

Automated Writing Evaluation (AWE) in Higher Education: Indonesian EFL Students' Perceptions about Grammarly Use across Student Cohorts

Delsa Miranty¹, Utami Widiati^{2*}

¹Graduate Program in ELT, Universitas Negeri Malang, alan Semarang 5, Malang, East Java, Indonesia

²Universitas Negeri Malang, +62341-551312, Jalan Semarang 5, Malang, East Java, Indonesia

ABSTRACT

Automated Writing Evaluation (AWE) has been considered a potential pedagogical technique that exploits technology to assist the students' writing. However, little attention has been devoted to examining students' perceptions of Grammarly use in higher education context. This paper aims to obtain information regarding the writing process and the merits and drawbacks of Grammarly use among Indonesian undergraduate EFL students. A hundred (n=100) students majoring in English education from a public university in Banten Province were involved in this research. They were divided into three groups of users, i.e., first-, second-, and third-year student cohorts, to test whether the frequency of using Grammarly can affect their perceptions of Grammarly use in a writing class. A questionnaire and an interview guide were used to obtain the data. While the questionnaire results were analyzed using SPSS version 20, the interview results were coded and categorized based on common themes. The findings showed that there is no difference among student cohorts in their perceiving that the use of Grammarly was considered necessary to compose and revise their writing because they still dealt with several writing constraints. They thought they got immediate and comprehensive feedback, notifications of errors, and suggestions to revise the errors. Furthermore, the frequency of using Grammarly, as viewed from the student cohorts, may affect their perceptions of usefulness and drawbacks of Grammarly use, especially whether Grammarly's feedback is always helpful or not. Examining the opportunities of using Grammarly in sparking a constructive learning atmosphere in EFL writing class is worth-researching further.

Keywords: Automated Writing Evaluation (AWE), Grammarly use, Indonesian EFL writing students

INTRODUCTION

The interests among scholars in researching the ways corrective feedback is informed when evaluating English as a foreign language (EFL) students' academic writing have partly been motivated by the importance of grammatical accuracy in academic writing. Grammatical accuracy helps students convey the intended meaning by considering the grammar forms, which will help determine the quality of the whole text. The forms might include tenses, word choices, word order, subject-verb agreement, punctuations, and spellings. The importance of performing an adequate grammar accuracy level is also articulated in the Common European Framework of Reference (Council of Europe, 2001) under the linguistic repertoire category to succeed in the meaning-making process. Azar (2007) highlights that grammatical accuracy might contribute to the increased intelligibility in dealing with the four English skills, including writing. Cavaleri and Dianati (2016) also argued that quality academic writing is influenced by grammatical accuracy. It can maintain the clarity and precision of the writers' intended ideas expressed in the writing. Hence, providing corrective feedback to improve the students' writing appears to be an endless and essential discourse.

Corrective feedback serves as the gatekeeper of the quality of students' writing. The students can recognize and measure

their capacity in composing writing from the results of the corrective feedback. They can also maintain their writing performance from the reflection on the feedback given by the teachers. The benefits of corrective feedback have been empirically justified by previous scholars (e.g., Isnawati, Sulisty, Widiati & Suryati, 2019; Wang, Shang, & Briody, 2013; Yang, 2010), asserting that feedback can influence students' writing quality. The study by Isnawati et al. (2019) revealed that when teacher corrective feedback was combined with a teacher-student conference, the students under the study had a greater mean score of writing performance, implying

Corresponding Author e-mail: utami.widiati.fs@um.ac.id

https://orcid.org/0000-0002-8603-4556

How to cite this article: Miranty D, Widiati U, (2021). Automated Writing Evaluation (AWE) in Higher Education: Indonesian EFL Students' Perceptions about Grammarly Use across Student Cohorts. Pegem Journal of Education and Instruction, Vol. 11, No. 4, 2021, 126-137

Source of support: Nil

Conflict of interest: None.

DOI: 10.47750/pegegog.11.04.12

Received: 23.05.2021

Accepted: 03.08.2021

Publication: 01.10.2021

a possibility of positive impacts of teacher written corrective feedback combined with a teacher-student conference.

Recent rapid technological changes have affected how corrective feedback can be provided, especially in writing classes. Lim and Phua (2019) stated that digital tools can improve the students' ability to write well and promote constructive feedback in the learning context. Remarkably, the technological development in this globalized world has triggered teachers, especially in the higher education context, to utilize automated writing evaluation (AWE) tools to execute written corrective feedback.

AWE in L2 Classrooms: Potentials and Challenges

By definition, automated writing evaluation (AWE) assesses written texts using automatic rating engines developed by computational linguistics (Wilson & Andrada, 2016). The assessment may range from lexical, syntactic, and grammatical levels to semantic and discourse levels (Chen & Cheng, 2008). The inception of AWE was in the 1960s, widely known as Page Essay Grade (PEG). The program applied multiple regression analysis to measure the writing quality using previously hand-scored essays as the referred texts (Shermis, Mzumara, Olson, & Harrington, 2001). Then, Educational Testing Service (ETS) collaborated with Vantage Learning developed e-rater and Intellimetric in the 1990s. Both scoring programs were considered more thorough in evaluating the lexical, syntactic, and discourse aspects of writings (Elliot & Mikulas, 2004). Pearson Knowledge Technologies also purchased Intelligent Essay Assessor to score essays. What differentiates this program from other AWE programs is that this program uses latent semantic analysis. The semantic meaning of the analyzed texts is compared to a corpus under a similar topic. Nowadays, several AWE tools have been used to rate and evaluate different forms of written genres, such as Criterion and My Access (Li, Link, & Hegelheimer, 2015). Previously, AWE tools were used for high-stakes testing, such as summative tests (Cotos, 2014; Stevenson & Phakiti, 2019). In line with technological advancement and the importance of technology in providing quick and thorough corrective feedback, many AWE tools have been used as an alternative pedagogical technique in L2 writing classrooms.

