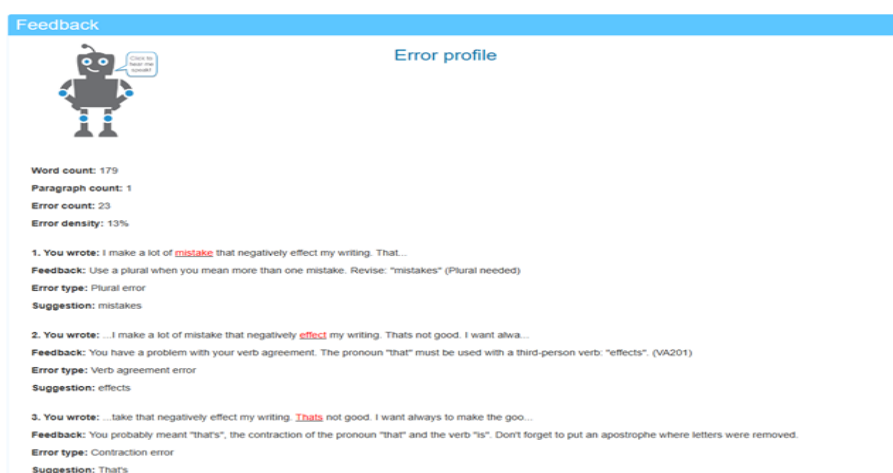**Figure 2. A Screenshot of the Virtual Writing Tutor**

Figure 3. A Screenshot of the Virtual Writing Tutor Feedback

The last AI GC is Grammarly program (Figure 4), which is a widely known program among AI GCs. Feedback is presented in real-time, but membership is required even to use the free version. The free version provides feedback only on general correctness, so feedback is limited. The free version presents only an underline of the wrong parts related to clarity, engagement, and delivery. The correct expression or detailed feedback on these parts is presented only in the paid version.

3.3.4 Pre- and post-questionnaires

Participants in the study used a free version of one specific AI GC, depending on the experimental group. In order to investigate students' changes of perceptions towards AI GC, 20 multiple-choice questions were included in the pre-and post-survey. In addition, the post-survey included two open-ended questions that wrote the specific AI GC's advantages and disadvantages to investigate more detailed opinions.
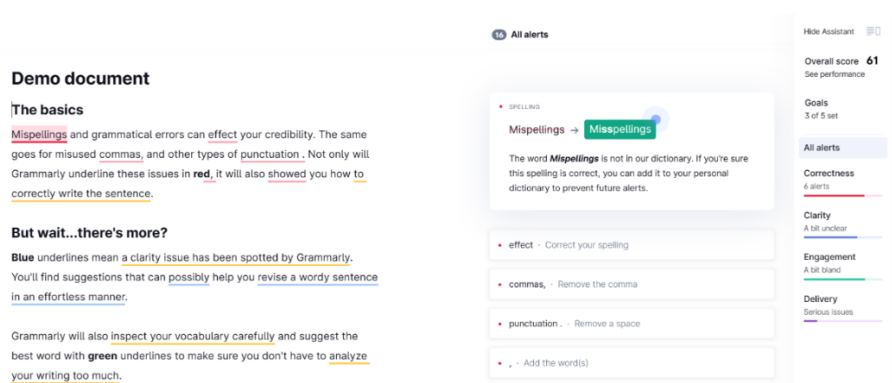
Figure 4. A Screenshot of the Grammarly Feedback

3.4 Data Analysis

In order to facilitate the data analysis process, the present study used several assessment instruments, including the TOEIC examination, grammar assessment, and questionnaires. The TOEIC examination was initially employed to establish whether similar English levels existed in pre-test scores among the three distinct groups. Subsequently, pre- and post-grammar tests were conducted to investigate significant differences in grammatical proficiency among the three groups. An analysis of variance (ANOVA) was conducted to ascertain intergroup distinctions. Furthermore, within each group, a paired-sample *t*-test was conducted to identify potential notable improvements in grammatical proficiency by assessing differential performance across the means of pre-test and post-test measurements. All the data was analyzed using both descriptive and inferential statistical methods by SPSS 23.0.

Surveys were administered both before and after the study to assess changes in learners' viewpoints regarding AI GCs. The resulting shifts in attitudes within each group were evaluated through paired-sample *t*-tests. Lastly, the qualitative components of the study encompassed the analysis of open-ended responses. Responses from participants within each group were aggregated and categorized by thematic correspondence, facilitating a comprehensive exploration of shared viewpoints and insights.

4. Results and Discussion

4.1 Effects of Three AI Grammar Checkers on Grammar

4.1.1 Changes in grammar learning using AI grammar checkers

This study aimed to investigate the impact of using three different AI GCs (SpellCheckPlus, Virtual Writing Tutor, Grammarly) on grammar knowledge while engaging in English writing activities. The main objective was to determine how these AI GCs influenced participants' understanding and usage of grammar rules. To achieve this, the study involved three groups, SpellCheckPlus, Virtual Writing Tutor, Grammarly, using a different AI GC.

Before the study began, all participants from the three groups took a grammar test to establish their initial grammar knowledge. After completing the pre-test, each group engaged in English writing task each week using their respective AI GC. During this phase, participants had access to the AI GCs to help them identify and correct

grammar errors in their writing process. Once the writing activities were completed, all participants took a post-test. The post-test aimed to assess how much their grammar knowledge had improved or changed after using the AI GCs during the writing activities (Table 3).

