



Automated Writing Evaluation Tools in the Improvement of the Writing Skill

Lorena Parra G.

Assoc. Prof. Faculty of Education, Language Department, Universidad Técnica de Ambato, Ecuador, lparra@uta.edu.ec

Ximena Calero S.

Assoc. Prof. Faculty of Education, Language Department, Universidad Técnica de Ambato, Ecuador, xcalero@uta.edu.ec

The improvement of the writing skill with the help of modern computer programmes has gained considerable attention in the English as a Foreign Language writing learning process and the better understanding about the contributions of these technological programmes are needed in the field of education. This paper sought to shed light on the effects of implementing free Automated Writing Evaluation tools on the writing performance of the students of an English Teacher Training Programme. Ecuadorian undergraduate learners were encouraged to use Grammarly and Grammarly as free Automated Writing Evaluation tools in the improvement of the writing skill as part of the experimentation of this study. These AWE tools were used as complement of teacher's feedback given to 28 learners on their writing performance during a semester of studies. Quantitative method was used to conduct this practical research by applying a t-test technique and a survey. A pre-test/post-test experimental research design was applied to examine students' improvement on the writing skill. The results revealed positive effects on the improvement of the students writing. This study confirms the benefits of free Automated Writing Evaluation tools in the improvement of the writing skill as it is discussed in the conclusion of this paper followed by some recommendations.

Keywords: automated writing evaluation tools, writing skill, EFL students, learning, evaluation tools

INTRODUCTION

Automated Essay Scoring (AES), Automated Writing Evaluation (AWE) or Automated Essay Evaluation as they are referred, are innovative computer tools used in different contexts as in the educational system used to give evaluative feedback in the writing

Citation: Parra G., L., & Calero S., X. (2019). Automated Writing Evaluation Tools in the Improvement of the Writing Skill. *International Journal of Instruction*, 12(2), 209-226. <https://doi.org/10.29333/iji.2019.12214a>

process. “Automated essay scoring (AES) is the ability of computer technology to evaluate and score written prose” (Shermis & Burstein, 2003, p. 37). This technological program has the capability to provide immediate qualitative and quantitative feedback by scoring the text, analyzing the text structure and generating a detailed evaluation of the written text (Khoii & Doroudian, 2014).

Early work in automated scoring application has been developed since 1960 with the purpose of saving time when grading written work and serving as a support for teachers when generating feedback on the students’ essays (Wilson & Czik, 2016). The design of early programs to evaluate writing features has improved rapidly since the mid-1990s thanks to the contribution of artificial intelligence technology which has served positively in the process of natural language and the tutoring of intelligent language system (Chen & Cheng, 2008).

In the field of language education, computer technology has contributed with new trends of language instruction and language assessment. Natural language processing technology is actively used not only for holistic scoring, but also for purposes of writing evaluation (Feng, Saricaoglu, & Chukharev-Hudilainen, 2015). One of these new implementations of technology in language training is automated writing evaluation (AWE) which has provided noteworthy support to meet from the needs of diagnostic feedback to the needs of summative assessment on aspects of writing (Warschauer & Grimes, 2008). These type of systems generate immediate diagnostic on a large number of essays, contributing to the teachers’ work reduction (Bai & Hu, 2017). Among the advantages AWE technology provides to its users is consistent explanations and immediate feedback (Hyland & Hyland, 2006; Warschauer & Ware, 2006). According to Zhang & Hyland (2018) these two characteristics have allowed to break the barriers of time and accelerate the feedback practice allowing the students to develop their writing skill at their own pace which can benefit the autonomous participation and interaction in the language classroom (Liao, 2016; Warschauer, 2000). Furthermore, the consistency and objectivity with which this tool evaluates the students’ writing performance makes the learners to improve the writing mechanics and accuracy (Wang, Shang & Briody, 2013).

The adoption of artificial intelligence system has been increasing rapidly in language learning process in the last decade generating doubts in the ability of computers to evaluate accurately and effectively written text which demands more investigations on the effectiveness of these tools on the students’ improvements in their writing skill. In Wang, Shang & Briody (2013), the use of AWE limits the students’ performance in a real and meaningful interaction which can affect the appropriate learners’ improvement in the overall writing process. Moreover, AWE program has become a writing assistance tool due to its online resources and editing features that facilitates the work for the students which may have a negative influence in the autonomy learning (Chen & Cheng, 2008).

As the implementation of these computer programs in the learning language system around the world is being used more often for different reasons as to achieve the desirable outcomes in the development of writing skill, further research is necessary to

investigate the implementation of AWE programs in the field of English as a Foreign Language (EFL) and the benefits these computer tools provide in terms of improving students' overall writing (Liu & Kunnan, 2015). Furthermore, most of the studies have been done in relation to commercial AWE tools being paramount to investigate the use of free AWE tools which are available for people who need support from these programs. (Ranalli, Link, & Chukharev-Hudilainen, 2017)

The present study aimed to examine the EFL undergraduate students' perceptions and effects of using AWE program as a medium to enhance their writing performance. To carry out this research, two Free Automated Writing Evaluation Tools as Grammark and Grammarly were used on the writing performance of Ecuadorian students of an English Teacher Training Program. In view of the study purpose, the following research questions were posed:

1. To what extent will the AWE tools help the students to improve their writing performance?
2. Are there any significant writing improvement differences between the students who use Grammark and Grammarly?
3. What are the students' attitudes toward using AWE tools?