At least, there are two essential components of AWE. First, AWE tools generate automated scores. They are designed to provide quick and thorough evaluation results of the analyzed texts so that the evaluation process is not time-consuming. Teachers can use the scores as a source of reflection to decide on future actions. Second, AWE tools provide automated, written corrective feedback (Ranalli, 2018). The students can use the feedback to revise their papers autonomously, but the teachers' role is still vital to clarify the feedback. Such a teacher role implies the benefits of implementing teacher-student conferences as explored by Isnawati et al. (2019).

Debates on the use of AWE tools remain in the context of L2 learning. Some scholars (Bai & Hu, 2017; Kern & Warschauer, 2000; Rezaei & Lovorn, 2010; Wang, Shang, & Briody, 2013; Ware & Warschauer, 2006; Zhang & Hyland, 2018) consider AWE a helpful means of assessing the students' writing. The automated rating engines allow the students to get immediate scores and written corrective feedback. The scores and feedback are considered more objective and consistent than human raters. The feedback can be quantitative and qualitative. Therefore, AWE tools can elevate students' motivation and raise their awareness to do several revision stages. It is also designed to be consistent over time so that the evaluation process can be less time-consuming. Alternatively, the students can use the score and feedback to have self-reflection and self-revision; in other words, they can be trained to develop their autonomy. The opportunity to do self-reflection and self-revision through AWE tools can also lessen their writing anxiety because it is self-paced and personalized. The personalization of the evaluation process makes the students not worry about negative judgment from either teachers or peers. From the teachers' perspectives, AWE tools can help teachers provide corrective feedback for the students. Detailed evaluation results from AWE tools enable the teachers to interact with the students to deal with the revision process. While the tools can inform the form-focused feedback such as grammatical accuracy and mechanics, the teachers can devote their time and efforts to evaluate the text-focused feedback such as content and organization.

However, doubts about the use of AWE tools to foster learner autonomy remain on the discussion. They were found not to help develop the students' attitudes toward learner autonomy. Cheng and Cheng (2008) asserted that the AWE tools did not necessarily foster the students' autonomy because the feedback was not concerned with content. Hence, the negotiation process and interaction about the feedback might be even discriminating against the students who are not familiar with the technology. Consequently, the motivation to improve their writing can be less evident. The independent use of the tools was also questioned because the students' writing would be unauthentic. The students tended to rely on the feedback given by the tools much. Inspired by these inconclusive previous research findings, the present study demystifies potential benefits of AWE tools for EFL writing instruction.

Grammarly in L2 Classrooms

As a form of AWE, Grammarly is the center of attention in academic writing in the higher education context. It is an American technology company in the English language digital writing tool, developed by the original Ukrainian authors Alex Shevchenko, Max Lytvyn, and Dmytro Lider in 2009. This tool combines artificial intelligence and natural

language processing under the machine and deep learning algorithms to provide real-time, quick, and comprehensive writing evaluation results. Grammarly's products offer not only grammar checking but also spell checking and plagiarism detection. Now, six products have been launched: The Grammarly Editor, Browser Extensions, Grammarly for Microsoft Office®, Grammarly for Your Desktop, The Grammarly Keyboard, and Grammarly for iPad. Generally, the tools are intended to help people strengthen their writing abilities and express ideas clearly and correctly (Grammarly, 2020). In Indonesia, the use of Grammarly is not obligatory especially in the context of higher education. The universities are given a discretion to determine automated writing evaluation tools beneficial for their language courses. Hence, Grammarly seems not immensely used, considering no specific regulation for such purposes.

A considerable body of research has examined the merits and drawbacks of using Grammarly tool in EFL writing classes. Regarding the merits of using Grammarly in writing classes, Japos (2013) stated that Grammarly could substantially decrease grammar errors. Likewise, Cavaleri and Dianati (2016) argued that Grammarly was perceived by the students as a valuable, helpful, and easy-to-operate AWE tool. It helped them understand grammar rules. Grammarly allows the students to understand the errors and the explanation; they can consider the suggestions given by the tool to revise their writing. Darayani, Karyuatry, and Rizqan (2018) and O'Neill and Russell (2019) justified the previous research that the students perceived Grammarly as a powerful tool enabling user writing to be easily checked for grammar possible stylistic errors. Their study also revealed that the students were more pleased with the feedback from Grammarly than that from teachers. Hence, writing quality can be increased. This positive effect of using Grammarly was reported in several studies (El-Ebyary & Windeatt, 2010; Li, Link, & Hegelheimer, 2015; Liao, 2015; Parra & Calero, 2019; Wang et al., 2013). The significance of using Grammarly in developing the students' attitudes and behaviors was also evident.

Cavaleri and Dianati (2016) found that the students were willing to continue using the tool. Ventayen and Orlanda-Ventayen (2018) unveiled that the students could get valuable lessons about grammatical errors so that their grammatical awareness could lead toward a better revision process. It is corroborated by Parra and Calero (2019) that the students' understanding of the errors increased. As a result, their writing accuracy can be enhanced autonomously. In the Indonesian context, Ghufron and Rosyida (2018) examined Grammarly's usefulness by conducting a quasi-experimental study involving 40 third-semester EFL students. The results showed that in addition to the reduced errors in diction, grammar, spelling, and punctuation, Grammarly could enhance the students' motivation to be autonomous learners. Nova's (2018) study

echoed the previous research in that using Grammarly encouraged the users to activate independent learning and self-revision process in their writing. This all highlights writing as a process.