Table 3. Result of Paired Samples *t*-tests

Group	Test	*M*	*SD*	*df*	*t*	*p*
G1(N=44)	Pre-test	54.66	11.37	43	-4.00	.00**
	Post-test	60.97	11.67			
G2(N=42)	Pre-test	52.80	10.60	41	-7.45	.00**
	Post-test	62.14	9.44			
G3(N=39)	Pre-test	49.47	12.67	38	-5.20	.00**
	Post-test	57.73	14.09			

G1: SpellCheckPlus, G2: Virtual Writing Tutor, G3: Grammarly

The first group using SpellCheckPlus consisted of 44 participants. In the pre-test, their mean score was 54.66 with a standard deviation of 11.37. After the post-test, their mean score increased to 60.97 (SD = 11.67). The difference in scores between in the pre- and post-test was 6.31, and this improvement was found to be statistically significant (*p* < .01). The second group with Virtual Writing Tutor had 42 participants, and their mean score in the pre-test was 52.80, with a standard deviation of 10.60. In the post-test, their mean score improved to 62.14(SD = 9.44). The difference in scores for the second group was 9.34, and this improvement was statistically significant (*p* < .01). The last group using Grammarly included 39 participants, and their pre-test mean score was 49.47 (SD = 12.67), and after the post-test, their mean score increased to 57.73. The increase in scores amounted to 8.26, and this progress was deemed statistically significant with a significance level of *p* < .01.

The result provided evidence that all three groups showed significant improvement in grammar knowledge from the pre-test to the post-test after using three different AI GCs in their English writing activities. Previous studies (Augustin and Wulandari, 2022, Im, 2021, Song and Kim, 2021), it was reported that university students demonstrated a significant improvement in English grammar proficiency through the utilization of AI-based Grammarly. Subsequently, Song and Kim (2021) indicated that the use of AI feedback from Grammarly contributed to enhanced post-assessment grammar knowledge. In addition, Yang (2018) demonstrated that using SpellCheckPlus for L2 writing resulted in a significant reduction in grammar errors. Similarly, findings from the research conducted by Badi et al. (2020) indicated improvement in task attachment, grammar, and punctuation when employing the Virtual Writing Tutor for L2 academic writing. Correspondingly, this present study validates enhanced grammar knowledge when utilizing the three AI GCs programs for L2 writing, similar to these prior research outcomes.

4.1.2 Effects of different AI grammar checkers on grammar learning

The investigation focused on the impact of using different AI GCs on their grammar. The primary objective was to assess whether the choice of AI GC had any meaningful effect on the students' grammar knowledge. By conducting one-way ANOVA (Analysis of Variance), the researchers could compare the means of the post-test scores among the three groups. This allowed them to evaluate if there were statistically significant differences in the grammar knowledge resulting from the use of the different AI GCs.

The outcomes of this analysis are presented in Table 4. Table 4 displays the test scores for grammar tests of three different groups. G1 had an average score of 54.66, G2 scored an average of 52.80, and G3 obtained an average score of 49.47 in the pre-test. The results indicated that there were not any significant differences in the grammar scores among the groups. This suggests that the performance of the groups on both the TOEIC and grammar tests

was similar to each other before experiment.

Table 4. Result of ANOVA Analysis

Stages	Task	Group	*M*	*SD*	*F*	*p*
Pre	Grammar	G1	54.66	11.37	2.12	.12
		G2	52.80	10.60		
		G3	49.47	12.67		
Post	Grammar	G1	60.97	11.67	1.50	.23
		G2	62.14	9.44		
		G3	57.73	14.09		

G1: SpellCheckPlus, G2: Virtual Writing Tutor, G3: Grammarly

Table 4 also provides a summary of the grammar test scores for three different groups in the post-test. As for the G1 group, the students achieved a mean score of 60.97, with a standard deviation of 11.67. The mean score for the G2 was 62.14, while the G3 achieved a mean score of 57.73. However, it is important to note that there was no statistical significance indicated by $p < .01$ among the three groups. This shows that the observed differences in the scores are not considered statistically meaningful.

Drawing from the demonstrated effectiveness of the widely recognized AI GC, Grammarly, this study aimed to investigate the efficacy of grammar learning through two other AI GCs: SpellCheckPlus and Virtual Writing Tutor. Although affirming the overall positive impact of AI GCs of the present study on grammar learning, the result did not uncover significant disparities in their effects on grammar improvement regardless of three different AI GC tools. That is, regardless of different features of three AI GCs, there were no significant outcomes among three groups. In other words, the preceding research emphasizes the efficacy and importance of Grammarly underscoring its utility as a valuable asset in the context of L2 writing (Im, 2021, Long, 2022, Park and Yang, 2020, Song and Kim, 2021). The present study aligns with these prior investigations enhanced grammatical knowledge not only Grammarly but also SpellCheckPlus and Virtual Writing Tutor. Consistent with these findings, this study further solidified the notion that all three AI GCs contribute to enhancing grammar proficiency irrespective of the particular program chosen.

4.2 Students' Perspectives Toward AI Grammar Checkers

4.2.1 Results of pre- and post-surveys

Students' perceptions of English grammar learning with the aid of AI GCs were investigated using both pre- and post-surveys. The survey consisted of sixteen items related to English grammar learning and the utilization of AI GCs for grammar-related tasks. Table 5-7 presents the mean scores for each group and assesses whether there were any noteworthy differences between the pre- and post-stages of the survey.