LITERATURE REVIEW

Assessing writing

According to many teachers' perceptions, writing is one of the productive skills of the language most difficult to learn and therefore to teach that demands careful attention on accuracy due to its complex process of communication (Hyland & Hyland, 2006). Teacher and peer feedback is not always consistent due to human error and for that reason identifying the same written problems is not possible, leading students to be confused on the messages they received from their teachers (Lavolette, 2015; Zhang, 2016; Ranalli et al., 2017;). For language teachers, the learners' outcomes are essential to give appropriate instructional feedback that provides information about the levels of correctness as well as means of improvement (Shim, 2013). However, instructional feedback may result in a process that demands an enormous workload for teachers who have to correct a large number of written work done by their students and therefore, it becomes time-consuming, instructor-centered and problematic to provide accurate and consistent holistic feedback (Warschauer & Grimes, 2008; Chapelle, Cotos, & Lee, 2015; Wilson & Czik, 2016).

In response to the need of breaking the barriers that teachers face with instructional feedback, research on the effects of using AWE program as a tool that supports writing improvements has been carried out during the last decade.

Automated Writing Evaluation

Many researches have been involved in the development of computer programs that contribute with the scoring and feedback of the writing skill. In Burstein, Chodorow, & Leacock (2004), writing is a language ability that can be best improved by frequent writing, appropriate and immediate feedback. Thanks to new technological inventions as

the AWE computer program, these checking processes have been automated serving as a support to teachers and as a tool that provides freedom and planning time to students in increasing their level of motivation (Shim, 2013).

The use of Automated Writing Evaluation has been increasing as a teacher assistant tool that provides high-level of feedback and writing quality (Wilson & Czik, 2016). This computer program helps to improve learners' writing quality due to its fast and individualized feedback with explanations of grammar, spelling, sentence and word usage which can contribute to learners' autonomy (Wang, Shang & Briody, 2013).

Several studies that promote the use of AWE have proven the benefits of this tool to the writing improvement in three dimensions. First, word processing facilitates the aspects of editing and revising grammar and spelling which contributes to learners' awareness in their writing (Wang & Wang, 2015). Second, an error correction program gives the students the opportunity to inspect their errors immediately and teachers the chance of interacting with their learners in specific error correction and feedback (Shim, 2013). Moreover, computerized feedback concentrates the learners' attention on sentence-level error which encourages them to enhance inaccurate usage and their capability to identify and reformulate errors when no human support is available which can encourage the autonomy learning (Wang, 2013). Third, artificial technology systems claim to be more objective and accurate when grading standardized essay tests, as human markers in the typical test score diverge by some points needing a third marker to have a final grade agreement (Warschauer & Grimes, 2008). In addition, human feedback is flexible and limited according to student's background and needs. However, the AWE can check large groups of essays and score them accurately due to the Latent Semantic Analysis technique used by newer AWE systems which compare semantic content of words used in essays (Khoii & Doroudian, 2013).

On the other hand, some researchers doubt the effectiveness and benefits of the usage of AWE pointing some disadvantages and limitations of this tool. In Chou, Moslehpour, & Yang (2016), AWE is restricted to give quantitative information on the length sentences, word distribution, word repetition and statistical analysis which leads to give general judgments and unclear feedback. In addition, AWE is limited to evaluate content, specific prompts of the program, process of writing and it discriminates students who are not familiar with technology (Khoii & Doroudian, 2013). In Burstein et al. (2004), teachers' feedback is the best way to enhance and motivate students' writing due to the capability of context analysis in which humans have to evaluate a process as opposed to a computer tool that evaluates according to what has been programmed to detect. Furthermore, AWE provides generic suggestions which is a limitation to contribute with formative learning, rich negotiation meaning and content development (Chen & Cheng, 2008).

Findings from previous studies evidenced positive and negative results of using commercial AWE tools; however, little research on the effects and perceptions of free AWE tools used to improve the writing skill has been done. This study attempts to examine the results of using free AWE tools as instruments that can contribute to the improvement of the learners' writing skill.

METHOD

This study involves quantitative research as data collection procedures resulted in numerical data that was analysed by statistical methods (Dörnyei & Griffee, 2010). A pre-test/post-test experimental research design was used considering subjects were assigned randomly to groups. The manipulation of one or more variables (independent variables) to determine the effect on another variable (dependent variable) corresponds to experimental studies (Mackey & Gass, 2015). Each experimental group worked with one Free Automated Writing Evaluation Tool (Grammarly and Grammarly) respectively to perform their writing skills in class hours and as outside-classroom activities during the semester of studies. The learners' writing improvement in the post-test was compared for both groups to identify whether there were differences between the groups in respect to the treatment. A survey with a Likert scale design was conducted to determine the learners' attitude towards the use of AWE tools to improve the writing skill.

Participants

The participants of this study were 28 undergraduate students (10 males and 18 females) of the English Teacher Training Program of the Faculty of Social Science in a public university in Ecuador. The population had been studying at the English Teacher Program for a year and a half, this was their third level of studies and the first time they took the subject "Writing I" as a part of their academic formation. This academic subject aims to give the students the concepts and skills they need to success in academic writing. The ages of the students ranged from 19 to 23 years old. Their English language proficiency levels were A2, B1 and B2 according to their average pre-test scores on Versant test. The researcher was assigned to teach Writing I during the whole term of class.

