

Online Open-Source Writing Aid as a Pedagogical Tool

Beata Lewis Sevcikova¹

¹ Applied Linguistics Department, College of Humanities, Prince Sultan University, Riyadh, Saudi Arabia

Correspondence: Beata Lewis Sevcikova, Applied Linguistics Department, College of Humanities, Prince Sultan University, Riyadh, Saudi Arabia. Tel: 966-569-285-139. E-mail: blewis@psu.edu.sa

Received: June 26, 2018 Accepted: July 23, 2018 Online Published: July 25, 2018

doi: 10.5539/elt.v11n8p126 URL: <http://doi.org/10.5539/elt.v11n8p126>

Abstract

The present research offers an assessment of the online open source tools used in the L2 academic writing, teaching, and learning environment. As fairly little research has been conducted on how to best use online automated proofreaders for educational purposes, the objective of this study is to examine the potential of such online tools. Unlike most studies focusing on Automated Writing Evaluation (AWE), this research concentrates only on the online, open-source writing aide, grammar, spelling and writing style improvement tools available either for free or as paid versions. The accessibility and ability to check language mistakes in academic writings such as college-level essays in real time motivates both, teachers and students. The findings of this empirical-based study indicate that despite some bias, computerized feedback facilitates language learning, assists in improving the quality of writing, and increases student confidence and motivation. The current study can help with the understanding of students' needs in writing, as well as in their perception of automated feedback.

Keywords: assessment, Automated Writing Evaluation (AWE), feedback, motivation

1. Introduction

Writing in English as a second (or foreign) language (L2) has a significant place in academia. From this perspective, the quality of teaching writing skills, as well as evaluating students' drafts, is demanding. Moreover, research writing for graduate students is the foundation of their academic success and future career. Learning how to write and then, maybe, publish in reputable journals make them competitive in the job market (Cotos, 2014).

Essentially, assessment of writing assignments aims at supporting and improving student learning. Assessment, as a term, stems from a movement towards accountability. According to Peha (2011), assessment is collecting information with the purpose of guiding future instruction and it is closely related to evaluation. He defines evaluation as any decision that is made based on the information which has already been gathered through assessment.

The students' aim is to become better writers or at least to perform sufficiently well in their classes to accomplish the level of success they desire. They want to have their performance in writing evaluated consistently. As it was mentioned in Blikstad-Balas, Roe and Klette (2018) "...student development as writers requires a supportive environment in which they receive sustained opportunities to write" (p. 119). Based on Beatty and Gerace study (2009, p. 152), within the conceptual changes in pedagogy, the main role of classroom discourse is to direct students' thinking and provide the environment to enhance thinking, cognitive skills and entire learning. Currently, two main scoring systems in writing are used for academic purposes:

- 1) The human rating system with two basic approaches, holistic and analytic.
- 2) Automated Writing Evaluation, AWE, based on Chen and Cheng research (2008), software is designed to provide instant computer-generated scores for a submitted essay along with diagnostic feedback.

Various studies report a high degree of agreement between human raters and AWE scoring systems: "Most of the research to date has been funded or carried out by developers of AWE software so as to demonstrate that it 'works' either in the sense that its scores correlate with those of human raters or that particular features or aspects function as desired" (Warschauer & Ware, 2006, p. 10). They also illustrate that the scores provided by AWE systems and human raters are well correlated (Dikli, 2006; Keith, 2003; Phillips, 2007).

1.1 AWE versus Human Rating

AWE (or AEE – automated essay evaluation) is defined as “the process of evaluating and scoring written prose via computer programs” (Shermis & Burstein, 2013, p. 1). It has witnessed a growing interest in the field of L2 writing (Chen & Cheng, 2008; Warschauer & Ware, 2006) over the past years. Up to now, the hands-on benefits of automated feedback for teachers, such as timesaving and pedagogical effectiveness have been researched in specific instructional contexts. Automated feedback is the most promising point of contact between the areas of AWE and L2, which some studies recommend applying to complement teacher’s comments (Shermis & Burstein, 2013; Warschauer, 2000).

The latest research on automated essay scoring conducted by Blood (2011) focuses on operationalizing and scoring writing assignments, as well as defining what is ‘excellent writing’. Training, as well as marking, is a time-consuming process. Online tools shorten the period needed to obtain scores (Livingston, 2009).

In addition, maintaining the reliability and construct validity of scores, is challenging, as instructors or testers differ in their comprehension of candidate performances and their tendencies towards severity and leniency (Blood, 2011). Although machines do not understand and evaluate writings using the same cognitive abilities as humans, the concerns about AWE might be allayed as various systems have seemed to achieve similar scoring prediction outcomes (Shermis & Hamner, 2012).

Previous research has documented that employing well-trained human raters is not always feasible. Significantly, biased ratings usually appeared in cases where candidates showed very high or low ability of writing. In these cases, the raters were not consistent in their application of rating criteria. Kondo-Brown (2002) states that the rater training may improve inter-rater reliability and the consistency of ratings, although it may not have much influence on other unwanted rater characteristics, such as increased or decreased severity.

Likewise, Shi (2001) concludes although raters do not differ significantly in the scores they gave, they differed considerably in the justifications they provided for their ratings. Various raters weighted certain essay categories more heavily than others (Weigle, 1999). Native speakers are inclined to focus more on content and language, while non-native speakers focused on organization and length. Similarly, the research conducted by Wolfe, Song and Jiao, (2016, p. 1) foresee the difficulty in the accuracy of scoring an essay by experienced and inexperienced raters.

1.2 Benefits and Challenges of AWE

Even though the question of whether automatic essay scoring tools apply scoring rules accurately and consistently, it has not been a matter of great professional discussion because internal reliability is an inherent trait of such systems. Many critics of automatic essay scoring have not succeeded in contextualizing their criticism within “an interpretive argument for the validity of a specific testing context, thus relegating their objections to the realm of the abstract” (Blood, 2011, p. 53). For example, Ericsson and Haswell’s (2006) anthology doubts the implementation of AWE in writing classes. They question computers’ ability to understand texts and consequently to evaluate the quality of writing. On the contrary, it seems, intelligent feedback features such as individualized student approach, showing the error type, explicitly explaining the error, leads to student self-correction. For these reasons, they perceive AWE as an objective solution.