However, the Grammarly tool is not without flaws. This tool seems to lack comprehensive results, especially related to the quality of the content and organization. It cannot capture the coherence level between sentences and paragraphs. Misleading feedbacks were also noted in Nova's (2018) study in which the students identified inaccuracies because the results did not conform to their intention. Grammarly's over-checking feature is another challenge because it may consider the reference writing incorrect, mostly the authors' names. As Nova (2018) found out, this tool's most noticeable weakness is that Grammarly can only approve American English to automatically consider British English, particularly the spelling incorrect.

Moreover, doubts also remain on whether this technology is beneficial for actualizing meaningful learning. Interaction, discussion, and collaborative reflection as vital elements in the writing process seem to be not adequately addressed using Grammarly since AWE programs are designed to be used as self-study and self-paced learning media. In other words, it lessens the opportunity for the students and their teachers to talk about the feedback provided by the tool as a critical factor in developing writing quality and second language acquisition (Zhang & Hyland, 2018).

Although the integration of Grammarly in writing classes has been quite extensively examined, little attention is devoted to examining this issue in the Indonesian higher education context. There is no obligation for the university or higher education institution in Indonesia to use AWE tools such as Grammarly because it is at the teacher's discretion. It is also not compulsory for the students to use AWE tools such as Grammarly before submitting writing tasks to the faculty. Grammarly's incorporation might be new for some undergraduate students because their writing is usually checked manually by their teachers. Raised by this circumstance, the efforts to delve into the students' perceptions of using Grammarly in writing classes can help evaluate the usability of the tool. Additionally, the results can reflect the students' attitudes toward Grammarly use so that teachers can make evaluative actions to maintain the usefulness in writing classes. In short, this study seeks to explore the perceptions of Indonesian undergraduate EFL students from different cohorts after using Grammarly in their writing class. The following research questions drive the present study.

RESEARCH QUESTIONS

- **RQ 1:** Is there any difference between the students across cohorts regarding their perceptions about the writing process?

- RQ 2: Is there any difference between the students across cohorts regarding their perceptions about the use of Grammarly in the writing class?

METHOD

Research Design

This study employed a quantitative approach in obtaining a numerical description of the students' perceptions about the use of Grammarly in writing classes. Inferential statistical results were also brought to reveal the differences between groups of students. Additionally, a qualitative approach was used to examine further the students' more profound and broader perspectives regarding the use of Grammarly as a form of AWE tools. It appears crucial to understand whether there are variations of perceptions among the students of different year cohorts.

Study Group

The study involved at first 133 undergraduate students (students of the first year = 51 students; second year = 51 students; third year = 31 students) from the English education department of a public university in Banten Province, Indonesia. Accessibility was the main reason for involving those research participants. This study compared the perceptions of different student cohorts, that is, to examine whether there were significant differences among the students who used Grammarly for the first time and those who had already used it before, and among those who used Grammarly purposefully in a writing course. While the first-year students experienced using Grammarly for the first time, the second and the third-year students may have experienced using such an AWE tool for more than one time.

Data Collection Tools

Two instruments were employed to obtain adequate data to address the research questions. A 5-point Likert-scale questionnaire was adapted from Cavaleri and Dianati (2016); the questionnaire consisted of fifteen items. Table 1 depicts the details of the questionnaire items. The questionnaire was formulated initially in English for two reasons. First, the English education department students already had basic English knowledge to understand the questionnaire items. Second, as the questionnaire was adapted from Cavaleri's and Dianati's (2016) study, the English-version questionnaire items were preferred to avoid losing the original meaning from the primary source. The validity and reliability tests of the questionnaire items were performed by using SPSS 20. The Pearson correlation coefficient value is above the r table .194 with the significance level 0.05 (two-tailed) for the items with either positive or negative meaning. The Cronbach's Alpha value is .767 for the items with positive meaning and .461 for the items with negative meaning. It means that the

Table 1: Description of the questionnaire items.

| Number of items | Type | Domain |
|-----------------|----------------------|--|
| Five items | 5-point Likert scale | Students' perceptions about the writing process |
| Four items | | Students' perceptions about the usefulness of Grammarly in writing classes |
| Six items | | Students' perceptions about the drawbacks of Grammarly in writing classes |

questionnaire items are valid and moderately reliable. Then, an interview guide was formulated based on the questionnaire items to obtain more comprehensive data.

Data Collection

The data collection procedure started with getting the students' consent to follow the overall research process. Then, the questionnaire was sent to the students using Google Forms. Out of 133 students, 100 students resent the questionnaire completely, comprising 40 from the first year, 38 from the second, and 22 from the third. The remaining thirty-three questionnaires were then discarded from the analysis. Fifteen students who returned the questionnaire sheets were further involved in the semi-structured interview sessions. The students were selected based on the categorization of their GPA (1.00-4.00): lower-achievers (2.95-3.45), middle-achievers (3.46-3.71), and higher-achievers (3.72-4.00). Five students from each category were invited to participate in the interview sessions.

Data Analysis

The data analysis procedure started with the responses to the questionnaire items. The results were calculated descriptively using SPSS version 20 to obtain the frequency and the mean score of each questionnaire item. The mean scores were then used to determine three categories regarding the frequency of Grammarly use as a learning strategy. A Strategy Inventory of Language Learning (SILL) developed by Oxford (1995) was employed to divide the mean scores of the questionnaire items into three categories: high, medium, and low levels of Grammarly use, as depicted in Table 2. Based on the categorization from SILL, the categories identified in the analysis were divided into three levels: (1) High level of Grammarly use (HG); (2) Medium level of Grammarly use (MG); and (3) Low level of Grammarly use (LG).