Instruments

To carry out this investigation, three main instruments served for data collection. First, the students were given the official Versant standardized placement test from Pearson as a pre-test and post-test to measure their overall language proficiency and specifically their writing performance. The test comprehends nine parts: A: Read Aloud, B: Repeat, C: Sentence Builds, D: Conversations, E: Typing, F: Sentence Completion, G: Dictation, H: Passage Reconstruction, and I: Summary and Opinion (Pearson, 2016). The scores specifically for the writing section were considered. Second, the students' writing samples and the reports obtained from the AWE tools Grammarly and Grammarly were analyzed to verify the students' improvement on the writing skill (see Appendix A). Both tools are open grammar checkers that generally detect potential mistakes in writing. Grammarly automatically detects wordiness, use of articles use of conjunctions, spelling, punctuation, word choice, and style (see Fig. 1). Grammarly finds problems related to passive voice, word phrase, run-on sentence, and transition (see Fig. 2). Third, a questionnaire was designed and adapted by the studies of Wang, Shang & Briody (2013) and applied to the students to know their attitude regarding the use of the AWE tools. The reliability of the Likert – type items in the questionnaire was measured

by using the Cronbach's alpha with a result of 0.74 showing that the instrument is reliable.

Figure 1

Automated Writing Evaluation Tool – Grammarly

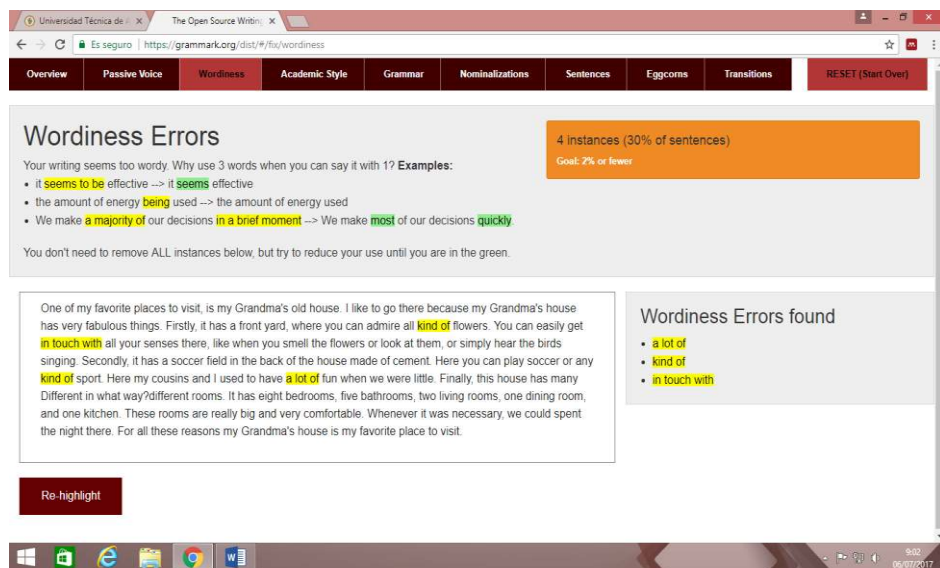
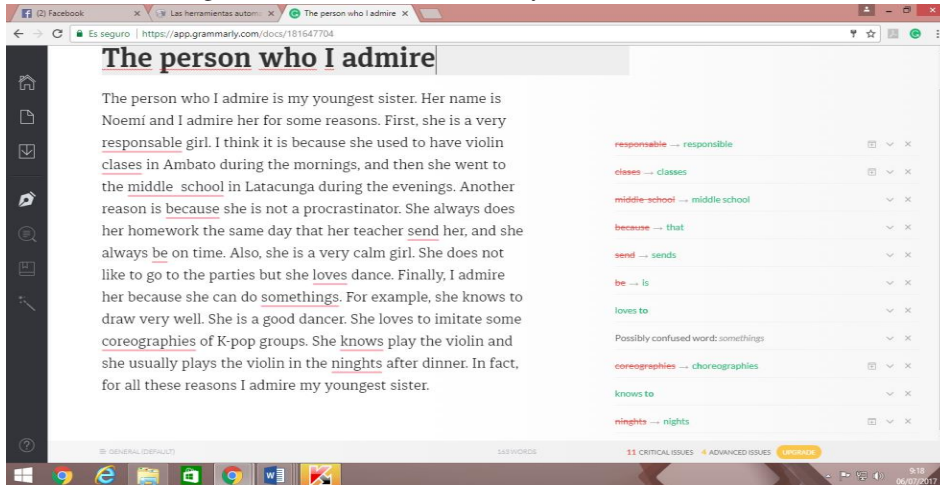


Figure 2

Automated Writing Evaluation Tool - Grammarly

At the beginning of the semester of March 2017, the official Versant test was applied to 28 students who participated in this study to have evidence of their language performance, taking their writing results as the starting point of this research. This

computed pre-test/ post- test experimental design was applied as a single group to guarantee its equivalence. Learners were assigned to work with Grammark and Grammarly, the free Automated Writing Evaluations tools. A total of eight weeks sessions of three hours were carried out during the treatment. At the end of the same semester, June 2017, the same official Versant test was applied as a post-test which was conducted in both groups, the students who used Grammarly and the others who used Grammark, paying particular attention to their writing results.

The two free AWE tools, Grammark and Grammarly, were socialized to the students in the computer laboratory that was used during the experimentation. The participants were divided randomly into two groups, the students to use Grammarly and the other students to use Grammark. The two free AWE tools were selected due to their ease of access to work without restrictions. The students were given the same topics to write their paragraphs in Word-processing programme. Each class, the students wrote a text and it was copied or uploaded to be corrected by the AWE tool.