Although some studies strongly reject machine-scoring software, some believe that use of a variety of software packages is beneficial for the 21st century academic settings because they generate student knowledge in EFL courses.

1.3 AWE Potentials for the Learning Process

One reason to explore the use of AWE in and out of the classroom is that it has the potential to provide more feedback, independent revision and learning opportunities for students. This type of feedback can often lead to better overall quality of writing. It also provides a space in which writers can learn more about evaluation criteria and improve their skills as writers and revisers (MacArthur, 2007; 2012).

Another reason to study AWE is that this technology is poised to take on a more significant role in academic writing education. In addition, due to errors, writing feedback from human raters is often inconsistent. Educators, even if writing is peer-reviewed, may not always recognize the same problems on a student paper each time they evaluate it. This inconsistency can bewilder students. In contrast, AWE feedback provides consistent feedback on the same types of errors for every draft a student submits (Moore & MacArthur, 2016).

Existing research focuses on technical qualities of various AWE tools. It also focuses on the students’ writing performance from a faculty perception of student writing proficiency. Nevertheless, it has not fully explored the

possible impact on student learning and perception (Burstein, Elliot, & Molloy, 2016).

In this research, I analyze Saudi Arabia female student responses to human and AWE rating systems. Regarding the situation in Saudi Arabia, Alkhatib (2015) points out that instructors use rubrics to assess L2 writing but female students at Saudi Arabia universities are rarely satisfied with their Written Corrective Feedback (WCF), and some of them take teacher comments ‘personally’ (Alkhatib, 2015). According to Salebi (2004), female students usually provide comments on their writing errors, claiming that they had made mistakes because of test anxiety, concentration on content rather than form and the limited time allotted to the test. This type of anxiety can cause dissatisfaction and demotivation among students learning English as their L2. Students agree with instructors that written responses may affect student writing and attitude. Although written comments are time-consuming; nevertheless, teachers use them because they want to help to improve the writer’s performance (Leki, 1990).

1.4 Research Focus

More exploration and discussion are required to examine the potential support of online AWE tools in teaching and learning English as a foreign language. The AWE should be used to achieve more desirable learning outcomes while avoiding potential harm from the technology. The present study, therefore, seeks for the answers on two central questions that guided the research:

- a) Can an online open-source writing aide tool be successfully implemented in writing classes to promote learning?
- b) What are the most important factors an open-source writing aide offers which influence student perception, motivation and learning?

2. Method

This classroom-based study investigated the impact of the development and use of AWE on writing performance across two groups of English as a foreign language (EFL) students in academic settings. The study followed the process-product research approach (Warschauer & Ware, 2006), employing a mixed-method design. Through both, quantitative and qualitative research methodologies, data were generated to back the evidence from the researched tools.

The qualitative part incorporated the study of students’ perception of the AWE feedback on essays in a natural setting (classrooms) and its interpretation and exploration of the experience of participants. The research sought to explain ‘if’ and ‘why’ an AWE promotes student motivation and learning in this environment. As a qualitative research method, (a) open-ended survey questions (see Appendix A where question 5 is the only open-ended one, question 10 and 15 required clarification and justification of students’ opinion and answers), (b) participant observation, and (c) informal interviews were used. The aim was to discover how they reacted to the human marker’s and AWE tools feedback. During the interviews at the end of both semesters 161 and 162, 61 students from four classes answered the informal interview questions. The aim was to discover how they reacted to the human rater’s and AWE tools feedback. During this interview, students answered the following questions:

- a) Could you use three adjectives describing your experience with AWE?
- b) Could you describe your feelings about the teacher’s holistic feedback and AWE feedback?
- c) Do you think AWE helped you to learn and improve writing?
- d) What did you like most/least about the fact we used AWE for reviewing process?
- e) Would you recommend continued use? Why? Why not?

Quantitative data gathered from the 15-question survey and the student’s samples marks are expressed in a numerical form which was analyzed and explained in figures (see Appendix B, Figures 1 and 2).

2.1 The Educational Context of the Study

The research was conducted in four student groups at Prince Sultan University (PSU), Saudi Arabia, academic year 2016/2017 (semesters 161 and 162). This University is one of the leading private university in the Middle East. The aim is to provide a quality education to the highest international standards.

Overall, there were 43 students in the senior group (group A) and 18 students in the sophomore group (group B). The language of instruction is English. Before starting studies, students must either finish the Preparatory Year Program (which has got three levels, from beginners to pre-intermediate) or come to the university with 5.5 IELTS band.

Female students from the College of Humanities, the TEFL and the Translation Programs were part of the research. There are various writing courses throughout the BA (Bachelor of Arts) studies. All writing courses are mandatory for both Programs. For my research, the intermediate writing course students (group B) and the advanced writing course students (group A) took part. The course objectives of both groups were to learn how to argue and reason in a piece of research. While the sophomore group (B) ‘touches’ the basics of research, the senior students’ semester assignment (A) is to conduct a comprehensive mini research or case study.

All four writing classes which took part in the research were second and fourth-year English major students. Both groups shared some standard features in their writing instruction. Their course objectives aimed to develop students’ academic writing skills in research, they covered similar content but in different depths. A similar writing approach, including model essay reading activities followed by pre-writing, drafting and revising activities was applied.

2.2 Data Collection

Both researched groups had to complete and submit two drafts (the first and the final) of an assignment (either argumentative essay – group ‘B’ or research – group ‘A’). Essay samples one and two used for this study were from sophomore students; essay three and four from senior students which served as examples of their performance. The criterion for choosing essays for this part of the research was students’ performance – one essay below expectations (poor), one excellent and two average essays (meets expectations) were chosen. The other students’ essays served as control essays; they were evaluated by both, human raters and AWE tools.

Four sample essays (two from semester 161 and two from semester 162) were chosen for the research and ten human markers, who were chosen randomly from two different institutions, evaluated them. Six markers were native speakers and four non-native speakers. Subsequently, the essays were assessed by online AWE tools. The markers, four males and six females marked the essays per holistic rubric (see Appendix C), and their marking was compared against the online grammar and spelling checkers, (see Appendix D).