Table 5 presents survey responses related to students' perceptions of using an AI GC, SpellCheckPlus (SC), for English writing. The mean scores and standard deviations for each item were measured before and after using the program. The mean score for the statement "SC is helpful for English writing" was 4.32 before, and after, it increased to 4.84. The mean score for "I will perform English writing more actively with SC" was 4.19 before, and after, it increased to 4.52. Similarly, for the statement "SC makes me know clearly about my English grammar errors," the mean score was 4.34 before, and after, it remained the same at 4.84. The mean score for "SC helps me check the grammar of English writing" was 4.48 before, and after, it was 5.00. Additionally, for "SC is convenient to use," the mean score was 4.25 before, and significantly increased to 5.09 after. For "SC is easy to use", the mean score was 4.07 before, and significantly increased to 5.30 after. Finally, for the statement "I will use SC in the

future," the mean score was 4.14 before, and significantly increased to 5.02 after. The differences in means were found to be statistically significant.

Table 5. Results of Pre- & Post-Survey: SpellCheckPlus

Item	Pre		Post			
	M	SD	M	SD	t	p
1. SC is helpful for English writing.	4.32	.98	4.84	.78	43	.00**
2. I will perform English writing more actively with SC.	4.19	.99	4.52	.90	43	.03*
3. I am interested in English writing with SC.	3.82	1.02	4.20	1.05	43	.06
4. I am motivated for English writing with SC.	4.05	1.01	4.14	.95	43	.66
5. SC makes me know clearly about my English grammar errors.	4.34	.86	4.84	.83	43	.01*
6. SC helps me check the grammar of English writing.	4.48	1.05	5.00	.78	43	.00**
7. SC is helpful for detailed feedback.	4.14	1.21	4.52	1.05	43	.09
8. SC is convenient to use.	4.25	.87	5.09	.77	43	.00**
9. SC is easy to use.	4.07	.90	5.30	.70	43	.00**
10. I will use SC in the future.	4.14	.85	5.02	.82	43	.00**

SC: SpellCheckPlus   *$p < .05$, **$p < .01$

Overall, students' perceptions of using the AI GC, SpellCheckPlus, were generally positive, with significant improvements in several areas, such as its helpfulness, performing English writing actively, clear identification of grammar errors, checking grammar, ease of use, and future use. This result aligns with the findings from the positive perception shown by learners after using SpellCheckPlus in the previous study (Yang, 2018). Nonetheless, there were no significant changes observed in aspects such as students' motivation, interest, and specific feedback pertaining to the utilization of SpellCheckPlus.

Table 6 presents survey responses related to students' perceptions of an AI GC, Virtual Writing Tutor, for English writing. The mean scores and standard deviations for each item were measured before and after using the program. Among the responses of using Virtual Writing Tutor (VW), certain items exhibited significant differences, while others did not show any notable variations. The mean score for the statement "VW is helpful for English writing" increased from 4.43 before to 4.76 after. The mean score for "VW makes me know clearly about my English grammar errors" increased from 4.48 before to 4.88 after, and for "VW helps me check the grammar of English writing," it increased from 4.64 before to 5.05 after. Additionally, for "VW is convenient to use," the mean score significantly increased from 4.69 before to 5.17 after. Likewise, for "VW is easy to use," the mean score significantly increased from 4.69 before to 5.36 after. Finally, for the statement "I will use VW in the future," the mean score rose from 4.36 before to 4.88 after. The differences in scores were statistically significant with $p < .01$.

In summary, students generally perceived the Virtual Writing Tutor positively, and significant improvements were observed in its helpfulness, future usage intent, clarity in identifying grammar errors, checking grammar, convenience, and ease of use. The results of this survey are consistent with the findings of Badi et al. (2020), which demonstrated that learners showed a positive perception after using Virtual Writing Tutor in writing classes. However, some aspects, such as active engagement, interest in English writing, and detailed feedback, did not show statistically significant differences before and after using the program.

Table 6. Results of Pre- & Post-Survey: Virtual Writing Tutor

Item	Pre		Post			
	M	SD	M	SD	t	p
1. VW is helpful for English writing.	4.43	.83	4.76	.91	41	.03*
2. I will perform English writing more actively with VW.	4.33	.79	4.45	1.09	41	.51
3. I am more interested in English writing with VW.	4.17	.96	3.79	1.14	41	.07
4. I am motivated for English writing with VW.	4.14	.95	3.93	1.11	41	.32
5. VW makes me know clearly about my English grammar errors.	4.48	.99	4.88	.86	41	.04*
6. VW helps me check the grammar of English writing.	4.64	.88	5.05	.85	41	.03*
7. VW is helpful for detailed feedback.	4.48	.89	4.67	1.00	41	.32
8. VW is convenient to use.	4.69	.75	5.17	.96	41	.00**
9. VW is easy to use.	4.69	.84	5.36	.96	41	.00**
10. I will use VW in the future.	4.36	.85	4.88	.97	41	.00**

VW: Virtual Writing Tutor   *$p < .05$, **$p < .01$

Table 7 shows the results related to students' perceptions of an AI GC, Grammarly (GR), for English writing. The mean scores and *p*-values for each item were measured before and after using the program. The mean score for the statement "GR is helpful for English writing" significantly increased from 4.36 before to 4.97 after. The mean score for "I will perform English writing more actively with GR" increased from 4.28 before to 4.72 after, and for "GR makes me know clearly about my English grammar errors", it increased from 4.44 before to 4.82 after. Additionally, for "GR helps me check the grammar of English writing", the mean score increased from 4.51 before to 4.95 after. Likewise, for "GR is convenient to use", the mean score significantly increased from 4.15 before to 4.97 after. For the item, "GR is easy to use", the mean score significantly increased from 4.46 before to 5.10 after. Finally, for the statement "I will use GR in the future," the mean score significantly increased from 4.15 before to 4.95 after. The differences in means were statistically significant with $p < .05$.

