

The Effects of Grammarly-Assisted Written Feedback on ESL Undergraduate Students' Academic Writing Accuracy

Vaasala Kaliappa^{1,*}, Julie Indra Krishnan²

¹ General Studies Unit, Arau Community College, Arau, Perlis, Malaysia
Jalan Tengku Budriah, Tambun Tulang, 02600 Arau, Perlis

² Faculty of Education, Open University Malaysia, Pulau Pinang, Malaysia
No. 1, Lebuhraya 2, Pusat Bandar Seberang Jaya, 13700 Pulau Pinang, Malaysia.
E-mail: ^{1,*}vaasala@staf.kkarau.edu.my, ²indrajulie3684@gmail.com

Correspondence Author Email: vaasala@staf.kkarau.edu.my

Submitted: 16/05/2026; Accepted: 12/06/2026; Published: 30/06/2026

Abstract—The application of artificial intelligence (AI) powered tools in the context of second language writing classes has grown rapidly in recent years. One such tool that has changed the way written corrective feedback given in English language classrooms is Grammarly. It is an automated writing evaluation (AWE) system that is gaining widespread popularity as a tool that supports language accuracy. However, limited empirical evidence is available on its effectiveness in improving writing accuracy of English as Second Language (ESL) undergraduates in Malaysian context. Therefore, further controlled studies particularly classroom-based intervention studies are required in order to establish whether written feedback from Grammarly can enhance the accuracy of academic writing among ESL undergraduate students in Malaysia. This study focused on the effects of feedback from Grammarly on grammatical accuracy in argumentative academic essays. Quasi-experimental design with pre-test and post-test control group was employed for this study. The participants used in the study were sixty-four ESL undergraduate students. The experimental group was provided with written feedback from Grammarly during eight weeks, and the control group received feedback from a teacher. Grammatical error ratio (Errors per 100 words), percentage of error-free T-units, and syntactic accuracy were used as measurements of writing accuracy. Findings indicated that despite the improvement in both groups, the experimental group assisted by Grammarly had more reductions in grammatical errors, and higher increases in error-free units relative to the control condition. These results are consistent with the recent studies that indicate that corrective feedback mediated by Grammarly and other AWE systems can be effective in the development of writing, provided it is applied in a well-structured writing revision. Nevertheless, the intervention lasted only for eight weeks and long-term retention of writing accuracy was not investigated after the feedback tool was withdrawn. Therefore, the generalizability of the findings is limited. Future studies should examine the long-term effects of employing Grammarly as a supplementary instructional resource in ESL undergraduate writing classes that improve writing accuracy.

Keywords: Grammarly; Automated Writing Evaluation (AWE); ESL Academic Writing; Grammatical Accuracy; Written Corrective Feedback; L2 Writing Development

1. INTRODUCTION

The role of digital tools in writing instruction has been growing at a very fast pace in recent years, and nowadays it is heavily influenced by artificial intelligence (AI). A thematic review of technology in writing instruction indicates that technology is no longer a means of typing or formatting instead technology is becoming more useful in the learning process, including drafting, revising, and language accuracy (Eragamreddy & Joseph, 2025). In this more general online transformation, automated writing evaluation (AWE) software such as Grammarly has received particular focus in English language instruction since the devices are capable of offering instant feedback and facilitating rewrites. The systematic review of AWE tools in argumentative writing also indicates that the tools are gaining popularity in EFL settings, but their results are affected by how learners perceive and utilize feedback (Haddadian et al., 2025). Meanwhile, the idea of applying AWE and AI-related tools to the writing practice in higher education is increasingly gaining momentum, as well as the more structured integration of these tools into classroom activities and their daily use in student writing activities (Alshehri, 2025). It is a combination of these trends that AI-powered feedback is becoming a viable aspect of ESL/EFL writing classrooms, particularly on a university level.

Although more writing materials are now available, accuracy continues to be a significant problem for ESL students, particularly in academic writing. The most striking fact is that many undergraduate students still commit a lot of grammatical mistakes in a variety of drafts, and certain ones cannot be eliminated with time, as learners can still commit them despite feedback (Haddadian et al., 2025). In academic work, recurring mistakes may decrease the clarity and performance in assessment-based conditions. It is due to this fact that written corrective feedback has continued to be a significant subject in second language writing studies. Nevertheless, teacher-based feedback may be restricted by time, teaching classes, and labor. In case of delayed or unstable feedback, learners cannot revise effectively, or they tend to commit the same errors in their next assignments. This problem leads to the need to find feedback methods that are more frequent and easier to offer, yet do not rely on superficial correction but promote genuine learning.