Furthermore, a one-way ANOVA test was run to measure the extent to which each item of the questionnaire responses was different within each group of the students or between the three groups, i.e., first, second, and third years. The mean scores of each item were recalculated. Here are the hypotheses to be tested.

- H_0 : There is no difference of perceptions between and within first-, second-, third-year student groups about the use of Grammarly in the writing class.
- H_1 : There is a difference of perceptions between and within first-, second-, third-year student groups about the use of Grammarly in the writing class.

After the questionnaire responses were processed using descriptive and inferential statistical measurements, the researchers analyzed the interview transcripts to enrich the quantitative data. The researchers manually analyzed the interview transcripts. First, the interview audio recordings were transcribed. Because the interview sessions were conducted in English, the transcript was not edited unless there were severe grammatical errors that influenced its clarity. Second, the transcript was coded verbatim. Third, the coding results were categorized into several groups to corroborate the questionnaire results to address the research questions. A member checking was conducted by having the transcripts proofread and reviewed by the respondents to avoid biased meaning and completeness. The respondents agreed that the transcripts were already eligible to be analyzed and presented.

FINDINGS

- **RQ 1:** Is there any difference between the students across cohorts regarding their perceptions about the writing process?

Table 3 depicts the questionnaire results regarding the students' perceptions of themselves in the writing process,

Table 2: Categorization of Grammarly Use.

| Category | Indicator | Interpretation |
|----------|------------|------------------------------|
| High | 4.5 to 5.0 | Always or almost always used |
| | 3.5 to 4.4 | Usually used |
| Medium | 2.5 to 3.4 | Sometimes used |
| Low | 1.5 to 2.4 | Generally not used |
| | 1.0 to 1.4 | Never or almost never used |

including the mean scores and standard deviation values of each questionnaire item.

The results showed that the students from all groups perceived proofreading service as not considerably essential factor in the writing process ($\bar{x} = 2.41$,). They still preferred feedback from others to assist them to write in English. However, based on the results of questionnaire item 5, the students still believed that the feedback given by their lecturers was understandable enough in the writing class ($\bar{x} = 2.61$,). It might imply that they needed a tool to offer feedback in their writing class, particularly their papers' grammar. As perceived by the students, the importance of proofreading service was associated with the students' writing difficulties. Most of the students also agreed that their vocabulary and grammar knowledge was still not excellent ($\bar{x} = 3.47$, and expressing thoughts in writing was not an easy task ($\bar{x} = 3.37$, , regardless of their year level. The students' need to get assistance was related to the fact that they did not have adequate confidence in formulating correct sentences $\bar{x} = 3.49$.

The ANOVA test results revealed that the significance value for each of the five questionnaire items is lower than $\alpha=0.05$ (Item 1 .994; Item 2 .365; Item 3 .343; Item 4 .565; Item 5 .591). The results reinforce that they find the teacher's assistance in the writing process is crucial for them. They are still not confident enough with their vocabulary and grammar knowledge as well as their ability to express ideas in their writing, regardless of their year level. They also perceived that the teacher's feedback on their writing is understandable and easy to follow.

The fact that most of the students were still not excel at grammar and vocabulary was corroborated by the following interview excerpts.

My English grammar and vocabulary are not too weak and not too advanced. It's kind of medium level. Sometimes, I'm confused to use them. It's all based on what I want to express. (A third-year student)

Sometimes, I am confused to express my ideas in writing because my weaknesses in English Grammar. So, I think very

Table 3: The Students' Perceptions of Statements about Themselves in the Writing Process.

| Statement | Group A | | Group B | | Group C | | \bar{x} Mean | \bar{x} SD |
|---|---------|------|---------|-----|---------|------|----------------|--------------|
| | Mean | SD | Mean | SD | Mean | SD | | |
| 1. I do not need any help with writing in English; I just need a proofreading service | 2.40 | .97 | 2.42 | .93 | 2.40 | .79 | 2.41 | .90 |
| 2. My knowledge of English grammar and vocabulary is weak | 3.57 | .65 | 3.33 | .85 | 3.50 | .56 | 3.47 | .70 |
| 3. I do not always feel confident that I have written the correct sentences | 3.31 | 1.05 | 3.54 | .75 | 3.62 | .87 | 3.49 | .90 |
| 4. I am fine with English grammar, but I find it difficult to express my ideas in writing | 3.42 | 1.06 | 3.21 | .99 | 3.46 | 1.07 | 3.37 | 1.04 |
| 5. I do not always understand the feedback I get in my writing | 2.71 | .89 | 2.48 | .83 | 2.62 | 1.03 | 2.61 | .91 |

Note: Group A: 1st year, Group B: 2nd year, Group C: 3rd year

Table 4: The One-Way ANOVA Test Results.

| | | <i>Sum of Squares</i> | <i>df</i> | <i>Mean Square</i> | <i>F</i> | <i>Sig.</i> |
|--------|----------------|-----------------------|-----------|--------------------|----------|-------------|
| Item 1 | Between Groups | .011 | 2 | .005 | .006 | .994 |
| | Within Groups | 80.179 | 97 | .827 | | |
| | Total | 80.190 | 99 | | | |
| Item 2 | Between Groups | 1.005 | 2 | .503 | 1.018 | .365 |
| | Within Groups | 47.905 | 97 | .494 | | |
| | Total | 48.910 | 99 | | | |
| Item 3 | Between Groups | 1.765 | 2 | .883 | 1.081 | .343 |
| | Within Groups | 79.225 | 97 | .817 | | |
| | Total | 80.990 | 99 | | | |
| Item 4 | Between Groups | 1.255 | 2 | .627 | .574 | .565 |
| | Within Groups | 106.055 | 97 | 1.093 | | |
| | Total | 107.310 | 99 | | | |
| Item 5 | Between Groups | .905 | 2 | .452 | .529 | .591 |
| | Within Groups | 82.885 | 97 | .854 | | |
| | Total | 83.790 | 99 | | | |

important for mastering English Grammar before you start writing. (A third-year student)

Some of the students reported that their grammar skills and vocabulary mastery were not advanced enough. Both weaknesses affected their ability to write in which sometimes they felt confused at using both language aspects in their writing well. They also thought that mastering English grammar was of great importance when dealing with writing activity.