The questionnaire applied to collect the students' attitudes after using the AWE tools was elaborated using a five-point agreement Likert scale for the first eight questions of the survey, and checkboxes selection for the last three questions of the instrument.

Data collection procedure

Data was collected with the participation of former students who gave their consent to be part of this study. The instructor of the course used the first hours of the classes to teach about the parts of paragraphs and essays, writing conventions and to provide feedback to the students. The participants of the course took the official Versant placement test from Pearson as a pre- test and post- test to measure their overall English performance, the results of the writing skill were considered at the beginning and at the end of the experimentation to compare the results. The participants were divided in two groups to work with the assigned tool respectively. Having the results of the pre-test, the learners were divided homogeneously in two groups to use the AWE tools. They were trained on how to use Grammarly and Grammark during the second hours of the classes in the computer laboratory. After being familiarized with the AWE tools, the learners were given common topics to write their compositions which were uploaded to the AWE tools to be corrected. The students stored their original and corrected writings to their e-portfolio created and shared by Google Drive with the instructor. The instructor gave feedback on the errors that the AWE tools did not provide explanations as for example when and why to use a specific verb tense. The first and final drafts were analyzed by the learners and the instructor to be aware of the detected errors by the used computer tools. At the end of the experimentation, all the participants responded to a questionnaire to obtain information regarding their attitudes toward using the AWE tools to enhance their writing performance.

Data analysis

This study employed quantitative method. First, for the purpose of evaluating the effect of applying the AWE programs on their writing performance, the results from the pre-test and post-test were analyzed by using a paired-samples t-test. Second, in order to see if there was any difference between Grammark and Grammarly use, an independent samples t-test was applied. Finally, descriptive statistics were used to indicate the students' attitude of using AWE tools to improve their writing skills.

FINDINGS

This study investigated the effects of using free Automated Writing Evaluation Tools, Grammarly and Grammark, on the Ecuadorian students of an English Teacher Training Programme. The scores obtained from the participants' overall Versant pre-test and post-test were compared. The results are presented on based of the research questions. (1) To what extent will the AWE tools help the students to improve their writing performance? (2) Are there any significant writing improvement differences between the students who use Grammark and Grammarly? (3) What are the students' attitudes toward using AWE tools?

The results of an independent t-test performed on the grades obtained by students on a Versant test prior to the implementation of the experiment showed that the groups were very similar in terms of writing (see Table 1), as a p value of 0.9058 was obtained. Table 1 shows the similarity between scores in the writing section of the Versant test of the students prior to the application of the treatment.

Table 1

T-test results for the group's homogeneity

Group	No.	Mean	St. Deviation	T Value	Sig. (2 tailed)
Group A	14	47.0714	14,31763	-0,11	0,9058
Group B	14	47.6469	10,71022		

To answer the first research question, a paired samples t-test was used to find any statistically significant differences in the results of the post-test compared with those of the pre-test for both groups. For Group 1, Table 2 shows the results of the paired samples t-test for the writing section of the Versant test.

Table 2

Paired t-test Group 1 (differences between pre- and post-test)

Group	Variable	Test	No.	Mean	St. Deviation	T Value	Sig.(2tailed)
Group 1	Writing Versant	Pre	14	47.0714	14,31763	-3.42	0.0044
		Post	14	57.2857	14.31763		

Table 2 shows that Group 1 improved their writing performance significantly in the writing section of the Versant test ($p = 0.0044$). The AWE tools helped students to identify errors on writing style, grammar and encourage them to edit their drafts by giving the learners immediate feedback which made them aware of their errors and

enhanced those suggesting possible alternatives that can make their writings more accurate.

Table 3
Paired t-test Group 2 (differences between pre- and post-test)

Group	Variable	Test	No.	Mean	St. Deviation	T Value	Sig. (2 tailed)
Group 2	Writing Versant	Pre	14	47.6429	10.71022	-3.38	0.0048
		Post	14	55.0714	7.97558		

Similarly, Table 3 shows that Group 2 also improved their writing performance significantly in the writing section of the Versant test ($p= 0.0048$). The participants made fewer errors in the post-test after using the AWE program. Fast and individualized explanation may have motivated the students for autonomy learning which can encourage users to analyze their errors with privacy.

Table 4
T-test result (differences between Group 1-and Group 2)

Group	Variable	Test	No.	Mean
Group 1	Writing Versant	Post	14	57.2857
Group 2	Writing Versant	Post	14	55.0714

To answer the second research question, an independent samples t-test was used to find any statistically significant differences in the results of the post-test in group 1 compared with those of the post-test for Group 2, and no significant differences were found. The two AWE tools provided immediate feedback on similar aspects of grammar, punctuation and writing style contributing to learners' awareness on detecting and reformulating their errors. Both methodologies had a positive impact on the learners as Grammark and Grammarly have similar characteristics to evaluate writing performance.

Table 5 shows the descriptive statistics of participants' responses to the post-instruction survey. Most participants agree or strongly agree on the analysis that grammar is useful ($M=4.08$) and that it is user-friendly ($M=4.38$), as well as a high percentage of respondents also agree or strongly agree that their English writing ability improved after using Grammark ($M=4.23$). It was also reported that the users were satisfied with Grammark, while some respondents were not sure that its error analysis of usage was useful ($M=3.85$).