2.3 Participants Profile

Most students at university, aged 18 to 21, are Saudi nationals, 10-15% are non-Saudis. Students come either from local or international schools; while a few studied abroad.

2.4 Reasons for Choosing Online AWE Tools

- a) Both are accessible to the public and straightforward to use and understand.
- b) They provide holistic and analytic scores along with feedback on style, content, language use and mechanics.
- c) Both tools offer a wide range of writing correction ideas.
- d) AWE tools can be used as formative learning tools because they allow students to revise and edit their drafts. The analytic assessment and feedback were given to each essay draft submitted to the AWE tools.

2.5 Description of AWE Tools

1) Paid AWE tool ‘A’ tested in this study combines several characteristics in its feedback. It is immediate (less than 60 seconds from the time of submission) and intelligent (generated automatically by a natural language processing-based engine) online marking tool. It offers individual, student-based submission feedback in the respective discipline. This tool goes beyond fixing misspelled words, it even identifies words, albeit spelled correctly, in the wrong context so a user will not overlook any aspect of writing that can diminish its quality. The tool’s operation technology is both technical and contextual – it helps correct grammar, punctuation, and spelling mistakes. At the same time, it also improves vocabulary by pinpointing and fixing contextual errors. It also recommends enhancements of writing style.

2) The free AWE tool ‘B’ tested in this study is a marking tool to check students’ writing for language and style problems before the submission. The best output of this tool can be found in academic writing, including reports and essays. As designers of the program claim, it does not contain sophisticated online techniques, such as AWE software, it allows the users to have formative feedback and instant evaluative regarding their submissions.

2.6 Designing a Classroom Performance Task

It is not always easy for the instructor to design a classroom performance task. With this caveat in mind, I took into consideration:

- 1) The course objectives, skills and initial knowledge which I wanted students to gain as a result of completing a task.
- 2) A task specification (see Appendix E) –it was related to the students' Major (in this case Applied Linguistics and Translation).
- 3) Explicit performance criteria measuring the extent to which students have mastered the skills and knowledge were prepared. The holistic approach focuses on overall writers' performance; it does not identify writers' incompetence and errors. Ten human markers, during the Standardization session, earned the basic training in understanding the holistic rubric used for this research. The understanding of the holistic rubric was unified (see Appendix C) and the decision on marking style was taken.

3. Results

The study intended to observe the potential support of both automated proofreaders for teaching and learning academic writing purposes. The findings indicated that computerized feedback facilitated language learning, assisted in quality writing, and increased the confidence of students. The results suggested that both AWE tools tested in the research develop feedback that encouraged students' revision through the writing process, supported their feelings that the feedback was objective and fair, acted as an intermediary between an instructor and a student and made teaching and learning student-centered. As group 'A' students learned from the beginning how to use AWE tools to improve their writing, their submissions and re-submissions did not contain as many mechanical errors, especially in orthography (spelling and capitalization) and punctuation, grammar and usage mistakes (such as rendering word-for-word into English) as the other group's submissions. Group 'B' students submitted their first draft without using AWE tools. After the instructor's WCF, the AWE tools were introduced, and they acquired the chance to re-submit the first draft after using the AWE, and then, they received the written and AWE feedback again.

3.1 Data Analysis and Findings

In Tables A1 and A2 (see below) I summarized scores given by markers and AWE tools and provided the comparison of grades expressed in percentage. The study showed that apart from essay #2, grades displayed a wide discrepancy between the marks awarded by human raters and AWE tool 'B.' Referring to Tables A1a and b and A2, it can be seen that only Essay one (a sophomore student writing) was marked similarly, by both AWE tool 'A' and human markers. It was easier to compare AWE tool 'A' and human markers as the program marked similarly to them, it gave overall percentages for the evaluated essay and holistic comments and focused on genre-specific writing style (see Appendix F and Table A2). Tool 'B' uses a scale of 12 points and more analytical evaluation of areas such as paragraph, sentence, vocabulary, academic style, grammar and punctuation and general writing style rating complemented with smiley faces (see Appendix G). Its analytical style of evaluating writing detects and highlights issues related to academic styles such as vagueness of language, informal phrases, over-used expressions, comma splices, sentence fragments and many more, so the agreement between this program and the human raters' grades was very low. However, if attention is paid on types of mistakes in essay three, despite the differences in scoring (AWE 'A' 80%, AWE 'B' 59%) both tools detected and marked similar mistakes in the academic style. When analyzing closer essay 1 mistakes, Tool 'A' seemed to 'understand' that the essay was an argumentative one focusing on personal issues, not academic ones. Tool 'B' marked all personal pronouns as mistakes of an academic style and for that reason, it gave much lower grades (see Appendix B).

Table A1. Summary of evaluations by human raters (T – teachers)

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
ESSAY 1	D	F	D	C	C	D ⁺	D	D	F	D
ESSAY 2	B	C	A	C	C	D	B	D	F	B
ESSAY 3	C	C	A	A	A ⁺	C	B	C	B	C
ESSAY 4	A	B	B	B	C	C	B	C	B	B
ESSAY 4	A	B	A	A	B	A	A	A	A	B

Table A2. Summary of evaluations by both, human raters and AWE tools

	Marks from Teachers					MARKS FROM A WE TOOL "A"	MARKS FROM A WE TOOL "B"
	A	B	C	D	F		
ESSAY 1	0	0	2	6	2	64%(D)	4 out of 12 (points / 33%-F)
ESSAY 2	5	4	1	0	0	82%(C)	5 out of 12 (points / 41%-F)
ESSAY 3	0	4	6	0	0	80%(C)	8 out of 12 (points / 59%-F)
ESSAY 4	7	3	0	0	0	97%(A)	8 out of 12 (points / 59%-F)

3.2 Survey and Interview Analysis

At the end of both semesters, an informal, in-class interview was conducted with all 61 students from both groups. During the informal interview, students used more positive than negative adjectives for question (a). From the positive ones, they mentioned that using AWE tools was "exciting, new, interesting, stimulating, educational, encouraged thinking and independent work, objective and both tools helped them with the reviewing process." Negative adjectives such as "challenging, scary at the beginning, more time consuming because it was easier to ask an instructor and get the answer, were mentioned."