Recent research efforts on AI in L2 writing have not just widened their scope past accuracy but tend to concentrate on aspects such as engagement and motivation among learners, as well as affective outcomes. As an example, the study of how AI applications such as ChatGPT can be used to facilitate grit and persistence in writing courses has been undertaken (Bekdas, 2026). In other study literature, it is also found that AI-adaptive feedback can affect engagement and writing strategies, and behavioral reactions to feedback can lead to results in learners (Luo & Yusuf, 2025). These studies contributed valuable findings into the incorporation of AI in ESL writing, however, there are fewer studies that

specifically and directly concentrate on academic writing accuracy among ESL undergraduate learners, particularly with transparent metrics of accuracy and controlled intervention design. In addition, variations in writing tasks, feedback mechanisms and instructional contexts contributed to inconsistent findings regarding the effectiveness of AI-mediated writing correcting feedback (Rahim & Gregory, 2025). More importantly, impact of corrective feedback generated by Grammarly as the specific AWE tool to improve writing accuracy of Malaysian ESL undergraduate students in authentic classroom settings is limitedly examined through empirical studies. Therefore, this study addresses the gap by examining the impact of written feedback from Grammarly on the academic writing accuracy of ESL undergraduates through a quasi-experimental classroom design.

This study aims at investigating the effects of Grammarly-supported written feedback on the accuracy of ESL undergraduate students in academic writing based on measurable accuracy indicators in drafts compared with the results of a conventional teacher feedback condition.

The study was guided by the following objectives, to investigate if corrective feedback from Grammarly significantly reduces grammatical error rates over time, to measure the impact of AWE on structural and syntactic writing accuracy compared to traditional teacher feedback.

Written corrective feedback (WCF) is very common in writing in a second language, as it assists learners to observe errors and revision in their writing. Learners are provided opportunities to reflect on their errors and revise their texts accordingly. According to Ngyuen and Chu (2024) WCF has positive impact towards writing development where it makes learners aware of language forms and gradually improve the grammatical accuracy in ESL and EFL writing contexts. Furthermore, WCF serves both diagnostic and pedagogical function which enables learners to recognise recurring errors and eventually learners use more accurate language in their subsequent writing tasks.

One of the approaches within WCF research is Dynamic Written Corrective Feedback (DWCF) that focuses on the repetition of feedback and rewriting. The study of DWCF emphasizes that continuous and focused feedback provides the best chance of enhancing it, particularly in the case of errors that are likely to stabilize over time. DWCF encourages learners to engage repeatedly with feedback so that persistent grammatical errors can be identified and gradually reduced rather than treating correction as a one-time occurrence. A recent study by (Cheng & Xu, 2025) highlighted that immediate and repetitive corrective feedback encourages greater learner engagement and contributes to improvements in writing accuracy over time. Learners must actively process feedback and apply revision to subsequent writing task to make corrective feedback most effective. Error correction is then transformed into a meaningful learning process rather than a simple editing tasks. According to Hartshorn et al. (2023), repeated exposure to corrective feedback allows learners to notice recurring grammatical errors and gradually reduce them over time. We can reduce language errors from becoming stabilized through this approach. Learners have opportunities to repeatedly revise their writing draft and at the same time reflect on their mistakes and apply corrective knowledge in the next writing tasks. These principles provide theoretical frameworks to examine the effect of technology-mediated feedback especially Grammarly on the development of academic writing accuracy among ESL undergraduates.

Automated Writing Evaluation (AWE) tools are widely adopted in second language writing instruction as a result of increasing integration of educational technology into language learning. AWE systems provide learners with immediate feedback on various aspects of writing, such as grammar, spelling, punctuation, sentence structure, and language usage. Traditional teacher feedback requires considerable time to provide, while AWE tools enable students to receive corrective feedback instantly and revise their writing independently. Recent studies suggest that AWE systems can support writing development by increasing opportunities for revision and encouraging learners to engage more actively with the writing process (Wei et al., 2023; Jannath & Akter, 2026).