- **RQ 2:** Is there any difference between the students across cohorts regarding their perceptions about the use of Grammarly in the writing class?

Table 5 displays the questionnaire results regarding the use of Grammarly. The students agreed that Grammarly was not difficult to operate. In particular, using Grammarly could assist their writing process because the AWE tool gave them detailed feedback ($\bar{x} = 3.82$,). They also found that Grammarly provided explanations of the writing errors, not only the feedback ($\bar{x} = 3.97$). As a result, it was no surprise that the students perceived Grammarly as a helpful tool to understand the grammar rules ($\bar{x} = 3.88$), leading to the positive perception among them about the role of Grammarly's suggestions to revise and improve their works ($\bar{x} = 4.00$). All in all, the usefulness of using Grammarly was quite evident among the students ($\bar{x} = 3.65$,).

The mean scores were further linked with SILL to determine the frequency of using Grammarly per item. Table 2 in the method section, the questionnaire items number 6, 7, 8, 9, 14, 15 indicate that their considerably positive perceived usefulness of using Grammarly was because they usually used

the tool. These questionnaire items were under the category of high level.

Table 6 depicts the one-way ANOVA test results of the questionnaire items. The results revealed that the significance value of each of the questionnaire items presented before was lower than $\alpha=0.05$ (Item 6 .975; Item 7 .447; Item 8 .102; Item 9 .999; Item 14 .258; Item 15 .911). These findings imply that generally the students perceived usefulness of Grammarly regarding its details of feedback to understand grammar rules and suggestions to revise the writing. However, their perceived usefulness of the explanations of the errors given by such AWE tool still varies within or across the three groups, i.e., first-, second-, and third-year student groups. In other words, the "high category" of Grammarly use may not automatically result in conformed perceptions among the students of the usefulness of Grammarly in this case.

The interview results corroborated the perceived usefulness of using Grammarly in that the students experienced the benefits in writing class. They said that they used Grammarly to help them overcome problems related to the projects and materials provided in the writing class, support them by automatically correcting their paper, check their writing assignments, and improve their writing performances.

To be honest, for the first time I used Grammarly, I was shocked. Then, I decided to continue using Grammarly because for me it is very good and very quick to give corrections on my writing. (A first-year student)

Grammarly can be installed on our smartphone, and we can change our keyboard with Grammarly so that it can

Table 5: Students' Perceptions about Ways that Grammarly is Helpful.

| Statement | Group A | | Group B | | Group C | | \bar{X} Mean | \bar{X} SD |
|--|---------|-----|---------|------|---------|------|----------------|--------------|
| | Mean | SD | Mean | SD | Mean | SD | | |
| 6. Grammarly gives detailed feedback | 3.80 | .79 | 3.81 | .80 | 3.84 | .80 | 3.82 | .79 |
| 7. Grammarly makes helpful suggestions for improving my work | 4.05 | .72 | 3.87 | .69 | 4.06 | .56 | 4.00 | .66 |
| 8. Grammarly gives good explanations about my errors | 3.88 | .67 | 3.84 | .83 | 4.18 | .53 | 3.97 | .70 |
| 9. Grammarly has helped me understand grammar rules | 3.88 | .79 | 3.87 | .89 | 3.87 | .79 | 3.88 | .81 |
| 14. I do agree that student get usefulness with the use of Grammarly in writing class. | 3.45 | .88 | 3.81 | .80 | 3.68 | 1.02 | 3.65 | .91 |
| 15. Grammarly is easy to use, especially in writing class. | 3.65 | .83 | 3.75 | 1.06 | 3.71 | .99 | 3.71 | .95 |

Table 6: The One-Way ANOVA Test Results.

| | | Sum of Squares | df | Mean Square | F | Sig. |
|---------|----------------|----------------|----|-------------|-------|------|
| Item 6 | Between Groups | .032 | 2 | .016 | .025 | .975 |
| | Within Groups | 62.728 | 97 | .647 | | |
| | Total | 62.760 | 99 | | | |
| Item 7 | Between Groups | .724 | 2 | .362 | .812 | .447 |
| | Within Groups | 43.276 | 97 | .446 | | |
| | Total | 44.000 | 99 | | | |
| Item 8 | Between Groups | 2.250 | 2 | 1.125 | 2.338 | .102 |
| | Within Groups | 46.660 | 97 | .481 | | |
| | Total | 48.910 | 99 | | | |
| Item 9 | Between Groups | .002 | 2 | .001 | .001 | .999 |
| | Within Groups | 66.558 | 97 | .686 | | |
| | Total | 66.560 | 99 | | | |
| Item 14 | Between Groups | 2.280 | 2 | 1.140 | 1.374 | .258 |
| | Within Groups | 80.470 | 97 | .830 | | |
| | Total | 82.750 | 99 | | | |
| Item 15 | Between Groups | .175 | 2 | .087 | .094 | .911 |
| | Within Groups | 90.415 | 97 | .932 | | |

automatically correct my typing. It can be used anytime and anywhere we want, including in our writing class. (A third-year student)

Grammarly corrects my writing and I can learn from it what us the correct sentence structure. (A second-year student)

Meanwhile, the difference between the student cohorts in perceiving the helpfulness of Grammarly was corroborated by the following excerpts.