In their comments to question (b), they echoed their remarks from the survey – they understood that the combination of a teacher and AWE feedback enhanced feelings of objectivity and fairness. There were a few different responses (4 out of 18 students) from the sophomore students. They thought that it was too difficult to use AWE tools and preferred teacher evaluation. The process of reviewing was time-consuming for them.

Most of the students (57 out of 61) agreed that AWE promoted learning and helped in improving writing skills ("AWE explanation helped me to understand my mistakes"). Some students mentioned they did not like reviewing their essays. They even admitted ignoring the AWE review of their first draft ("I did not review my writing"), but once they saw that those students who utilized AWE got better feedback from the first draft, they made use of the AWE tool for their final submission. Most observed students (53 out of 61) did not want to purchase the paid tool 'A,' and so they used tool 'B.' They found it straightforward to use and understand, liked its 'Advice' feature and the 'Report card.' The most comprehensive and appreciated feature was the instant advice for improvement.

Those who used tool 'A' were keen on its discipline-specific feature as well as plagiarism checker. Some sophomore students reported problems with understanding the AWE tools explanation of the problem.

Almost all students (59 out of 61) suggested keeping these or similar tools because they feel it was motivational and easily mastered. In other words, both tools assisted them in improving vocabulary, using of prepositions, raising awareness of their weaknesses in writing. Some students also recommended implementing more training on how to understand AWE tool comments.

In a 15-question survey, both groups mostly agreed that only teachers should mark most of the assignments (question 1) and that writing is a useful skill (question 8). In question 4, sophomore students preferred direct correction by a teacher (55.55%), while many senior students (69.76%) thought a combination with AWE tool increased assessment motivation and objectivity (see Appendix A, Figures 1 and 2). Only 27.77% of sophomore students stated that AWE together with teacher's feedback was preferable. Another difference between groups was noted when students commented on rubrics (question 10). While senior students assumed, it was essential to get a rubric just before the assignment (41.86%), sophomore students believed it was more important at the beginning of the course (50%).

Questions 12 and 13, where I asked how students felt about the feedback given by an instructor and the AWE tool, both groups agreed that they felt the feedback, in this case, was more objective. Even though more than 72 % of senior students and 66.66 % of sophomore students felt that teacher feedback is objective, they stated, that a mix of teacher and AWE feedback, is not only motivational but more objective (79% of senior students and 72.22 % of sophomore students were in favor of this combination).

Questions 14 and 15 which echoed research questions and answers from students, supported the hypothesis that both tools were motivational and confidence building which supported their learning. Overall, 81.39% of senior students and 66.66% of sophomore students agreed that AWE tools were helpful and 7.41% thought using tools was motivational because: "You were the first instructor teaching us how to use AWE as well as how to correct our mistakes this way. It was interesting, and it motivated me to think how to do it" (question 15).

Question 5 asked students for justification of question four. Mostly, students did not prioritize peer feedback but expressed their belief that an AWE tool was a positive addition to an instructor's feedback, as it increased objectivity, supported learning and provided a fair evaluation.

4. Discussion

A few studies are investigating and analyzing AWE tools designed especially for teaching purposes. Despite the accessibility, online AWE tools have not been analyzed so far and therefore this study can start a discussion as to whether these accessible tools can or cannot be used in the classrooms. Overall, AWE tools under investigation were perceived positively by my students, and this despite their limitations in recognizing genre, style and context (especially in the unpaid version). However, some students echoed Ericsson and Haswell's (2006) doubts and they questioned a computer's ability to understand the content of the writing and accuracy of correcting mistakes.

This research showed that pedagogical practices with online AWE tools influence student perceptions in facilitating their learning of writing. The AWE implementation was accepted quite favorably when the program was used to help students with the revision of their first drafts.

As mentioned previously, AWE offers instant feedback which may vary in the degree of clarity and specificity. Such feedback leads students' attention to their linguistic inadequacies and motivates them to produce more comprehensible output. Based on Beatty and Gerace study (2009), within the conceptual changes in pedagogy, the main role of classroom discourse is to direct students' thinking and provide the environment to enhance thinking, cognitive skills and entire learning.

The purpose of this study, based on the research questions, is to empirically explore whether automated feedback can trigger learning processes and foster student/computer interaction that potentially leads to the revision of output and improvement in students' written performance. Based on the informal interviews conducted in both researched groups, most students agree the immediate AWE tool enhances their self-correction techniques. Although instructors can use AWE to operationalize the constructs of input and interaction, output has not yet been addressed by research. Considering this gap, the paper investigates the learning potential of automated feedback generated by two online grammar and spelling checkers.

Most students perceive AWE as an objective solution. Although some limitations of the programs' methods of assessment and writing support features could have negatively affected student attitude towards the usefulness of AWE online tools, both classes appreciate this new experience. The students' remarks on the suitability of the implementation of the both online tools centers around: (a) prioritizing either human or AWE online tools feedback, (b) understanding of AWE corrective feedback and scores, (c) enhancing motivation to learn, (d) addressing new, challenging but modern ways of learning writing via AWE tools.

As described in my literature synthesis, I use AWE tools for pedagogical purposes to enhance students' self-belief in writing, attitudes, and motivation and learner autonomy (Smith, 2008). To conclude, the use of online AWE tools is not a simple "yes" or "no" issue. This matter involves a complex combination of factors regarding software design, pedagogical practices, learning contexts and student willingness to accept assistance from a non-human marker. Despite any shortcomings AWE tools may have, the presented study supports using these tools for teaching and learning purposes as not all institutions can implement professional writing applications into classes. Using easily accessible tools, based on this study, promotes student learning and independence.

The development of AWE tools is intensifying, and it is believed that very soon, they will be able to achieve the accuracy, which will be indistinguishable from human examination. Also, the research strives to prove that students are more motivated to revise because of the 'technology' feedback. Moreover, the receipt of prompt results from AWE tools motivates students to revise extensively prior to submitting their writing for scoring. Hopefully, this paper will have the effect of encouraging decision-makers and instructors to explore and possibly implement the use of such tools in order to enhance student autonomy and overall writing improvement.