Table 7. Results of Pre- & Post-Survey: Grammarly

Item	Pre		Post			
	M	SD	M	SD	t	p
1. GR helpful for English writing.	4.36	.71	4.97	.87	38	.00**
2. I will perform English writing more actively with GR.	4.28	.69	4.72	1.10	38	.02*
3. I am more interested in English writing with GR.	3.97	.78	4.21	1.13	38	.25
4. I am motivated for English writing with GR.	4.03	.90	4.28	.83	38	.19
5. GR makes me know clearly about my English grammar errors.	4.44	.75	4.82	.97	38	.03*
6. GR helps me check the grammar of English writing.	4.51	.68	4.95	.83	38	.01*
7. GR is helpful for detailed feedback.	4.23	1.01	4.49	1.07	38	.31
8. GR is convenient to use.	4.15	.79	4.97	1.04	38	.02*
9 GR is easy to use.	4.46	.75	5.10	1.07	38	.00**
10. I will use GR in the future.	4.15	.81	4.95	1.30	38	.00**

GR: Grammarly   *$p < .05$, **$p < .01$

Overall, students' perceptions of the AI GC, Grammarly, were generally positive, with significant improvements observed in its perceived helpfulness, active engagement in English writing, clarity in identifying grammar errors, convenience, ease of use and future use. However, there were no significant differences in interests, motivation, and detailed feedback.

Consequently, the AI GCs yielded positive shifts in learners' perceptions, encompassing assistance in English composition, active engagement, facilitation of grammar learning, and error detection. Moreover, its user-friendly and easily accessible nature demonstrated a willingness among learners to continue utilizing it in the future. The results align with earlier studies (Long, 2022, Park and Yang, 2020, Song and Kim, 2021), which also emphasized the positive results linked to immediate feedback, highlighting its beneficial aspects for university students in relation to usefulness, contentment, and effectiveness. The outcomes of the present investigation substantiate prior research indicating that feedback conveyed through 'Grammarly' was intelligible, facilitating self-directed revisions in English writing and thereby fostering an increased sense of writing confidence (O'Neill and Russell, 2019).

Interestingly, however, a noteworthy consensus emerged that there was no statistically significant change in the perception of learners regarding motivation and interest in English writing, as well as the receipt of detailed feedback, subsequent to the utilization of an AI GC. It can be assumed that using AI checkers do not impact students' interest or motivation towards writing. This implies that the incorporation of an AI GC into English composition did not appear to enhance learners' enthusiasm and motivation in the realm of English writing. Moreover, students might not receive detailed feedback from the checkers, which underscores the necessity of feedback from teachers.

4.2.2 Students opinions towards AI Grammar Checkers

To investigate students' perceptions in greater depth, open-ended survey questions were utilized. Following the study, all the responses were collected and categorized based on their similarities within each group. A comparison was then made among the responses from the different groups.

Table 8. Advantages of SpellCheckPlus

	Opinions	#	%
1	**Identifying errors** - Writing errors can be found. - It is convenient to find an error that I could not find. - I think the biggest advantage was that I could easily find errors that were hard to find.	25	57%
2	**Suggesting writing feedback (error correction)** - I also learned better expressions to find and use grammar errors in a short time using AI GC. - It has the advantage of facilitating easy understanding of grammar by identifying confusing grammar, explaining why it is incorrect, and providing feedback.	7	16%
3	**Convenience of use** - The AI GC is user-friendly, making it simple and easy to use. Additionally, it performs detailed checks for accuracy.	6	14%
4	**Helpful in learning** -While writing in English, I received feedback regarding errors; however, at times, they did not provide suggestions for correction, prompting me to think independently. -Utilizing the AI GC allowed me to gain a better understanding of the format of academic writing.	3	7%
5	**No response**	3	7%
Total		44	100%

The analysis of feedback pertaining to the AI GC, SpellCheckPlus, underscores noteworthy observations (Table 8). Notably, 57% of users perceived a distinct advantage in error identification, particularly for errors that are challenging to discern through self-checking. Moreover, 16% acknowledged the program's capacity to provide constructive writing feedback and propose enhanced expressions to foster a deeper understanding of grammar. Additionally, the user-friendliness of the AI GC program resonated with 14% of participants, while 7% acknowledged its pedagogical value, despite occasional limitations in correction recommendations.

The feedback from users regarding AI GC, Virtual Writing Tutor, indicates several positive aspects, presented in Table 9. Firstly, 50% of participants found AI GC remarkably efficient in detecting grammar errors with ease. Secondly, 21% of users appreciated the convenience of using the program. Thirdly, 19% of respondents valued the writing feedback provided by the program, citing its clear explanations for their mistakes and the guidance it offered for corrections. Moreover, 5% of users expressed that the program was helpful in their learning process, enabling them to identify recurring grammar mistakes and gain a deeper understanding of English grammar. Finally, a small percentage (5%) did not provide a response to the survey. That is, the feedback from the students using Virtual Writing Tutor highlights the program's effectiveness in detecting errors, its user-friendly interface, and its valuable suggest to write better sentences.