To be honest, in the first time I used Grammarly, I was surprised. (A first-year student)

Well, that's a plus point from Grammarly. However, in my experience, sometimes Grammarly missed a few points in terms of plagiarism. (A second-year student)

It provides you with the basics you need to lay down your writing foundation, spelling, grammar, and punctuation checker, which is more powerful than the average spell

check you get with most word processors. (A third-year student)

I think the user can't rely on this application over and over again. It only addresses common mistakes. So, learning by ourselves would be better than by using an AI assistance. (A third-year student)

The excerpts above demonstrated that the first exposure to Grammarly experienced by the students in the second semester made them feel surprised. In a more comprehensive way, one of the students in the fourth semester argued that Grammarly was not considerably helpful in terms of checking the plagiarism rate of the content. As time goes by, the students perceived that Grammarly was a more powerful tool, as compared to the average similar tools. This discrepancy may indicate the role of repetition of exposure to get the students more familiar with the use of Grammarly to assist their writing process independently. What was also found from the student

Table 7: Students' Perceptions about the Drawbacks of Grammarly.

| Statement | Group A | | Group B | | Group C | | \bar{X} Mean | \bar{X} SD |
|---|---------|-----|---------|-----|---------|-----|----------------|--------------|
| | Mean | SD | Mean | SD | Mean | SD | | |
| 10. The feedback is not always helpful | 3.05 | .76 | 2.57 | .79 | 2.59 | .75 | 2.75 | .79 |
| 11. I do not agree with some of the suggestions | 3.08 | .50 | 2.72 | .83 | 2.75 | .76 | 2.86 | .72 |
| 12. I cannot understand the explanations | 2.68 | .96 | 2.51 | .75 | 2.56 | .66 | 2.59 | .80 |
| 13. I have technical issues with Grammarly | 3.20 | .83 | 3.03 | .76 | 2.75 | .84 | 3.00 | .82 |

Table 8: The One-Way ANOVA Test Results.

| | | Sum of Squares | df | Mean Square | F | Sig. |
|---------|----------------|----------------|----|-------------|-------|------|
| Item 10 | Between Groups | 5.085 | 2 | 2.542 | 4.277 | .017 |
| | Within Groups | 57.665 | 97 | .594 | | |
| | Total | 62.750 | 99 | | | |
| Item 11 | Between Groups | 2.752 | 2 | 1.376 | 2.708 | .072 |
| | Within Groups | 49.288 | 97 | .508 | | |
| | Total | 52.040 | 99 | | | |
| Item 12 | Between Groups | .530 | 2 | .265 | .404 | .669 |
| | Within Groups | 63.660 | 97 | .656 | | |
| | Total | 64.190 | 99 | | | |
| Item 13 | Between Groups | 3.430 | 2 | 1.715 | 2.577 | .081 |
| | Within Groups | 64.570 | 97 | .666 | | |

cohort in the sixth semester was another student articulating a contrasting view on Grammarly. The student suggested the users not relying much on that application. Instead, the student preferred learning by themselves to utilizing the AI-based writing correction tool to assist their writing process and progress. This evidence indicates that the length of experience may not determine the conformity of perceptions among the students in the same cohort.

Furthermore, Table 7 shows the students' perceptions about drawbacks related to Grammarly use. The questionnaire results regarding the perceived drawbacks of using Grammarly in writing class were more intriguing. Some of the students generally viewed that despite the positive results of Grammarly's perceived usefulness, they still experienced technical errors in using the tool ($\bar{X} = 3.00$). The feedback was not always perceived useful by some students ($\bar{X} = 2.75$). Also, some suggestions were seen not in line with what they expected ($\bar{X} = 2.59$), or some explanations were not understandable enough ($\bar{X} = 2.59$). Apart from that, the results revealed that most students still perceived Grammarly as an easy-to-use and helpful AWE tool to obtain valuable feedback and explanation regarding the errors to refer to them in revising their works.

The ANOVA test results of the questionnaire items are showcased in Table 8. The results showed that the significance value of questionnaire item 12 was higher than $\alpha = 0.05$ (.669), representing no difference between the students across cohorts in their perceptions about the clarity of Grammarly explanation. The significance value of questionnaire items

11 and 13 was slightly higher than $\alpha = 0.05$ (.072 and .081, respectively), indicating slightly similar responses between the three groups about their disagreement with either some suggestions provided by Grammarly or its ease of use. However, there was a difference between the students across year levels in terms of how Grammarly's feedback was considered not always helpful because the significance value of questionnaire item 10 was lower than $\alpha = 0.05$ (.017). It means that more students were still in between whether they experienced the drawbacks or not.

DISCUSSION

This study aimed to explore the perceived usefulness and drawbacks of Grammarly as an AWE tool among first-, second-, and third-year Indonesian undergraduate EFL students. Regarding the first research question, "Is there any difference between the students across cohorts regarding their perceptions about the writing process?" there was no difference in terms of perceptions of Grammarly use among students of different cohorts. The students under study perceived that proofreading service is needed to produce good writing and that one of the uses of Grammarly is for proofreading services. The fact that there is no significant difference in mean across the three student cohorts indicates that length of learning English in higher education settings may not influence the increased perceived self-efficacy in writing in English. It means that several students are already confident enough in writing in English without any assistance even a proofreading

service. They admitted the need for automated feedback because they still dealt with several writing deficiencies, such as lack of confidence and abilities to express their ideas in a well-structured manner. They felt that their knowledge of English grammar and vocabulary seems not sufficient enough to write in English properly. However, all student cohorts do not considerably agree that the feedback given by the lecturers is difficult to understand. This may imply that albeit understandable, the lecturers' feedback seems not influential enough to encourage the students to improve writing skills because the sixth-semester students still have a tendency for not being confident at their writing and grammar skills. Such circumstance again results in the need for help from a proofreading service, instead of relying merely on the lecturers' feedback. The reason for this can be associated with the clarity, depth, and comprehensiveness of feedback given by the lecturers or the techniques employed to convey the feedback to the students.