Table 5
Descriptive statistics of the post-instruction survey of attitudes on Grammark

Statement	Median	Mean*	SD
1. I think its error analysis of grammar is useful (e.g., subject- verb agreement, possessive errors, wrong word, pronoun errors)	4.00	4.08	0.86
2. I think its error analysis of usage is useful (article, prepositions, word choice)	4.00	3.85	0.80
3. I think its error analysis of style is useful (e.g., repeated words, long/short sentences, passive voice)	4.00	4.15	0.69
4. I think its error analysis of organization development is useful	4.00	4.15	0.69
5. I think my English writing ability has improved after using Grammark	4.00	4.23	0.60
6. I think Grammark is user-friendly	4.00	4.38	0.65
7. I am willing to use Grammark again in the future if I have the chance	4.00	4.31	0.63
8. Generally speaking, I am satisfied with Grammark	4.00	4.08	0.76

Table 6 shows the descriptive statistics of the post-instruction survey of perceptions on Grammarly. Numerous participating students agree or strongly agree that their English writing improved after using Grammarly (M=4.00) and that this tool was user-friendly (M=4.00), as well as, they would like to use this tool again (M=4.20). However, respondents were not totally convinced that error analysis of usage was useful (M=3,93) nor they totally agreed of its analysis of style (M=3,67).

Table 6
Descriptive statistics of the post-instruction survey of attitudes on Grammarly

Statement	Median	Mean*	SD
1. I think its error analysis of grammar is useful (e.g., subject- verb agreement, possessive errors, wrong word, pronoun errors)	4,00	3,87	0,74
2. I think its error analysis of usage is useful (article, prepositions, word choice)	4,00	3,93	1,03
3. I think its error analysis of style is useful (e.g., repeated words, long/short sentences, passive voice)	4,00	3,67	0,90
4. I think its error analysis of organization development is useful	4,00	3,80	1,01
5. I think my English writing ability has improved after using Grammarly	4,00	4,00	1,07
6. I think Grammarly is user-friendly	4,00	4,00	0,85
7. I am willing to use Grammarly again in the future if I have the chance	4,00	4,20	0,68
8. Generally speaking, I am satisfied with Grammarly	4,00	3,87	0,52

On the same post-instruction survey, 3 additional questions were included, and their analysis are reported as follows: Figure 2 reports on the answers for question 9, "I think Grammark/Grammarly assisted me in the following aspects". An average of 64% of students agreed that the two AWE tools assisted them in vocabulary; the 78.5% of

respondents indicated to have had feedback on grammar as well as the 32% of learners stated that the two computer tools corrected the sentence coherence. 14.35% of participants informed to have received feedback on generating ideas. In terms of logic development, 40% of learners who used Grammarly agreed to have received assistance on this aspect, while the 23.1 % of students who used Grammarly identified this type of error correction. Moreover, Grammarly provided more feedback on mechanics and style than Grammarly did. On the other hand, Grammarly gave feedback on punctuation according to the 100% of participants, meanwhile 23% of users of Grammarly identified to have received feedback on this aspect.

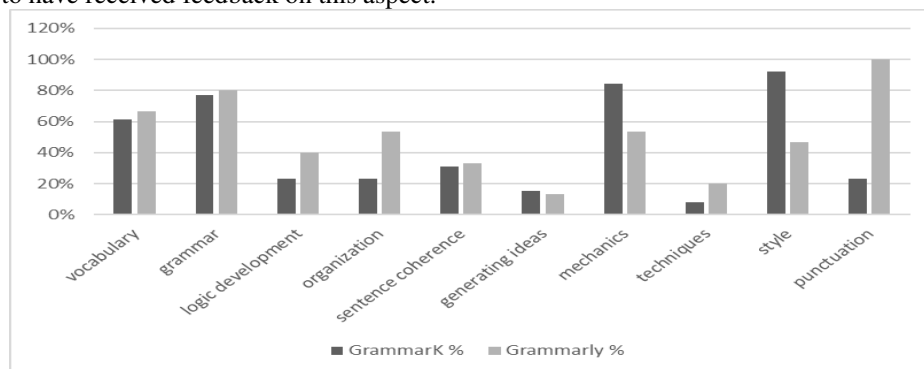


Figure 2

Comparison of percentage of respondents who believe that the grammar tools assisted them in various aspects of their English learning process during the instruction period.

Figure 3 reports on the answers for question 10, “What are the best parts of the grammar tool used?”. According to 64.3% of the participants who used Grammarly, this tool helped them to figure out grammatical errors that they had not noticed before, while for GrammarK, 46.2% of users agreed with this statement. Both AWE tools seem to have helped the learners to improve their writing skills as reported by 61.5% who used GrammarK and 71.4% who used Grammarly. Similarly, the participants agreed that GrammarK made them to understand writing (46.2%) as well as Grammarly did (50%). In terms of accuracy, Grammarly appears with 64.3% to have contributed more than GrammarK with 38.5% of agreement. Moreover, 57.1% of learners who used Grammarly indicated that this tool gave them instant feedback, it was faster to interact with this software and that they could examine their writings by themselves. Unlike 15.4% of users who stated that GrammarK provided instant feedback as well as 38.5 % of participants who reported that this software was interactive and allowed them to examine the writing by themselves.