Table 9. Advantages of Virtual Writing Tutor

	Opinions	#	%
1	**Identifying errors**	21	50%
	- It is effortless to detect grammar errors.		
	- It is beneficial to have the opportunity to review and identify the grammatical mistakes that I had overlooked previously.		
2	**Convenience of use**	9	21%
	- It is comfortable to use it.		
3	**Suggesting writing feedback (error correction)**	8	19%
	- The feedback was easy to comprehend as it provided clear explanations for my mistakes.		
	- It was beneficial to see the areas where I made errors and receive guidance on how to correct them through the provided feedback.		
4	**Helpful in learning**	2	5%
	- It was valuable to identify the specific grammar mistakes I repeatedly made, allowing me to be more mindful and avoid them in the future.		
	- I believe this experience provided me with a deeper understanding of English grammar, which was previously vague to me.		
5	**No response.**	2	5%
	Total	42	100%

The feedback of Grammarly reveals in Table 10. 46% of users acknowledged the AI GC as a valuable resource for identifying grammatical errors that would have otherwise evaded detection through self-assessment. A notable 28% of respondents expressed appreciation for the program's convenience and user-friendly interface. Furthermore, 15% of participants attested to its efficacy in enhancing their writing proficiency, particularly in the identification of intricate grammatical errors. Additionally, 8% of respondents emphasized the program's merits in providing expeditious and accurate feedback, accompanied by lucid explanations. Overall, the feedback highlights the error detection, convenience, and support in English writing.

Table 10. Advantages of Grammarly

	Opinions	#	%
1	**Identifying errors**	18	46%
	- It was highly beneficial to discover grammar errors that I couldn't identify through self-checking.		
	- It was advantageous to catch grammar mistakes that I had overlooked.		
2	**Convenience of use**	11	28%
	- It is highly convenient and accessible, providing a user-friendly experience for checking grammatical errors.		
	- Its good accessibility and simple usability are commendable features.		
3	**Helpful in learning**	6	15%
	- AI assisted me in enhancing my English writing ability by pinpointing complex English grammar errors that are challenging to detect independently, without the aid of someone else.		
	- It enabled me to rectify the grammar mistakes I had made, thereby allowing me to construct improved sentences.		
4	**Suggesting writing feedback (error correction)**	3	8%
	- I received prompt and accurate feedback on my essays.		
	- I appreciated the fact that it not only pointed out my errors but also provided clear explanations and guidance on how to correct them properly.		
5	**No response**	1	3%
	Total	39	100%

As a result, the participants using SpellCheckPlus highlighted its main advantages as identifying errors, suggesting writing feedback, and convenience of use. Similarly, the Virtual Writing Tutor group pointed out identifying errors, convenience of use, and suggesting writing feedback as its primary positive aspects. The last group in Grammarly was highlighted for its primary strengths in identifying errors, convenience of use, and its contribution to learning. Intriguingly, the Grammarly was deemed beneficial for learning, with a specific comment stating, "enhancing my English writing ability by pinpointing complex English grammar errors that are challenging to detect independently", "rectifying the grammar mistakes I had made, thereby allowing me to construct improved sentences." In other words, all three AI GCs excel in promptly identifying and rectifying errors while providing user-friendliness, whereas Grammarly stands out for affording learners opportunities for learning.

The analysis of SpellCheckPlus's feedback has yielded the following insights (see Table 11). Firstly, a substantial number of responses (27%) mentioned that the use of this program has no disadvantages. 27% of participants, encountered confusion stemming from the system's tendency to classify accurate text as erroneous and erroneously flag proper nouns. They identified instances of inaccuracies. In addition, 9% of users articulated their dissatisfaction with the limitation on the number of words to perform a check on all sentences simultaneously. Furthermore, an equivalent 9% expressed discontentment with the dearth of comprehensive feedback. A discernible 7% of participants acknowledged a heightened reliance on external grammar assistance as a result of their interaction with the tool. Lastly, 5% of respondents delineated constraints in the tool's capacity to comprehend contextual nuances.

Table 11. Disadvantages of SpellCheckPlus

	Opinions	#	%
1	**None**	12	27%
2	**Low accuracy**	12	27%
	- Proper nouns were also flagged as errors, which raised concerns about instances where it marked something as an error when it was not.		
	- Sometimes, it lacks accuracy.		
3	**Inconvenience of use**	4	9%
	- The limitation on the number of words makes it inconvenient for the system to perform a comprehensive check on all sentences simultaneously.		
4	**Insufficient feedback (explanation)**	4	9%
	- At times, its explanations are challenging to comprehend.		
5	**Increased tool dependency.**	3	7%
	- I felt somewhat reliant on external assistance.		
6	**Limitations in understanding contextual meaning**	2	5%
	- There are limitations to understanding contextual meaning.		
7	**No response**	7	16%
Total		44	100%

Table 12 displays the viewpoints concerning the drawbacks associated with utilizing the Virtual Writing Tutor. Approximately 41% of users raised apprehensions regarding the system's diminished accuracy in error detection and its propensity to misconstrue intended meanings. Subsequently, 15% of respondents perceived the feedback explanations as insufficient. A smaller portion, accounting for 7%, indicated a potential for augmented reliance on the system, while 2% identified a deficiency in contextual comprehension.