Grammarly is one of the most widely used platforms in higher education among the various AWE tools currently available. Learners receive automated feedback on grammatical accuracy, mechanics, clarity, and language use from Grammarly which allows them to identify and correct errors during the drafting and revision stages of writing. Studies show that when Grammarly was used as part of a structured revision process, learners make fewer surface-level language errors and have improved linguistic accuracy of their written texts (Wijaya, 2025; Setiawan & Alkhowarizmi, 2025). However, there are several limitations with AWE systems. Immediate corrective suggestions can make learners highly dependent on automated corrections provided by Grammarly or simply accept recommendations without analysing the underlying grammatical rules. The effectiveness of Grammarly correlates with learners' ability to critically interpret and apply the suggestions during revision phase as well as the quality of the feedback provided. Therefore, Grammarly should be incorporated as a supplementary instructional tool in ESL academic writing instruction rather than replacement for teacher feedback.

The growing adoption of Grammarly in academic writing classrooms, lead researchers to increasingly examine its effectiveness in improving writing accuracy among ESL and EFL learners. Previous studies suggest that Grammarly positively contributes to the reduction of grammatical errors which is prevalent only when there is active engagement between the learners and the feedback provided during the revision process. Learners are able to identify language errors and efficiently revise their writing when Grammarly as an automated writing evaluation tool offers immediate corrective feedback.

There are findings from several empirical studies that highlighted improvements in grammatical accuracy after using Grammarly. Automated feedback significantly improved intermediate EFL learners writing accuracy, particularly when students revised their drafts based on the feedback received (Dehkordi et al. (2025). Similarly, Ng et al. (2025) reported that targeted feedback on language use and mechanics from Grammarly assisted learners in reducing grammatical

errors in academic writing. In addition, writing accuracy improved if corrective feedbacks are available and when learners are willing to review, comprehend and apply corrective suggestions provided by automated evaluation tools during revision process (Lekamge & Smith, 2025).

Though the studies indicate positive findings there is an emphasis by the researchers that the effectiveness of Grammarly is influenced by how learners interact with the feedback. A particular draft could be improved by simply accepting automated corrections without the necessarily leading to long-term language development. The importance of learner engagement and informed revision practices when using automated feedback systems have been highlighted in recent studies (Zhang & Hyland, 2023). Nevertheless, available evidence suggests that Grammarly could improve writing accuracy. Further controlled studies should be employed to determine its effectiveness among ESL undergraduate learners in authentic academic writing contexts.

The quality of the feedback and the way learners engage with the suggestions generated by the system affect the effectiveness of Grammarly-assisted feedback. When learners actively review, interpret and apply the feedback received during the revision process, they achieve greater improvements in writing. Zhang and Hyland (2023) mentioned that learners must reflect and make informed revisions to their writing for meaningful improvement to occur rather and result in learning does not automatically observable when simply receive corrective feedback.

One of the important factors that influence learners' engagement with Grammarly feedback is digital literacy. Stronger digital literacy helps learners to evaluate automated suggestions critically and decide whether the recommendations are appropriate or not for the specified writing context. Students with active engagement with AWE feedback showed higher levels of revision quality and improved their writing in comparison to those accepted corrections without reflection (Zhang & Hyland, 2023). Similarly, studies on Grammarly and other automated correction tools suggest learners get the most benefits when treat AWE tools as a learning resource rather than just a proofreading mechanism (Lekamge & Smith, 2025; Setiawan & Alkhowarizmi, 2025).

Grammarly serves as a tool to immediately detect errors nevertheless one should not overly depend on it. Overreliance on automated feedback systems inhibit learners independent editing skills and deeper grammatical understanding. Teachers play an important role to guide students to interpret feedback appropriately and develop long-term writing competence.

There are a few limitations in the existing literature though the studies on Grammarly and AWE tools are growing in numbers. First, studies conducted in ESL undergraduate contexts are limited and many studies have been conducted in EFL settings. Therefore, findings cannot be generalized to learners who use English as second language. In addition, fewer studies related to Grammarly focused on examining writing accuracy such as grammatical error rates, error-free T-units, and syntactic accuracy compared to those focusing on learners' perceptions, attitudes, satisfaction, or writing quality (Ng et al., 2025; Wijaya, 2025).