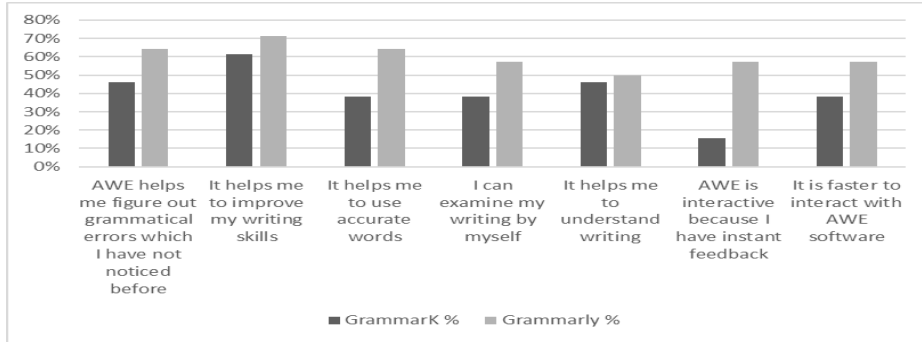


Figure 3

Comparison of percentage of respondents to the question: “What are the best parts of the grammar tools used?”

Figure 4 reports on the answers for question 11, “What are the worst parts of the grammar tool used” Participants who used both AWE tools stated that these computer programs did not have meaningful interaction (58%). For 33.3% of users of GrammarK and for 16.7% of users of Grammarly, it did not help them to revise their writing accurately. The participants reported that they could not organize their own learning with GrammarK (16.7%) nor with Grammarly (33.3). It was also indicated that a small percentage (8.3%) of learners who used Grammarly did not understand the tool, while another 8.3% who used GrammarK agreed it was a waste of time using AWE program. Moreover, a quarter of GrammarK users as well as 8.3% of Grammarly users indicated that AWE feedback was not a support for them.

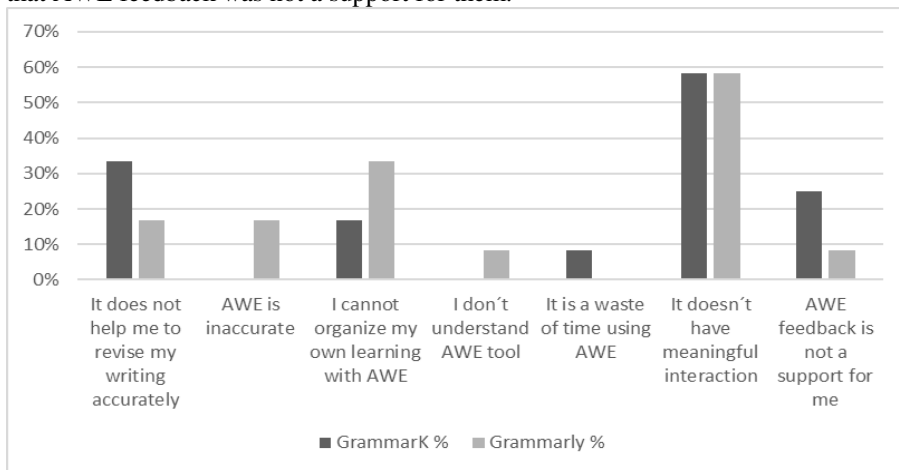


Figure 4

Comparison of percentage of respondents to the question: “What are the worst parts of the grammar tool used?”

DISCUSSION AND CONCLUSION

This study examined the effects and students' attitudes of using Grammarly and Grammarly AWE tools as a medium to improve the learners' writing performance. This study also aimed to investigate whether there was a significant difference between the pre-test and post-test to determine the extent the AWE tools helped them to enhance their writing performance. Grammarly and Grammarly were used to evaluate if there was a significant improvement difference between the two groups who used the computer programs respectively. The major results are discussed as follows.

First of all, the findings demonstrate that after using the AWE tools, the learners writing performance significantly increased according to the results of the pre-test and post-test. These results revealed that new trends of computer technology such as computer AWE tools helped to enhance the language learning process, language instruction and language assessment matching previous studies which state that processing technology is not only used for holistic scoring, but also for purposed of writing evaluation (Feng, Saricaoglu, & Churkharev-Hidilaonen, 2015). The learners made fewer errors after receiving immediate feedback provided by the AWE tools. Their awareness increased as word processing allows users to revise aspects of editing, spelling and grammar error in an autonomous way (Wang & Wang, 2015). Such findings confirm that AWE tools are computer programs that provide explanations along with pertinent diagnostic feedback that can contribute with the development of the writing skill (Dikli & Bleyle, 2014).

Secondly, the two AWE tools, Grammarly and Grammarly, provided feedback on similar aspects of writing that benefited the participants to improve their grammar, punctuation, mechanics, and style, and therefore, their writing accuracy. These findings confirm the t-test results which reveal the positive effects for the two groups of students who used the proposed AWE programs respectively (Wang & Wang, 2015). Such results match previous studies suggesting AWE as a classroom assistant tool that can improve learners' writing quality (Wang, Shang & Briody, 2013). However, for the aspects of writing that were not assisted by the AWE tools, teacher's feedback was necessary, affirming that the technology serves as an aid to enhance the learners' improvement. All the essays produced by the students were corrected in a moment, saving time and an enormous workload for the teacher which may become problematic to provide consistent feedback in a short time (Warschauer & Grimes, 2008; Chapelle, Cotos, & Lee, 2015; Wilson & Czik, 2016).