Table 12. Disadvantages of Virtual Writing Tutor

	Opinions	#	%
1	**None**	11	26%
2	**Low accuracy**	17	41%
	- In certain instances, the accuracy was relatively poor, making it difficult to distinguish between non-errors and actual errors, which was confusing.		
3	**Insufficient feedback (explanation)**	6	15%
	- It was unfortunate that there was no detailed explanation provided, specifying which part was incorrect and the reason behind the error.		
4	**Increased tool dependency.**	3	7%
	- I feel that as I continue to use it for English writing, the tendency to rely on it will likely increase.		
5	**Lack of contextual meaning**	1	2%
	- There have been instances where the system did not provide suitable solutions based on the context.		
6	**Inconvenience of use**	1	2%
	- It is troublesome to repeatedly access the site.		
7	**No response**	3	7%
Total		42	100%

The examination of feedback data related to Grammarly exposes several notable concerns (Table 13). A total of 36% of the students responded that there were no disadvantages. Subsequently, 15% of participants expressed apprehension concerning the potential increase of reliance on the AI GC. An additional 13% encountered intermittent instances of errors and misidentifications. Among the respondents, 8% found the explanations

accompanying the feedback to be inadequately informative. A minor 5% highlighted concerns associated with advertising in the free version. Similarly, 5% pointed out limitations in the program's contextual understanding. In parallel, an equivalent 5% reported inconvenience stemming from the program's automatic activation.

Table 13 Disadvantages of Grammarly

	Opinions	#	%
1	**None**	14	36%
2	**Increased dependency** - Instead of self-checking, I found myself relying more on this program. - Depending too much on it might create a mindset where AI GC will correct your writing.	6	15%
3	**Low Accuracy** - There were occasional errors, which was inadequate. - It occasionally misidentifies the correct parts as wrong.	5	13%
4	**Insufficient feedback (explanation)** - The free version does not offer feedback on sentence errors.	3	8%
5	**Advertising (premium) problems** - The free version lacks the ability to change synonyms, and it only corrects errors without providing grammar information.	2	5%
6	**Limitations in understanding contextual meaning** - Sometimes, it suggests grammar changes that alter the intended meaning of my writing.	2	5%
7	**Inconvenience of use** - The program automatically activates even when I don't intend to use it.	2	5%
8	**No response**	5	13%
	Total	39	100%

As a result, after using SpellCheckPlus, learners mentioned drawbacks including low accuracy, inconvenience of use, inadequate feedback (explanation), heightened dependency on the tool, and limitations in comprehending contextual meaning. Similarly, for those who utilized Virtual Writing Tutor, identified disadvantages encompassed low accuracy, insufficient feedback (explanation), increased reliance on the tool, absence of contextual meaning, and inconvenience of use. However, learners using Grammarly presented differing opinions compared to the other two groups. Among the group utilizing Grammarly, the highlighted drawbacks comprised increased dependency, low accuracy, inadequate feedback (explanation), issues related to advertising (premium), limitations in understanding contextual meaning, and inconvenience of use. This discrepancy arises from the utilization of Grammarly, wherein learners expressed diverse viewpoints. The variations between the free and premium versions of the tool, along with the inconvenience posed by its advertisements, contributed to these differences. While some learners acknowledged its beneficial impact on learning, these advantages were juxtaposed with concerns about the possibility of excessive reliance on the tool if used extensively.

Considering the limitations observed following the utilization of AI GCs, it is clear that there is a need for improvements in areas such as enhancing accuracy, expanding feedback capabilities, deepening contextual understanding, and optimizing user-friendliness. In conclusion, these results exhibit parallels with a previous investigation conducted by Park (2019), which indicates that Grammarly's feedback aligns with accuracy in relation to specific errors. In essence, while AI GCs swiftly and conveniently detect errors and assist in their L2 writing, considering aspects such as the context and the tendency to flag even accurate segments as errors, it would be advisable to utilize them in conjunction with teacher or peer feedback.

5. Conclusion

This study aimed to explore the efficacy of three different AI GCs in the context of English writing, aiming to understand their grammar improvement and learners' perceptions. Previous research predominantly compared and analyzed the effects of professor feedback or self-feedback against those of AI GCs, examining the areas where writing improved with AI GCs' feedback. Some studies also have investigated students' perspectives toward AI GCs. However, given the prevalence of AI GC use, particularly Grammarly, this study sought to investigate the effects and perceptions of grammar improvement after utilizing Grammarly as well as other AI GCs (SpellCheckPlus and Virtual Writing Tutor) during the writing process.

For the first research question, "To what extent does self-correction using three different AI GCs (SpellCheckPlus, Virtual Writing Tutor, Grammarly) help students improve their grammar proficiency?", compared to the pre-grammar test, all three groups of participating students exhibited grammar improvement in the post-grammar test after the experiment. However, there were no significant differences in post-test grammar scores among the three AI GC groups after the experiment. In other words, while grammar proficiency improved after using AI GCs during the English composition process, no discernible differences in grammar improvement were observed across the different AI GCs used. These align with previous research, indicating that AI GCs positively impact both writing and grammar. However, no significant variance was observed in grammar improvement based on the different AI GC used during the writing process. Prior research underscores the effectiveness and significance of Grammarly, highlighting its valuable contribution to L2 writing (Im, 2021, Long, 2022, Song and Kim, 2021). This current study aligns with these previous investigations by showcasing improved grammar proficiency across three distinct groups utilizing different AI GCs (Yang, 2018, Badi et al. 2020). By these findings, this study further reinforces the notion that all three AI GCs contribute to boosting grammatical proficiency, regardless of the specific AI GC.

The second research question aimed to investigate learners' perceptions regarding the usage of three distinct AI GCs. The analysis of responses from the open-ended questionnaire yields the following results. All the groups exhibited favorable attitudes toward AI GCs, namely SpellCheckPlus, Virtual Writing Tutor, and Grammarly as evidenced by considerable enhancements across dimensions like perceived effectiveness, more active involvement in English writing, precision in grammatical error identification, convenience, user-friendliness, and future use to employ the tool. However, specific aspects such as active engagement and the receipt of comprehensive feedback demonstrated no statistically significant alterations prior to and after employing the AI GCs. Consequently, this study also substantiates the effectiveness of all three AI GCs in facilitating error detection and correction across their respective functionalities (John and Woll, 2020).