Acknowledgments

The researcher thanks Prince Sultan University and the research group (Language Learning and Teaching Group RG-CH-2016/11/11).

References

Alkhatib, N. (2015). *Written corrective feedback at a Saudi University: English language teachers' beliefs, students' preferences, and teachers' practices*. (Thesis (PhD), University of Essex, Essex, UK. Retrieved

- from <http://repository.essex.ac.uk/15382/1/THESIS%206Oct..pdf>
- Beatty, I. D., & Gerace, W. J. (2009). Technology-enhanced formative assessment: A research-based pedagogy for teaching science with classroom response technology. *Journal of Science Education and Technology*, 18(2), 146-162. <https://doi.org/10.1007/s10956-008-9140-4>
- Blikstad-Balas, M., Roe, A., & Klette, K. (2018). Opportunities to Write: An Exploration of Student Writing During Language Arts Lessons in Norwegian Lower Secondary Classrooms. *Written Communication*, 35(2), 119-154. <https://doi.org/10.1177/0741088317751123>
- Blood, I. A. (2011). Automated Essay Scoring: A Literature Review. *Working Papers in TESOL & Applied Linguistics*, 11(2), 1-25.
- Burstein, J., Elliot, N., & Molloy, H. (2016). Informing automated writing evaluation using the lens of genre: Two studies. *Calico Journal*, 33(1), 117.
- Chen, C.-F. E., & Cheng, W.-Y. E. C. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning and Technology* 12(2), 94-112.
- Cotos, E. (2014). *Genre-based automated writing evaluation for L2 research writing: From design to evaluation and enhancement*: UK: Palgrave Macmillan. <https://doi.org/10.1057/9781137333377>
- Dikli, S. (2006). An overview of automated scoring of essays. *The Journal of Technology, Learning and Assessment*, 5(1), 1-35.
- Ericsson, P. F., & Haswell, R. H. (2006). *Machine scoring of student essays: Truth and consequences*: Utah State University Press.
- Keith, T. Z. (2003). Validity and automated essay scoring systems. In *Automated essay scoring: A cross-disciplinary perspective* (pp. 147-167): Lawrence Erlbaum Associates.
- Kondo-Brown, K. (2002). A FACETS analysis of rater bias in measuring Japanese second language writing performance. *Language Testing*, 19(1), 3-31. <https://doi.org/10.1191/0265532202lt218oa>
- Leki, I. (1990). *Coaching from the Margins: Issues in Written Response*. Cambridge, UK: Cambridge University Press.
- Livingston, S. A. (2009). Constructed-Response Test Questions: Why We Use Them; How We Score Them. R&D Connections. Number 11. *Educational Testing Service*, 1-8.
- MacArthur, C. A. (2007). Best practices in teaching evaluation and revision. In *Best practices in writing instruction* (pp. 141-162). New York, US: The Guilford Press.
- MacArthur, C. A. (2012). Evaluation and revision. In I. V. W. B. (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 461-483). London: Psychology Press.
- Moore, N. S., & MacArthur, C. A. (2016). Student Use of Automated Essay Evaluation Technology During Revision. *Journal of Writing Research*, 8(1), 149-175. <https://doi.org/10.17239/jowr-2016.08.01.05>
- Peha, S. (2011). Assessing writing, assessing writers. Retrieved on March 4, 2014. Retrieved from Teaching that makes sense website: [https://www.ttms.org/PDFs/09%20Writing%20Assessment%20v001%20\(Full\).pdf](https://www.ttms.org/PDFs/09%20Writing%20Assessment%20v001%20(Full).pdf)
- Salebi, M. Y. (2004). Saudi College Students' Perception of Their Errors in Written English. *Scientific Journal of King Faisal University (Humanities and Management Sciences)*, 5(2), 1425.
- Shermis, M. D., & Burstein, J. (2013). *Handbook of automated essay evaluation: Current applications and new directions*. New York, US: Routledge.
- Shermis, M. D., & Hamner, B. (2012). *Contrasting state-of-the-art automated scoring of essays: Analysis*. Paper presented at the annual meeting of the National Council of Measurement in Education, Vancouver, BC, Canada. Retrieved from <http://www.journalofwritingassessment.org/article.php?article=69>
- Shi, L. (2001). Native-and nonnative-speaking EFL teachers' evaluation of Chinese students' English writing. *Language Testing*, 18(3), 303-325. <https://doi.org/10.1191/026553201680188988>
- Smith, R. (2008). Learner autonomy. *ELT journal*, 62(4), 395-397. <https://doi.org/10.1093/elt/ccn038>
- Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Language teaching research*, 10(2), 157-180. <https://doi.org/10.1191/1362168806lr190oa>
- Warschauer, M. (2000). The changing global economy and the future of English teaching. *Tesol Quarterly*, 34(3),

511-535. <https://doi.org/10.2307/3587741>

Weigle, S. C. (1999). Investigating rater/prompt interactions in writing assessment: Quantitative and qualitative approaches. *Assessing Writing*, 6(2), 145-178. [https://doi.org/10.1016/S1075-2935\(00\)00010-6](https://doi.org/10.1016/S1075-2935(00)00010-6)

Wolfe, E. W., Song, T., & Jiao, H. (2016). Features of difficult-to-score essays. *Assessing Writing*, 27, 1-10. <https://doi.org/10.1016/j.asw.2015.06.002>

Appendix A

Survey

COURSE CODE _____

1. Should the teacher mark all your writing assignments?

- a) All
- b) Most of them
- c) Some
- d) None

2. Should the Automated Essay Scoring system be used to score all your writing assignments?

- a) All
- b) Most of them
- c) Some
- d) None

3. Should other students mark your work?

- a) All
- b) Most of them
- c) Some
- d) None

4. Would you prefer your writing assignments to be marked by:

- a) The teacher
- b) The Automated Essay Scoring system
- c) The other student (peer scoring)
- d) The combination of 'a' and 'b'
- e) The combination of 'a' and 'b' and 'c'