Thirdly, regarding the students' perceptions of using the AWE tools, the findings indicate that the learners hold positive attitudes toward the effect of using AWE on the improvement of some aspects of writing skills. The guidance received from the computer programs encouraged the learners to produce better written texts and to correct their errors by themselves at their own pace benefiting autonomous participation, and motivating to classroom interaction (Liao, 2016; Warschauer, 2000). Nevertheless, the lack of meaningful interaction between the students and the AWE was one of the weaknesses these computer programs revealed as well as the limitations on aspects of content and organization. These results support earlier research that advice

that computerized feedback concentrates the learners' attention on sentence-level error which encourage them to enhance inaccurate usage and their capability to identify and reformulate errors when no human support is available which can encourage the autonomy learning (Wang & Wang, 2015)

In conclusion, according to the results of this research, the AWE programs Grammarly and Grammarly enhanced significant aspects of the writing skill on the students' writing performance, motivating them to reformulate their errors by the sense of awareness and autonomy that the learners are conscious when using these tools without the support of the teacher's feedback. However, it is worthy to say that human guidance is required to compensate the limitations of AWE programs have, making it clear that these technological tools are classroom aids that complement the writers' improvement. In addition, it is important to understand that writing motivation and writing quality are characteristics that are linked with the objectives and learners needs. Although, AWE tools have demonstrated to have its limitations such as content development, these tools may have benefits for both teachers and students with appropriate teacher's monitoring and guidance.

It is suggested that teachers should provide students extra practice by giving them a bank of topics or letting students to select topics of their interest to be developed as additional assignments in order to provide learners more autonomy when writing. Additional teachers' feedback is required to clarify doubts that learners may have when using AWE tools as well as models of written texts should be provided to get more concrete ideas about how to compose a better content and organized written text. Considering the use of AWE tools in EFL and peer feedback for revision of aspects the learners may misunderstand might lead to future studies.

REFERENCES

- Bai, L., & Hu, G. (2017). In the face of fallible AWE feedback: How do students respond? *Educational Psychology*, 37(1), 67–81. <https://doi.org/10.1080/01443410.2016.122327>.
- Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated essay evaluation: The criterion online writing service. *AI Magazine*, 25(3), 27. <https://doi.org/10.1002/rcm.5057>.
- Chapelle, C. A., Cotos, E., & Lee, J. (2015). Validity arguments for diagnostic assessment using automated writing evaluation. *Language Testing*, 32(3). <https://doi.org/10.1177/0265532214565386>.
- Chapelle, C. A., Cotos, E., & Lee, J. (2015). Validity arguments for diagnostic assessment using automated writing evaluation. *Language Testing*, 32(3), 385–405. <https://doi.org/10.1177/0265532214565386>.
- Chen, C.-F., & Cheng, W.-Y. E. (2008). Beyond the design of Automated Writing Evaluation: Pedagogical Practices and Perceived Learning Effectiveness in EFL Writing Classes. *Language Learning & Technology*, 12(2), 94–112.

Chou, H. C., Moslehpour, M., & Yang, C.-Y. (2016). My Access and Writing Error Corrections of EFL College Pre-intermediate Students. *International Journal of Education*, 8(1), 144. <https://doi.org/10.5296/ije.v8i1.9209>.

Dikli, S., & Bleyle, S. (2014). Automated Essay Scoring feedback for second language writers: How does it compare to instructor feedback? *Assessing Writing*, 22, 1–17. <https://doi.org/10.1016/j.asw.2014.03.006>.

Dörnyei, Z., & Griffee, D. T. (2010). Research Methods in Applied Linguistics. *TESOL Journal*. <https://doi.org/10.5054/tj.2010.215611>.

Feng, H.-H., Saricaoglu, A., & Chukharev-Hudilainen, E. (2015). Automated error detection for developing grammar proficiency of ESL learners. *CALICO Journal*, 0(0). <https://doi.org/10.1558/cj.v33i1.26507>.

Hyland, K., & Hyland, F. (2006). *Feedback in second language writing: Contexts and issues*. Cambridge University.

Khoii, R., & Doroudian, A. (2013). Automated Scoring of EFL Learners' Written Performance: a Torture or a Blessing? In *Conference proceedings. ICT for language learning* (p. 367).

Khoii, R., & Doroudian, A. (2014). Automated scoring of efl learners ' written performance : a torture or a blessing? *Conference Proceedings. ICT for Language Learning*, (July), 5146–5155.

Lavolette, E. (2015). The accuracy of computer- assisted feedback and students. Retrieved from <https://www.mendeley.com/catalogue/accuracy-computer-assisted-feedback-students-responses-it/>.

Liao, H.-C. (2016). Enhancing the grammatical accuracy of EFL writing by using an AWE-assisted process approach. *System*, 62, 77–92. <https://doi.org/10.1016/j.system.2016.02.007>.

Liao, H. C. (2016). Using automated writing evaluation to reduce grammar errors in writing. *ELT Journal*, 70(3), 308–319. <https://doi.org/10.1093/elt/ccv058>.

Liu, S., & Kunnan, A. J. (2015). Investigating the Application of Automated Writing Evaluation to Chinese Undergraduate English Majors: A Case Study of WriteToLearn. *CALICO Journal*, 0(0). <https://doi.org/10.1558/cj.v33i1.26380>.

Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design*. Routledge.

Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology*, 37(1). <https://doi.org/10.1080/01443410.2015.1136407>.

Shannon, P. (2016). Films, governmentality and agency in the struggle over reading education. In *Exploring Teachers in Fiction and Film: Saviors, Scapegoats and*

Schoolmarms (pp. 183–191). <https://doi.org/10.4324/9781315671949>.