Regarding the second research question, "Is there any difference between the students across cohorts regarding their perceptions about the use of Grammarly in the writing class?" the present study also showed that Grammarly use is perceived useful because it provides practical explanations and suggestions for improving the writing among the students. The one-way ANOVA test results demonstrated that the hypothesis H0: There is no difference of perceptions between and within first-, second-, and third-year student groups regarding the use of Grammarly in the writing classroom was accepted regarding the students' perceptions of the usefulness of Grammarly use. The students' responses further explain that the common categories of errors provided by Grammarly encompassed spellings and punctuations, subject-verb agreement, and structure, as justified by Daniels and Leslie (2013). In other words, the students across cohorts generally view Grammarly as useful automated writing correction tool because they can get explicit results of the grammar errors contained in their writing. The findings corroborated Salteh and Sadeghi (2010) that the students preferred explicit error corrections.

However, differences were identified in which one of the students from the 2nd-semester student cohort feels shocked because it is the first time to use Grammarly. Quite similarly, a student from the 4th-semester student cohort views such AI-based AWE tool not considerably helpful in terms of identifying plagiarism rate. Contrastingly, the students from the 6th-semester student cohort perceive Grammarly as a helpful tool to understand the basic writing foundation, albeit not mentioning the issue of plagiarism rate. This discrepancy may indicate the role of repetition of exposure to get the students more familiar with the use of Grammarly to assist their writing process independently. However, another intriguing perception is that one student from the same 6th-semester student cohort does not considerably agree with

the usefulness of Grammarly. The student recommends the users to better learn from themselves, not relying much on the software since it only corrects common writing errors. This may counter the role of repetition of exposure toward increased familiarity with the utilization of the software. It contrasts with Parra and Calero's (2019) study that the students' perception of the AWE tools, Grammark and Grammarly, is positive toward the development of their writing skills to some extent. One of the reasons of such discrepancy might be associated with the students' diverse learning styles or strategies in the present study. The unfavored perception of Grammarly use may also be connected with the value of interactions between the students and the lecturers. Since the feedback given by Grammarly is based on the writing being analyzed, there is no further explanation for the students of why the results can be so. This confirms Zhang and Hyland (2018) that opportunities for open interactions are considerably helpful for students to enhance their second language acquisition and writing quality. Considering the diverse learning styles or strategies, the teachers should take into account various techniques to deliver the feedback as well.

The perceived less comprehensive correction provided by the software may be caused by the student's preferences on certain types of corrective feedback. Focused and comprehensive written corrective feedback needs to be revisited because both strategies of giving feedback to students still receive ample attention among scholars (see Lee, 2020 for a more comprehensive review on this issue). She further recommends the possibility of combining focused and comprehensive WCF. On the one hand, the time allocation in the classroom can indeed influence the teachers to focus on major issues in the writing. On the other hand, the students' needs may contradict to the teachers' focused WCF given. The present study accords with Lee's (2020) final thought that in an authentic writing classroom, addressing what really matters according to the students' needs in an authentic writing classroom is essential.

Grammarly gave the explanations and corrective feedback that the students can use as self-paced learning resources. This finding confirms Dodgson et al. (2016) that using English Language Learning Websites as learning materials helps students cope with indirect corrective feedback because they need references to the learning process's corrective feedback. Qassemzadeh and Soleimani (2016) also argue that the feedback from AWE tools can significantly affect learning passive forms. The feedback has a statistically significant influence on the retention of passive forms from Grammarly and teachers among Iranian EFL students.

Another benefit perceived by the students is the easy-to-operate feature of Grammarly. Cavaleri and Dianati (2016) found in their study that more than 80% of the participating students give a positive evaluation of the tool. The reason is

that Grammarly can be used as either a stand-alone website or an add-on in several information-processing tools such as Word, Search Engine, and educational websites. Grammarly is realistic, which has colors that make the students recognize the errors in their writing. The automated written corrective feedback also helped them understand their mistakes. In short, it was beneficial to decide what to revise and what to write from the feedback. Like an AI-based AWE tool, Grammarly was considered easy-to-operate because the students could use it anytime and anywhere. Hence, most of the three groups accepted Grammarly use to assist their learning and writing processes.

Nova (2018) justified the perceived ease of use of Grammarly that the students can get immediate access to and monitor the feedback in their accounts to save more time to self-edit and self-revise their works. In sum, Grammarly use in a writing class for higher education is perceived quite well because it helps the students get immediate and comprehensive feedback, eases the access to the explanations of the errors, improves knowledge, particularly about grammar, and provides some suggestions to revise the papers. The development of AWE tools has allowed the students to understand what they want to learn and acquire the knowledge they need so that they know their potentials and weaknesses in writing and in fulfilling their own learning needs independently; in other words, the use of Grammarly as revealed in this study might stimulate EFL students in writing classes to become autonomous.