Second, often descriptive, survey-based or exploratory research designs were employed in previous studies. These research designs limit the researchers from making clear conclusions on the effectiveness of Grammarly in improving writing accuracy. Other than that, not many studies adopted controlled intervention designs to compare teacher feedback with feedback from Grammarly using measurable linguistic indicators (Dehkordi et al., 2025). As a result, it is difficult to establish to what extent Grammarly contributes to improvements in writing accuracy.

Finally, limited empirical evidence is available regarding the effectiveness of Grammarly among Malaysian ESL undergraduate students. Differences in educational context, language proficiency, and instructional practices may influence how learners engage with automated feedback and apply corrective suggestions during revision. Therefore, further classroom-based studies employing controlled research designs and objective accuracy measures are needed to determine the impact of Grammarly-assisted feedback on academic writing accuracy among ESL undergraduate learners. This study addresses these gaps by examining the effectiveness of Grammarly-assisted feedback through a quasi-experimental comparison with teacher feedback using multiple indicators of writing accuracy.

2. RESEARCH METHODOLOGY

In this research, the quasi-experimental pre-test and post-test control group design was employed in order to investigate the impact of written feedback from Grammarly on the accuracy of academic writing. Two intact undergraduate writing courses were picked. One of the classes became the experimental one which employs Grammarly-assisted feedback and the other became the control one that employs teacher-only feedback. Quasi-experimental design is employed to preserve natural, undisturbed classroom ecosystem as well as institutional constraints. Though the lack of true randomization leads to the potential of selection bias, ecological validity is prioritized through this approach to ensure that the observed interaction of Grammarly reflect authentic academic environment. Selection bias is minimized through the comparison of pre-test writing accuracy scores of experimental and control groups prior to the intervention. The population possessed comparable levels of writing proficiency before treatment as there are no statistically significant differences between the groups at baseline.

The groups were exposed to a pre-test writing exercise at the very start of the research and an after-treatment exercise. The design will enable comparison of improvement in each group and differences across groups. There are also similar experimental designs, which were applied in the studies comparing AI-mediated and teacher-mediated corrective feedback in the context of L2 writing (Abdi Tabari et al., 2025). Moreover, the structured pre-test and post-test designs

were also used in the generative AI feedback research studies investigating how many times participants revised and how well they wrote, which were also used to determine the change in performance over time (Ng et al., 2025)

The independent variable was the feedback type which are from Grammarly and teacher. The dependent variable was the accuracy of writing in academic writing, which was measured by distinct linguistic indicators.

2.1 Participants

The study population consisted of 64 ESL undergraduate students, who were taking a mandatory course in academic writing at a private university. The students were second-year students and had intermediate levels of proficiency in English according to institutional placement tests. The experimental group had 32 students, and the control group also had 32 students. The instructor used to teach both classes to minimize variation due to teachers. The sample size is similar to earlier AI writing intervention experiments, which were carried out in institutions of higher learning, and class-based samples were employed to test the application of automated writing instruments in EFL/ESL teaching (Alshehri, 2025). During the process of data collection, all participants agreed to participate in the study.

2.2 Instruments and Writing Tasks

Two academic writing tasks were administered to students in the form of pre-test and post-test under timed conditions. All the students had to write a 300-word argumentative essay on a broad scholarly topic that would be appropriate to write on at the undergraduate level. The 300-word essay length was intentionally selected as a controlled diagnostic instrument. The main purpose of the writing task was to obtain a manageable and comparable sample of students' writing for linguistic accuracy analysis and it was not to evaluate the overall quality of extended academic essays of ESL undergraduate students. Similar controlled writing tasks have been employed in second language writing research to facilitate the systematic calculation of grammatical errors, error-free T-units, and syntactic accuracy measures. The standardized word limit also ensured consistency across participants and reduced potential variation resulting from essay length (Barkaoui & Woodworth, 2023; Broussard et al., 2001). The themes were similar in level and format so that they could be compared. Regular writing assignments in AI and AWE research are based on the principle of argumentative writing, as it is possible to conduct systematic assessments of language accuracy and structure (Zhang & Hyland, 2025; Haddadian et al., 2025).

a. Accuracy Measures

Three indicators were used to test the accuracy of writing:

b. Grammatical Error Ratio (GER):

Division of the total number of grammatical errors by the total number of words x 100 (error per 100 words). It has been applied in AI-based accuracy research to monitor the decrease in error rate (Dehkordi et al., 2025).