Similar opinions emerged for all three AI GCs regarding the advantages and disadvantages after using AI GCs. As for the advantages, the participants utilizing SpellCheckPlus emphasized its primary benefits in terms of error identification, suggesting writing feedback, and ease of use. The Virtual Writing Tutor group highlighted identifying errors, convenience of use, and providing writing feedback as its core positive attributes. The third group using Grammarly, was recognized for excelling in identifying errors, convenience, and enhancing learning. Interestingly, Grammarly proved valuable for learning, with a specific comment mentioning its effectiveness in addressing intricate grammar errors that are difficult to spot independently, and enabling the correction of mistakes for improved sentence construction. In essence, all three AI GCs excel in swiftly detecting and correcting errors while offering user-friendly interfaces. However, Grammarly stands out by affording learners opportunities for educational advancement.

Students presented a variety of opinions regarding various drawbacks as well. After using SpellCheckPlus,

learners mentioned drawbacks like accuracy issues, inconvenience, inadequate feedback, reliance on the tool, and challenges in understanding context. Similarly, Virtual Writing Tutor users noted downsides including inaccuracies, insufficient feedback, increased tool dependency, limited contextual understanding, and user inconvenience. However, Grammarly users had differing opinions. They pointed out increased dependency, reduced accuracy, insufficient feedback, premium-related advertising concerns, limited contextual understanding, and user inconvenience. This diversity in views results from Grammarly's usage, reflecting various learner perspectives influenced by differences between its free and premium versions and the inconvenience of advertising. While some recognized the AI GC's benefits, concerns arose about excessive reliance through extensive use.

Given that the Writing Virtual Tutor encompasses features such as grammar checking, punctuation correction, spell checking, paraphrasing, vocabulary enhancement, as well as the ability to structure essays and compose hypertext narratives, this tool holds substantial potential for aiding in the writing process. However, Grammarly, widely used by many learners, presents inconveniences such as persistent advertisements and prompts for upgrading to its paid version. Therefore, it can be advantageous for learners to explore alternative tools with similar functionalities to address these drawbacks. In addition, due to the distinct advantages inherent in various programs, utilizing not only a single tool but also multiple tools concurrently based on the user's objectives could potentially serve as a beneficial learning approach (Jeon et. Al., 2021). Moreover, while AI GCs aid students' writing and grammar learning, their use should be combined with teacher or peer feedback (Min, 2020, O'Neill and Russell, 2019) to account for context and the risk of flagging accurate segments as errors.

The findings of this study have yielded compelling outcomes, yet several limitations warrant consideration. Firstly, a more assessment could be achieved by contrasting the performance of learners who engaged with AI GCs with those who did not, thereby facilitating a more comprehensive evaluation. Moreover, it is essential to acknowledge the restricted generalizability of the study's outcomes, attributed to the exclusive inclusion of students from a college in Korea, sharing comparable levels of English proficiency. Thus, to establish broader applicability, the inclusion of a more diverse student cohort representing varying proficiency levels and diverse backgrounds is imperative. Furthermore, while the research examined learners' perspectives and perceptions concerning AI GCs, a more comprehensive investigation is warranted through qualitative interviews, enabling an in-depth exploration of learners' nuanced insights and experiences.

Building upon the present study's outcomes, future research endeavors could contemplate a nuanced exploration of AI GC implementation, employing a stepwise feedback approach to assess its impact on learners' grammar and vocabulary knowledge. Exploring potential differences in the utilization of AI GCs corresponding to varying levels of learner proficiency could offer significant insights for pedagogical purposes. Furthermore, a promising avenue for investigation involves conducting analyses before and after learners' engagement with various AI GCs, which could illuminate the relative effectiveness of these tools in enhancing writing skills.

References

Al-Ahdal, A. 2020. Using computer software as a tool of error analysis: Giving EFL teachers and learners a much-needed impetus. *International Journal of Innovation, Creativity and Change* 12(2), 418-437.

Agustin, R. and S. Wulandari. 2022. The analysis of grammatical errors on students' essay writing by using Grammarly. *Jurnal Pendidikan Bahasa Inggris Proficiency* 4(1), 39-46.

Badi, A., M. Osman and A. Al-Mekhlafi. 2020. The impact of Virtual Writing Tutor on writing skills and attitudes of Omani college students. *The Journal of Education and Development* 4(3), 101-116.

Chun, H., S. Lee and I. Park. 2021. A systematic review of AI technology use in English education. *Multimedia-Assisted Language Learning* 24(1), 87-103.

Chung, S. 2018. Use of rubric-referenced self-assessment in EFL instruction: A tool for self-regulated learning. *Journal of Learner-Centered Curriculum and Instruction* 18(10), 767-789.

Dale, R. 2016. Checking in on grammar checking. *Natural Language Engineering* 22(3), 491-495.

Ferris, D. 2003. *Response to Writing: Implications for Second Language Students.* NJ: Lawrence Erlbaum.

Ferris, D. and B. Roberts. 2001. Error feedback in the L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing* 10, 161-184.

Ghufron, M. A. and F. Rosyida. 2018. The role of Grammarly assessing English as a foreign language (EFL) writing. *Lingua Cultura* 12(4), 395-403.

Hadiat, A.W.F., W. Tarwana and L. Irianti. 2022. The use of Grammarly to enhance students' accuracy in writing Descriptive text (A case study at eighth grade of a Junior High School in Ciamis). *Journal of English Education Program (JEEP)* 9(2), 1-10.