5. Justify your choice in the question 4. Give at least three reasons:

6. Do you expect to help other students improve their writing skills?

- a) Yes
- b) Sometimes
- c) No

7. Do you expect to work on a group writing task?
- Yes
 - Sometimes
 - No
 - I prefer individual writing
8. Will good writing skills help you in life?
- Yes
 - Sometimes
 - No
 - I do not know
9. How do you prefer submitting writing?
- Written by hand
 - As a soft copy
 - As a computer-based writing (written on a computer and printed as a hard copy)
10. I think marking guidelines and rubrics as means of student assessment performance should be given (handed out):
- At the beginning of the course
 - During the course
 - Just before the writing assignment
 - During the assignment
 - Shouldn't be given at all

Explain your choice:

11. As a student, I think the use of marking guidelines and rubrics as means of student performance assessment of is:
- Necessary and important
 - Necessary but not important
 - I don't know
 - Not necessary, nor important
12. When I get feedback on my writing from a teacher, I feel:
- The feedback is usually objective
 - The feedback is sometimes objective
 - The feedback is sometimes subjective (just a teacher's opinion)
 - The feedback is mostly subjective (just a teacher's opinion)
13. When I get feedback on my writing only from the Automated Essay Scoring system, I feel:
- Motivated and the feedback is usually objective

- b) The feedback is sometimes objective
- c) The feedback is sometimes subjective
- d) The feedback is mostly subjective
- e) I feel unmotivated because of mistakes I made

14. When I get feedback on my writing from a teacher and the Automated Essay Scoring system, I feel:

- a) The feedback is usually objective
- b) The feedback is sometimes objective
- c) Unmotivated and the feedback is subjective
- d) The feedback is mostly subjective
- e) Motivated and the feedback is mostly objective

15. Which do you think supports learning the most?

- a) Direct corrections (showing all your mistakes and possibly correcting them by a teacher)
- b) Peer corrections and group corrections
- c) Indirect feedback (based on Rubric and I provided feedback both on form and content).
- d) Just Automated Essay Scoring system correction
- e) Combination of Automated Essay Scoring system and teacher's feedback.

Explain your choice:

Appendix B

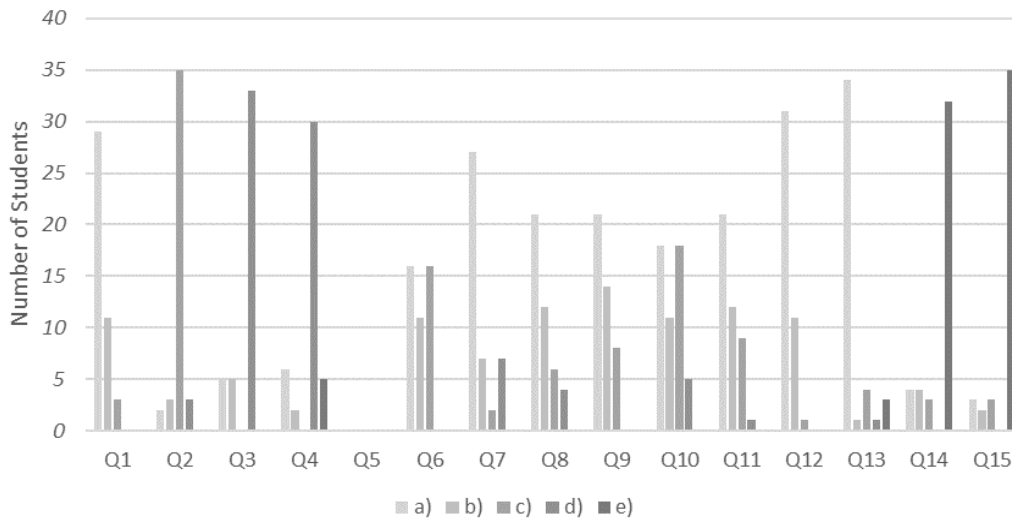
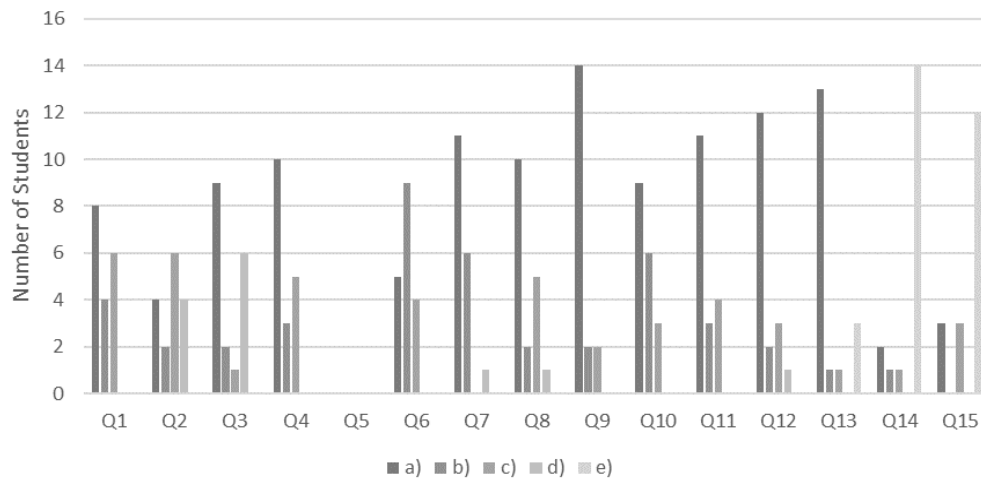