c. Error-Free T-Units (EFTU):-

Share of T-units that are grammatically correct. T-unit analysis is typically applied in writing research in L2 that contrasts the feedback type (Abdi Tabari et al., 2025).

d. Syntactic Accuracy Rate:

Ratio of the correctly constructed sentences to the total number of sentences. The action is consistent with the studies that investigate the stabilized grammatical mistakes and dynamic corrective feedback (Hartshorn et al., 2023)

e. Errors were classified in a certain way:

1. Article errors
2. Verb tense errors
3. Mistakes in subject-verb agreement.
4. Preposition errors
5. Word form errors

All the essays were coded by two trained raters. To be sure that there was inter-rater agreement, it was computed before final scoring.

2.3 Treatment Procedure

The treatment lasted 8 weeks. The two groups had four major writing assignments that were done during this time.

a. Experimental Group (Feedback from Grammarly)

Real-time feedback on grammar, spelling, punctuation, sentence structure, word choice and writing clarity was provided to the students in the experimental group using Grammarly, an automated writing evaluation (AWE) tool. Students submitted their essay drafts to Grammarly and reviewed the automated feedback generated by the system. Only language-related recommendations were considered during revision, while content development and idea generation remained the responsibility of the students. The system gave automatic recommendations on grammar, sentence structure, and other language uses. Students were required to:

1. Review Grammarly feedback
2. Revise their draft
3. Submit a second version

The settings of Grammarly were kept consistent throughout the intervention. Students were instructed not to use external generative AI tools such as ChatGPT during the revision process and are only allowed to access Grammarly platform using the standard academic writing mode. The feedback provided by Grammarly served solely as a

corrective support mechanism for language accuracy.

Every assignment had two revision cycles, and this enabled interaction with the feedback repeatedly. Similar AI-assisted writing training models have been used in previous L2 writing research (Luo et al., 2025). Multi-draft revision structures were also observed in comparison between AI and teacher feedback processes (Abdi Tabari et al., 2025).

The frequency of revision that is number of changes made in a draft was noted to be further analyzed, which is in line with research on generative AI revision (Mekheimer, 2025).

b. Control Group (Teacher Feedback)

The instructor provided the corrective feedback to the control in a written form with the help of a standardized correction code. Students made one revision of their drafts before each assignment due to the teacher's comments. Grammarly was not used. The two groups were given equal time for instruction and writing prompts.

2.4 Data Analysis

Quantitative statistical methods have been used to analyze data.

- T-tests were performed in pairs to compare pre-test accuracy and post-test accuracy scores on each of the groups to identify the difference between the two.
- The results of the post-test in the experimental and control groups were compared with independent-sample t-tests to establish whether there was a situation where Grammarly-assisted feedback led to a significantly large improvement.
- The measurement of the value of improvement was determined using the effect size where Cohen's d effect size was calculated and interpreted using conventional benchmarks (0.20 = small, 0.50 = medium, 0.80 = large).

The statistical method adheres to quantitative investigation models that are applicable in the studies of an AI writing tool that measures learner development and change in performance (Rahim & Gregory, 2025). All statistical analyses were done using statistical software, and the level of significance was $p < .05$.

3. RESULTS AND DISCUSSION

3.1 Pre-Test Comparison

Each of the two groups was pre-tested with a 300-word argumentative essay before the therapy started. The findings revealed that there was no difference between the two groups in the initial writing accuracy in the study. The experimental condition where the feedback is from Grammarly obtained a mean error ratio of grammar of 14.8 errors per 100 words (SD = 2.1). The mean of control group that received teacher feedback was 15.1 errors (SD = 2.3) in 100 words. An independent-samples t-test revealed no significant difference between the two groups at pre-test ($t(62) = 0.54, p = .59$).

There was also a similarity in the percentage of error-free T-units. Error-free T-units in the experimental group were 38%, and those in the control group was 36%. This was not a statistically significant difference ($p > .05$). The results have shown that the two groups have had a similar level of accuracy.