Hedgecock, J. 2005. Taking stock of research and pedagogy in L2 writing. In E. Hinkel, ed., *Handbook of Research in Second Language Teaching and Learning*, 597-613. Mahwah, NJ: Lawrence Erlbaum.

Hyland, F. 2003. Focusing on form: Student engagement with teacher feedback? *System* 31(2), 217-230.

Hyland, K. and F. Hyland. 2006. *Feedback in Second Language Writing.* Cambridge: Cambridge University Press.

Im, H. 2021. The use of an online grammar checker in English writing learning. *Journal of Digital Convergence* 19(1), 51-58.

John, P. and N. Woll. 2020. Using grammar checkers in an ESL context: An investigation of automatic corrective feedback. *CALICO Journal* 37(2), 169-192.

Joo, H. 2020. An investigation of the effect of self-correction on error and syntactic complexity of Korean adults' English writing. *English Language Assessment,* 15(2), 215-233.

Junqueira, L. and C. Payant. 2015. I just want to do it right, but it's so hard: A novice teacher's written feedback beliefs and practices. *Journal of Second Language Writing* 27, 19-36.

Kang, D. 2008. Feedback on EFL writing: Teacher, peer, and self-review. *Foreign Languages Education* 15(1), 1-22.

Kim, Y. 2013. Exploring English writing feedback employed Korean collegiate students. *The Mirae Journal of English Language and Literature* 18(1), 177-205.

Kim, Y. and N. Kim. 2013. The types and characteristics of teacher feedback and peer feedback in process-based writing classes. *Studies in Linguistics* 29, 1-26.

Lee, I., P. Mak and R. Yuan. 2019. Assessment as learning in primary writing classrooms: An exploratory study. *Studies in Educational Evaluation* 62, 72-81.

Lindblom-ylanne, S., H. Pihlajamaki and T. Kotkas. 2006. Self-, peer- and teacher-assessment of student essay. *The Journal of the Institute for Learning and Teaching* 7(1), 51-62.

Link, S., A. Dursun, K. Karakaya and V. Hegelheimer. 2014. Towards best ESL practices for implementing automated writing evaluation. *CALICO Journal* 31(3), 323-344.

Long, R. 2022. Online grammar checkers versus self-editing: An investigation of error correction rates and writing quality. *Journal of Nusantara Studies* 7(1), 441-458.

Mansoor, G. 2009. Teacher-correction, peer-correction and self-correction: Their impacts on Iranian students' IELTS essay writing performance. *The Journal of Asia TEFL* 6(1), 117-139.

Min, J. 2020. Korean university students' perceptions of feedback activities on English composition: Self-review, peer feedback, and computer feedback. *Korea Journal of English Language and Linguistics* 20(1), 335-362.

Moon, D. 2021. Evaluating corrective feedback generated by an AI-powered online grammar checker. *International Journal of Internet, Broadcasting and Communication* 13(4), 22-29.

Nazari, N., M. Shabbir and R. Setiawan. 2021. Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial. *Heliyon* 7(5), 1-9.

O'Neill, R. and A. M. T. Russell. 2019. Stop! Grammar time: University students' perceptions of the automated feedback program Grammarly. *Australian Journal of Educational Technology* 35(1), 42-56.

Park, C. and H. Kim. 2016. The effect of self and peer feedback: Learner autonomy in high school writing classes. *Modern English Education* 17(3), 53-77.

Park, J. 2019. An AI-based English grammar checker vs. human raters in evaluation EFL learners' writing. *Multimedia-Assisted Language Learning* 22(1), 112-131.

Park, J. and I. Yang. 2020. Utilizing an AI-based grammar checker in an EFL writing classroom. *Korean Journal of Applied Linguistics* 36(1), 97-120.

Ranalli, J. 2018. Automated written corrective feedback: How well can students make use of it? *Computer Assisted Language Learning* 31(7), 653-674.

Sahu, S., Y. Vishwakarma, J. Kori and J. Thakur. 2020. Evaluating performance of different grammar checking tools. *International Journal of Advanced Trends in Computer Science and Engineering* 9(2), 2227-2233.

Song, E. and H. Kim. 2021. The effects of AI grammar checker and self-feedback on English grammar of EFL college students. *Journal of Foreign Studies* 57, 91-120.

Tompkins, G. E. 2008. *Teaching Writing: Balancing Process and Product* (5th ed.). Upper Saddle River, NJ : Pearson Merrill Prentice-Hall.

Van Beuningen, C. G., N. H. De Jong and F. Kuiken. 2012. Evidence on the effectiveness of comprehensive error correction in Dutch multilingual classrooms. *Language Learning* 62, 1–41.

Wang, X. 2023. Chinese English learner's perspectives on automated writing feedback. *Multimedia-Assisted Language Learning* 26(2), 33-58.

Yang, H. 2010. SpellCheckPlus. *CALICO Journal* 28(1), 261-267.

Yang, H. 2018. Efficiency of online grammar checker in English writing performance and students' perceptions. *Korean Journal of English Language and Linguistics* 18(3), 328-348.

Zhang, Z. and K. Hyland. 2018. Students engagement with teacher and automated feedback on L2 writing. *Assessing Writing* 36, 90-10.

Zhang, J., H. Ozer and R. Bayazeed. 2020. Grammarly vs. face-to-face tutoring at the writing center: ESL student writer's perceptions. *Praxis: A Writing Center Journal* 17(2), 33-48.

Examples in: English
Applicable Languages: English
Applicable Level: Tertiary