Figure 1. Group A – Senior student survey results



Group B – Sophomore Student Survey Results

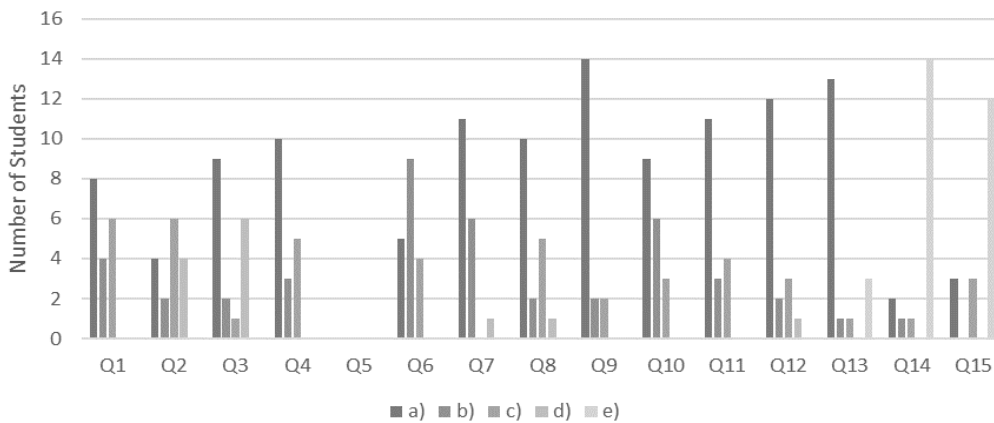


Figure 2. Group B – Sophomore student survey results

Appendix C

Holistic Rubric

“A” work: the argumentative essay is exceptional in every way. The essay is well organized, and all claims are supported. Content is detailed and lacking in no area. There are almost no errors in tone, format, mechanics, grammar, and content. (90-100%)

“B” work: the argumentative essay is sufficient in most operations. The essay is focused and clear, containing necessary details, but some areas require further development. While it may contain a few errors, they do not detract from the overall point. (80-89%)

“C” work: the argumentative essay is adequate but not exceptional in tone, format, mechanics, grammar, and content. The thesis statement is explicit although control and command lack organization. The paragraphs are undeveloped even though they provide support for the main ideas. There may be various errors in mechanics, grammar, content, and tone. These errors mostly do not detract from the overall writing. (70-79%)

“D” work: there are numerous errors in all areas of the argumentative essay. It is not organized or focused. The thesis is unclear, and the essay is not controlled. Ideas are underdeveloped, and there are numerous errors in tone, format, mechanics, grammar. The errors distract the message of the entire essay. Despite all of the errors, there appears to be a legitimate effort put forth by the writer. (60-69%)

“F” work: the argumentative essay has got significant problems throughout. There is no thesis, and if there is, it lacks argument and organization. The essay shows no understanding of basic essay structure, and there are significant errors in tone, format, mechanics, grammar, and content. The effort on the part of the writer is questionable, at best. (below 60%)

Appendix D

Essay 1 – Sample of Tool ‘A’ Scores

DOCUMENT	SCORE
	64 of 100
	ISSUES FOUND IN THIS TEXT
	49
	PLAGIARISM
	2%
Contextual Spelling	3
Confused Words	2
Misspelled Words	1
Grammar	13
Wrong or Missing Prepositions	4
Faulty Subject-Verb Agreement	2
Determiner Use (a/an/the/this, etc.)	2
Incorrect Phrasing	2
Conjunction Use	1
Pronoun Use	1
Incorrect Verb Forms	1
Punctuation	7
Punctuation in Compound/Complex Sentences	6
Misuse of Semicolons, Quotation Marks, etc.	1
Sentence Structure	1
Faulty Parallelism	1
Style	12
Improper Formatting	5
Passive Voice Misuse	3
Wordy Sentences	3
Politically Incorrect or Offensive Language	1
Vocabulary enhancement	13
Word Choice	13

First of all, I'd like to say that without my parents, I'd be nothing. My lovely parents always support and provide me with all desires and needs. They always help me in my life especially when I am a college student. Thanks to my parents for giving me strength and ability to understand, learn and complete this study. Really, I am very thoughtful and I ask Allah to bless them. Although I decided what to study and im loving it but I'm pretty sure I would never survive without having them standing beside me.

Generally, my parents are my first teachers. They are the most important and essential for me. I hope to be like them in all their characteristics. They have great and strong characters that have a great effect on me. They always help me in growing up and to be a healthy and useful one for my country and humanity. I'm very proud that I have parents like my parents. They do best efforts to be a good individual in my country. The parents bear a great responsibility to be a great symbol for me to show me what is right and what is wrong moreover to show me the good behavior and to keep bad behavior away from me. One can say that it is their parents' duty to guide their child on a suitable way but my parents completely sacrifice with everything for me. My parents' behaviors has influenced on me as they show me how to think clearly and give reasonable decisions about all my affairs especially when I am student young. They show me how to realize and identify my target and do my best to get it through understanding steps of all procedures and events that I can face. They learn me how to set my targets according to my desires and needs.

My parents provide me with necessary encouragement through all my life but I think that the best and remarkable encouragement is through my college. They never deprive me of anything one day but they always present a continuous motivation as parents are fully aware of their roles and they also aware of modern life is being. They bring up me in a natural and healthy environment. Their role is a very important to play through right bringing up. They solve all problems that face me and help me to

Unoriginal text: 9 words
www.facebook.com/ghatunasharikh...

Repetitive word: always

[Really]

[, and]

[me - I'm]

[but]

[Generally]

[characters-that - characters that]

Overused word: great

Faulty parallelism

[and-that - proud that]

possibly confused: do and make

Overused word: good

Overused word: great

Unoriginal text: 8 words
www.trustnetadmission.com/ghatunasharikh...

[me - in]

[, but]

[with]

[have - have]

[me]

[parents-provide - parents provide]

[, but]

Essay 1 – Sample of Tool ‘B’ Scores

Your Results

What the colors below mean:
Academic Style = Red | **Grammar / Spelling = Green** | **General Style = Blue**

PARAGRAPH 1:

First of all, I'd like to say that without my parents, I'd be nothing. My lovely parents always support and provide me with all desires and needs. They always help me in my life especially when I am a college student. Thanks to my parents for giving me strength and ability to understand, learn and complete this study. Really, I am very thoughtful and I ask Allah to bless them. Although I decided what to study and im loving it but I'm pretty sure I would never survive without having them standing beside me.

PARAGRAPH 2:

Generally, my parents are my first teachers. They are the most important and essential for me. I hope to be like them in all their characteristics. They have great and strong characters that have a great effect on me. They always help me in growing up and to

Report Card ?