3.2 Post-Test Improvement

At the end of the 8-week treatment, the two groups had improved. The enhancement was, however, higher in the Grammarly-assisted group. The experimental group has decreased the grammatical error ratio by 14.8 to 8.9 per 100 words, which is a 39.8 error reduction. The paired-samples t-test indicated that this was statistically significant ($t(31) = 11.42, p < .001, d = 1.95$). On the other hand, the control group minimized its error percentage to 11.7 errors per 100 words, which decreased by 22.5%. This difference was also high ($t(31) = 6.18, p < .001, d = 1.09$), but not as large as the effect of the experimental group.

In case of error-free T-units, the experimental group obtained an improvement of 38 to 61, and the control group obtained an improvement of 36 to 48. The comparison of the results in the post-test showed that the difference between groups was significant ($t(62) = 3.84, p < .001$), with the gains being stronger in the Grammarly-assisted group.

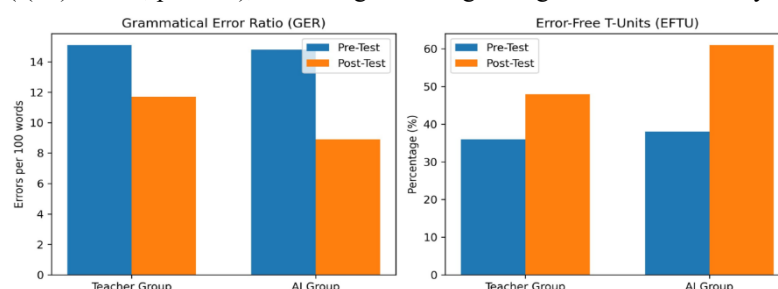


Figure 1. Pre-test and post-test writing accuracy outcomes for the teacher-feedback and Grammarly feedback groups

3.3 Between-Group Comparison

A t-test involving independent samples of post-test accuracy scores showed that the Grammarly-assisted group performed significantly higher than the teacher-feedback group ($p < .01$). The results can be attributed to the evidence that feedback from Grammarly may result in the objective improvement of writing accuracy (Dehkordi et al., 2025). They are also

consistent with the comparative studies that report differences between AI-mediated and teacher-mediated corrective feedback (Abdi Tabari et al., 2025). On the whole, the findings reveal that written feedback from Grammarly yielded more development in academic writing accuracy in ESL undergraduates.

3.4 Interpretation of Findings

The research results indicate that written feedback with the assistance of Grammarly has led to a significant increase in the grammatical accuracy of students. The first reason is that the feedback coming out of AWE tools specifically Grammarly provides instant correction, thereby enabling the students to be aware of their mistakes and amend them more attentively. Repeated exposure to feedback on drafts is likely to make learners identify patterns in their errors. This finding aligns with Shen (2025), found that effective interaction between human and AI encourages learners to engage cognitively with automated feedback. This interaction led learners to opt for more informed decisions rather than passively accepting of AI-generated suggestions. Although the improvement observed in the experimental group was substantial, the findings should be interpreted with caution. Students get repeated exposure to automated corrective feedback during the eight-week intervention period. It could have accelerated the reduction of surface-level grammatical errors. However, the present study was designed to measure writing accuracy outcomes rather than long-term language acquisition. Therefore, it cannot be conclusively determined whether the observed improvements reflected durable grammatical learning or effective utilization of Grammarly-generated corrections during the revision process. This is in favor of the concepts of Dynamic Written Corrective Feedback (DWCF) that focus on repetition feedback and revision to minimize errors that have been stabilized (Hartshorn et al., 2023). The findings are also consistent with the evidence that the AI dynamic feedback can enhance the writing performance when the learners actively revise their writing (Hartshorn & Pack, 2026).

3.5 Comparison with Previous Research

The current results align with the research of Mekheimer (2025), who discovered that writing performance was better in cases of generative AI-assisted feedback with the high frequency of revision. The findings also further add to those of Zhang and Hyland (2023), who reported that automated feedback combined with human feedback worked out to improve the quality of writing. Nonetheless, the study provides some specific evidence dedicated to undergraduate ESL students and the quality of academic writing, as opposed to the overall writing quality or interest.