Shermis, M., & Burstein, J. (2003). *Automated essay scoring: A cross-disciplinary perspective*. Routledge.

Shim, Y. (2013). The effects of online writing evaluation program. *Teaching English with Technology*, 13(3), 18–34.

Wang, P. (2013). Can Automated Writing Evaluation Programs Help Students Improve Their English Writing? *International Journal of Applied Linguistics & English Literature*, 2(1), 6–12. <https://doi.org/10.7575/ijalel.v.2n.1p.6>.

Wang, P., & Wang, P. (2015). Effects of an Automated Writing Evaluation Program : Student Experiences and Perceptions, 12(1), 79–100.

Wang, Y.-J., Shang, H.-F., & Briody, P. (2013). Exploring the impact of using automated writing evaluation in English as a foreign language university students' writing. *Computer Assisted Language Learning*, 26(3), 234–257. <https://doi.org/10.1080/09588221.2012.655300>.

Warschauer, M. (2000). *Network-based language teaching: Concepts and practice*. Cambridge University.

Warschauer, M., & Grimes, D. (2008). Automated Writing Assessment in the classroom. *Pedagogies: An International Journal*, 3(1), 22–36. <https://doi.org/10.1080/15544800701771580>.

Wilson, J., & Czik, A. (2016). Automated essay evaluation software in English Language Arts classrooms: Effects on teacher feedback, student motivation, and writing quality. *Computers and Education*, 100, 94–109. <https://doi.org/10.1016/j.compedu.2016.05.004>.

Zhang, Z. (Victor). (2016). Student engagement with computer-generated feedback: a case study. *ELT Journal*, ccw089. <https://doi.org/10.1093/elt/ccw089>.

Zhang, Z. (Victor), & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, 36, 90–102. <https://doi.org/10.1016/j.asw.2018.02.004>.

The person who I admire is my youngest sister. Her name is Noemí and I admire her for some reasons. First, she is a very responsible girl. I think it is because she used to have violin classes in Ambato during the mornings, and then she went to the middle school in Latacunga during the evenings. Another reason is because she is not a procrastinator. She always does her homework the same day that her teacher send her, and she always be on time. Also, she is a very calm girl. She does not like to go to the parties but she loves dance. Finally, I admire her because she can do somethings. For example, she knows to draw very well. She is a good dancer. She loves to imitate some coreographies of K-pop groups. She knows play the violin and she usually plays the violin in the ninghts after dinner. In fact, for all these reasons I admire my youngest sister.

responsible → responsible

classes → classes

middle-school → middle school

because → that

send → sends

be → is

loves to

Possibly confused word: somethings

coreographies → choreographies

knows to

ninghts → nights

Appendix A. Students' writing samples and the reports obtained from the AWE tools Grammar and Grammarly.

The person who I admire (student' text)

The person who I admire is my youngest sister. Her name is Noemí and I admire her for some reasons. First, she is a very responsible girl. I think it is because she used to have violin classes in Ambato during the mornings, and then she went to the middle school in Latacunga during the evenings. Another reason is because she is not a procrastinator. She always does her homework the same day that her teacher send her, and she always be on time. Also, she is a very calm girl. She does not like to go to the parties but she loves dance. Finally, I admire her because she can do somethings. For example, she knows to draw very well. She is a good dancer. She loves to imitate some coreographies of K-pop groups. She knows play the violin and she usually plays the violin in the ninghts after dinner. In fact, for all these reasons I admire my youngest sister.

Report from Grammarly

The best place to live

One of the best places for live is Korea for many reasons. First, it is a beautiful place. Korea has a lot of beautiful cities for visit. It has big malls; restaurants and places for

buy anything you want. For example, Myeongdong is one of the biggest street markets in Korea. Second, Korea has a lot of restaurant. Korean people love to eat so that's why

Wordiness Errors

7 instances (53% of sentences)
Goal: 2% or fewer

Your writing seems too wordy. Why use 3 words when you can say it with 1? Examples:

- it **seems to be** effective --> it **seems** effective
- the amount of energy **being** used --> the amount of energy used
- We make **a majority of** our decisions **in a brief moment** --> We make **most** of our decisions **quickly**.

You don't need to remove ALL instances below, but try to reduce your use until you are in the green.

One of the best places for live is Korea for many reasons. First, it is a beautiful place. Korea has a lot of beautiful cities to visit. It has big malls, restaurants and places for buy anything you want. For example, Myeongdong is one of the biggest street markets in Korea. Second, Korea has a lot of restaurant. Korean people love to eat so that's why they had a lot of different kind of food. One of the most popular food there is the Ramen. Finally, the entertainment there is very different. Korea is the capital of kpop, so most of the Young Korean wants to be a kpop start for that influence. And also like a consequence of kpop, many people from Latin America, USA, Europe and others, start to move to Korea for a good live there. For all that's reasons we can say that Korea is a good and interesting place for live.

Wordiness Errors found

- a lot of
- interesting
- kind of
- one of the most
- there is

Re-highlight

they had a lot of different kind of food. One of the most popular food there is the Ramen. Finally, the entertainment there is very different. Korea is the capital of kpop, so most of the Young Korean wants to be a kpop start for that influence. And also like a consequence of kpop, many people from Latin America, USA, Europe and others, start to move to Korea for a good live there. For all that's reasons we can say that Korea is a good and interesting place for live.

Report from Grammar