Paragraph Rating: 😞

Sentence Rating: 😞

Vocabulary Rating: 😞

Academic Style Rating: 😞

Grammar/Punctuation Rating: 😞

General Writing Style Rating: 😞

Overall Score: 4/12 😞

Advice

- This looks like it still needs a lot of work.
- In particular, it looks like you should focus on:
 - Paragraphs

Appendix E

Sophomore student (Group B) assignment:

Argumentative essay (20%) – your argumentative research paper needs to support your opinion on an issue. Analyze the problem based on the research topic and support your analysis with evidence.

Due date: 11th December 2016

Requirements:

- 1000 – 1200 word limit (5 paragraphs)
- APA style (Time New Roman 12 font) – double spaced
- Include at least 3 academic references and citations (books or scholarly articles)
- **First draft:** Submit the hard copy together with your peer reviewed draft (Due date 27th November)
- **Second draft:** in the week from 4th – 8th December, you will be assigned to meet with me during your class time so that I may review/edit your second draft. Time for the meeting will be limited, so please ensure you have already corrected your essay for grammar, spelling and punctuation mistakes.
- **Please use a computer program to check your writing before submitting your second draft.**
- **FINAL DRAFT:** Due date: 11th December 2016 (Submit hard and soft copy)

Failure to follow all the guidelines will result in a lower grade (no exceptions!!)

Thank you and all the best.

Senior student assignment (Group A):

You'll produce a research paper on a Major based issue of your choice. Your paper must be a critical analysis of a chosen topic, based in the appropriate theory and methodology.

Due date: 11th December 2016

Requirements:

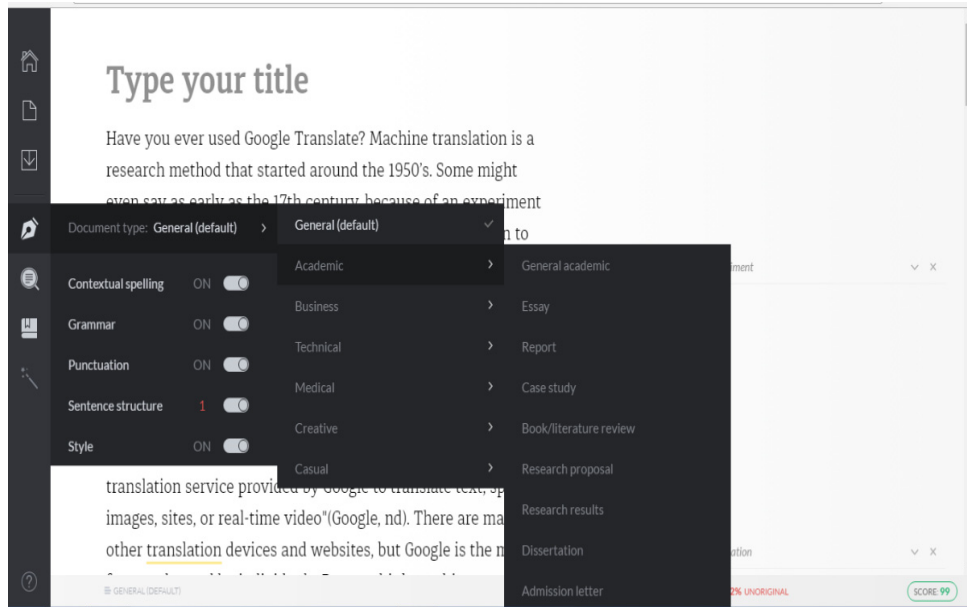
- 2500 – 3000-word limit
- Your research should be typed, double-spaced on standard-sized paper
- APA recommends using 12 pt. Times New Roman font.
- Include a **page header** (also known as the "**running head**") at the top of every page.
- Include at least 5 academic references and citations (books or scholarly articles)
- **First draft:** Submit the hard copy together with your peer reviewed draft (Due date 24th November)
- **Second draft:** in the week from 4th – 8th December, you will be assigned to meet with me during your class time so that I may review/edit your second draft. Time for the meeting will be limited, so please ensure you have already corrected your essay for grammar, spelling and punctuation mistakes.
- **Please use a computer program to check your writing before submitting your drafts.**
- **FINAL DRAFT:** Due date: 11th December 2016 (Submit hard and soft copy)

Failure to follow all the guidelines will result in a lower grade (no exceptions!!)

Thank you and all the best.

Appendix F

AWE tool ‘A’ – A Genre-specific Writing style. Sample.



Appendix G

AWE tool ‘B’ – A Feedback Sample With Explanation

Your Results

The colors below mean:

Academic Style = Red | **Grammar / Spelling = Green** | **General Style = Blue**

PARAGRAPH 1:

Have you ever used Google Translate? Machine translation is a research method that started around the 1950's. Some might even say as early as the 17th century, because of an experiment that helped translate from Chinese to English, this experiment. For those few who know, machine translation is, Translation devices are language search engines. Machine translation provides word meaning and synonyms from source language to the target language. Providing the user with a word for word translation by searching the web for word translations. "It is a free multilingual statistical machine translation service provided by Google to translate text, images, sites, or real-time video"(Google, nd). There are many other translation devices and websites, but Google is most frequently used by individuals. Do you think machine translation could replace human translators? Many people believe that computers such as Google Translate Toolkit, Babylon Translator, PROMPT, Yandex, Systran and so on will replace humans in translation. Therefore, making Translation as a huge waste of time. People debate whether or not there will still be job opportunities for freshly graduated translators. Human translators being replaced by machines in many aspects.

Report Card ?

Paragraph Rating: 😊

Sentence Rating: 😊

Vocabulary Rating: 😊

Academic Style Rating: 😞

Grammar/Punctuation Rating: 😞

General Writing Style Rating: 😞

Overall Score: 7/12 😊

Advice

- This looks like a reasonable piece of work, but it definitely needs to be checked through again. You should also make sure that it is appropriate for the assignment you have been given (e.g. does it answer your essay question?).
- In particular, it looks like you should focus on:
 - Paragraphs
 - Academic style
 - General writing style

Questions and exclamations are generally not considered acceptable in academic writing. Rewrite a question as a statement. Rewrite an exclamation to make the statement less emotional.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).