3.6 Pedagogical Implications

The findings indicate that the feedback provided with the assistance of Grammarly can be applied as a support resource in higher education writing classrooms. It is not to be used to substitute the teacher's feedback, but can lessen the teacher's workload and offer instant correction. Nevertheless, learners should be taught about the interpretation and use of Grammarly feedback. Digital literacy is significant to the level of successful interaction between the learners with automated feedback systems (Zhang & Hyland, 2023). Teachers must teach learners to critically use AWE tools rather than blindly accept the feedback provided. Grammarly might primarily function as AWE tool yet it is essential to understand that it can also make mistakes where suggestions may seem to be inappropriate, inaccurate or divert from the writer's intended meaning. Therefore, students must be taught to evaluate the recommendations generated by Grammarly before make any corrections. This is where teachers play important role in guiding learners to make informed revision decisions.

3.7 Limitations of the study

The findings of the study should be interpreted with several limitations. First of all, the intervention was conducted over a relatively short period of eight weeks. This limits the conclusions regarding the long-term retention of grammatical accuracy gains. Second, the learners' cognitive processes while using Grammarly were not examined during the study, instead it focused exclusively on quantitative measures of writing accuracy only. As a result, it is unclear whether students developed a deeper understanding of grammatical rules or simply implemented the corrections suggested by the system. Interviews, or reflective journals could be employed in future studies to investigate how learners interpret and apply feedback generated by Grammarly. Finally, the study was conducted within a single institutional context, thus limiting the generalizability of the findings to other ESL learning environments.

4. CONCLUSION

This study investigated the effects of Grammarly-assisted written feedback on the academic writing accuracy of ESL undergraduate students. The results showed that undergraduates that received written feedback from Grammarly demonstrated improvements in grammatical accuracy and error-free T-units compared to those received feedback from the teacher only. This suggests that writing revision in ESL contexts can be supported by Grammarly as a supplementary pedagogical tool with careful integration. However, the study did not examine the retention of the writing accuracy after Grammarly support was withdrawn or the corrections provided are fully understood by the students. Therefore, it is crucial to interpret the findings cautiously and learners should be encouraged to evaluate the feedback provided by AWE tools rather than solely depending on it. There is a need for future studies which investigate the effects of more time and whether the results of the accuracy gains can be sustained over time with longitudinal studies.

ACKNOWLEDGMENT

Thank you to those who have supported the implementation of this research.

REFERENCES

- Abdi Tabari, M., Kushki, A., & Wang, Y. (2025). Comparing the effects of teacher- and AI-mediated corrective feedback on accuracy, complexity, and quality in L2 written narratives. *Computer Assisted Language Learning*, 1–24. <https://doi.org/10.1080/09588221.2025.2561608>.
- Alshehri, M. (2025). Integrating Automated Writing Evaluation into Saudi EFL Students' Writing Practice. *Arab World English Journal*, 16(1), 70–87. <https://doi.org/10.24093/awej/vol16no1.5>
- Barkaoui, K., & Woodworth, J. (2023). An exploratory study of the construct measured by automated writing scores across task types and test occasions. *Studies in Language Assessment*, 26–63. <https://doi.org/10.58379/qcfs2805>
- Bekdaş, M. (2026). ChatGPT: the artificial intelligence for fostering grit in second language writing classes. *Humanities and Social Sciences Communications*, 13(1). <https://doi.org/10.1057/s41599-026-06557-w>
- Broussard, K. M., Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (2001). Second Language Development in Writing: Measures of Fluency, Accuracy and Complexity. *TESOL Quarterly*, 35(2), 342. <https://doi.org/10.2307/3587656>
- Cheng, X., & Xu, J. (2025). Engaging second language (L2) students with synchronous written corrective feedback in technology-enhanced learning contexts: A mixed-methods study. *Humanities and Social Sciences Communications*, 12(1). <https://doi.org/10.1057/s41599-025-05007-3>
- Dehkordi, M. B., Riahipour, P., & Esfahani, F. R. (2025). Charting Uncharted Waters: AI-Driven Feedback and Writing Accuracy of Intermediate Iranian EFL Learners. *Research on English Language Education*. <https://doi.org/10.71673/relp.2025.1203616>
- Eragamreddy, N., & Joseph, R. (2025). Digital Tools and Writing Education: A Thematic Analysis of Technology's Role in Writing Skills Development. *Arab World English Journal*, 16(3), 3–25. <https://doi.org/10.24093/awej/vol16no3.1>
- Haddadian, G., Kim, M. K., & Haddadian, N. (2025). A Systematic Review of Automated Writing Evaluation Tools in Argumentative Writing for English as Foreign Language Education. *Research Synthesis in Applied Linguistics*, 1–51. <https://doi.org/10.1080/29984475.2025.2598266>
- Hartshorn, K. J., & Pack, A. (2026). The effects of artificial intelligence-based dynamic written corrective feedback on second language writing and user sentiment. *RELC Journal*. <https://doi.org/10.1177/00336882251405498>
- Hartshorn, K. J., Rice, S. H., Eckstein, G., & Evans, N. W. (2023). Dynamic written corrective feedback frequency and its effects on ESL writing fluency, accuracy, and complexity. *Feedback Research in Second Language*, 1, 7–32. <https://doi.org/10.32038/frsl.2023.01.02>
- Jannath, H., & Akter, A. (2026). Artificial Intelligence (AI) for Automated Writing Evaluation in TESOL: Effectiveness and Challenges. *Journal of Languages and Language Teaching*, 14(1), 273–287. <https://doi.org/10.33394/jollt.v14i1.18165>
- Lekamge, R., & Smith, C. (2025). Impact of Auto-Correction Features in Text-Processing Software on the Academic Writing of ESL Learners. *ELOPE: English Language Overseas Perspectives and Enquiries*, 22(1), 69–91. <https://doi.org/10.4312/elope.22.1.69-91>
- Luo, L., & Yusuf, A. (2025). Bridging AI and pedagogy: how AI-adaptive feedback shapes Chinese EFL students' writing engagement, metacognitive writing strategies, and writing performance. *Assessment & Evaluation in Higher Education*, 1–24. <https://doi.org/10.1080/02602938.2025.2548919>
- Luo, M., Hu, X., & Zhong, C. (2025). The collaboration of AI and teacher in feedback provision and its impact on EFL learner's argumentative writing. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-025-13488-7>
- Luo, Ying, Yuwei Xia, and Xiaofei Lu. 2025. "A Systematic Review of the Role of Artificial Intelligence in Second Language Writing Education." *Digital Studies in Language and Literature* 2(2):302–25. doi:10.1515/dsll-2025-0001.
- Mekheimer, M. (2025). Generative AI-assisted feedback and EFL writing: a study on proficiency, revision frequency and writing quality. *Discover Education*, 4(1). <https://doi.org/10.1007/s44217-025-00602-7>
- Ng, W., Yunus, M. M., & Ismail, H. H. (2025). Investigating the effect of Grammarly on enhancing writing skills in ELT classrooms: A systematic literature review (2020–2024). *International Journal of Academic Research in Progressive Education and Development*, 14(1), 1454–1476. <https://doi.org/10.6007/IJARPED/v14-i1/24774>
- Rahim, S. A., & Gregory, N. A. (2025). AI Writing Tools as Catalysts for L2 Writing Development: A Quantitative Investigation of Students' Experiences. *International Journal of Academic Research in Business and Social Sciences*, 15(9). <https://doi.org/10.6007/ijarbss/v15-i9/26490>
- Setiawan, F & Alkhowarizmi, A. (2025). Exploring an Artificial Intelligence as Automated Feedback Program in EFL Writing. *ETERNAL (English Teaching Journal)*, 16(1), 202–224. <https://doi.org/10.26877/eternal.v16i1.1206>
- Shen, F. (2025). A Study of Human-AI Interaction Patterns in Artificial Intelligence-Assisted Second Language Writing. *International Journal of English Literature and Social Sciences*, 10(2), 041–052. <https://doi.org/10.22161/ijels.102.8>
- Nguyen, H. M., & Chu, T. A. (2024). Written Corrective Feedback in Second Language Writing: A Review of Research. *REFlections*, 31(2), 858–870. <https://doi.org/10.61508/refl.v31i2.275269>
- Wei, P., Wang, X., & Dong, H. (2023). The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: a randomized controlled trial. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1249991>
- Wijaya, K. F. (2025). The Role of Grammarly as an Automated Writing Evaluation Tool in EFL Writing Learning Process. *Datokarama English Education Journal*, 6(2), 61–76. <https://doi.org/10.24239/dee.v6i2.127>
- Zhang, Z. V., & Hyland, K. (2023). The role of digital literacy in student engagement with automated writing evaluation (AWE) feedback on second language writing. *Computer Assisted Language Learning*, 1–26. <https://doi.org/10.1080/09588221.2023.2256815>