However, several drawbacks of Grammarly are still identified in the present study. Several students still experienced technical errors in using Grammarly. This might be caused by the infrequent use of the writing class tool and the absence of learner training in using the technology. The questionnaire results also reveal that Grammarly's feedback does not always help some students because the feedback might be challenging to digest, not comprehensive enough, or not conform to their expectations. Previous research (Ghufron & Rosyida, 2018; Nova, 2018; Parra & Calero, 2019) resonates with the present study. Although Grammarly has been developed to detect run-on sentences, dangling expressions, sentence fragments, and difficult-to-understand sentences, such an AWE tool can still detect the organization, structure, or flow of the ideas or content. However, H_1 : *There is a difference of perception between and within second-, fourth-, sixth-semester student groups regarding the use of Grammarly in the writing classroom* is accepted for the questionnaire item 10 (Sig. value $.017 < \alpha=0.05$). The second- and third-year students disagree with the statement that Grammarly's feedback is not always helpful, which differs from the first-year students. This difference might be caused by the intensity level in using Grammarly in writing class so that the first exposure of Grammarly use to the first-year students may influence their views on its usefulness.

Moreover, the evaluation results are sometimes perceived as not satisfactory by some of the students. One example is noticed in Nova's (2018) study, where Grammarly, to date, is not capable of catering for varieties of English, only American English spellings and punctuations. It might be caused by the complex nature of writing and the variety of styles in using written English, which cannot be comprehensively evaluated by AI tools. This domain is beyond forms and semantics so that it needs the teacher's role to handle this weakness. Hence, interaction and discussion between students and their teachers can foster a festive learning atmosphere, contributing to meaningful students' writing class engagement, in line with the study by Isnawati et al. (2019). Although students' engagement with written corrective feedback has been explored quite extensively (see Zhang, 2017; Zheng & Yu, 2018), this issue remains underexplored in the context of the current development of AWE such as Grammarly.

CONCLUSION

The data sets have delineated the Indonesian undergraduate EFL students' perceptions of Grammarly use in writing classes. The findings reveal that the students across the year levels agree regarding their perceptions of Grammarly as a useful AWE tool because they are already aware of the need to obtain proofreading services. Non-conformities, however, are still identified across the student cohorts or among the students within the same cohort. This study posits that the length of experience in and exposure of using Grammarly seems not the only factor, which influences the students' perceptions of the usefulness of Grammarly. Apart from all the merits of Grammarly as an AI-based AWE tool, the drawbacks remain on the surface of discussion. The infrequent use of Grammarly, as viewed from the student cohorts, may affect their perceptions about whether the feedback given by Grammarly is always helpful.

SUGGESTION

The fact that Grammarly is not capable of detecting and evaluating unorganized content or ideas and of accommodating English varieties in this more globalized era raises a classical question on whether AWE can replace teachers' roles to maintain interaction and enhance critical thinking in the writing class. This issue needs to be further investigated.

LIMITATION

Considering the limitations of the present study, more rigorous research can be taken into account to reveal the influence of students' multidimensional attributes on their preferences and acceptance of using Grammarly in writing classes. More qualitatively, future research can investigate teachers' processes and strategies in balancing their roles and those of Grammarly

in developing students' knowledge of writing and grammar as well as raising their writing awareness.

REFERENCES

- Azar, B. (2007). Grammar-based teaching: A practitioner's perspective. *TESL-EJ*, 11(2), 1–12.
- 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>.
- Caveleri, M., & Dianati, S. (2016). You want me to recheck your grammar? The usefulness of an online grammar checker as perceived by students. *Journal of Academic Language and Learning*, 10(1), 223–236. <http://www.journal.aall.org.au/index.php/jall/article/view/393/246>
- Chen, C. F. E., & 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. Retrieved from <https://www.researchgate.net/publication/45681611>
- Cotos, E. (2014). *Genre-based automated writing evaluation for L2 research writing: From design to evaluation and enhancement*. New York, NY: Palgrave Macmillan.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. New York, NY: Cambridge University Press.
- Daniels, P., & Leslie, D. (2013). @ CUE Grammar Software Ready for EFL Writers? 391–401
- Darayani, N. A., Karyuatry, L. L., & Rizqan, M. D. A. (2018). Grammarly as a tool to improve students' writing quality (Free online proofreader across the boundaries). *Jurnal Sains Sosial dan Humaniora*, 2(1), 83–89.
- Dodgson, A. N., Tariq, B., Alauyah, M., & Yusof, M. (2016). The secondary school students' usage of English learning websites to self-correct writing errors. *Asian TEFL*, 1(11), 2503–2569. <http://doi.org/10.21462/asianteftl.v1i1.3>
- El-Ebyary, K., & Windeatt, S. (2010). The impact of computer-based feedback on students' written work. *International Journal of English Studies*, 10(2), 121–142.
- Elliot, S. M., & Mikulas, C. (2004, April). The impact of MY access™ use on student writing performance: A technology overview and four studies. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.
- Ghufroon, M. A., & Rosyida, F. (2018). The role of Grammarly in assessing English as a foreign language (EFL) writing. *Lingua Cultura*, 12(4), 395–403.
- Grammarly. (2020). About Grammarly. Retrieved from <https://support.grammarly.com/hc/en-us/categories/115000018611-About-Grammarly>
- Isnawati, I., Sulisty, G. H., Widiati, U., & Suryati, N. (2019). Impacts of teacher-written corrective feedback with teacher-student conference on students' revision. *International Journal of Instruction*, 12(1), 669–684.
- Japos, G. (2013). Effectiveness of coaching interventions using Grammarly software and plagiarism detection software in reducing grammatical errors and plagiarism of undergraduate researches. *JPAIR Institutional Research*, 1(1), 97–109.
- Kern, R., & Warschauer, M. (2000). Introduction. In M. Warschauer & R. Kern (Eds.), *Network-based language teaching: Concepts and practice* (pp. 1–19). London: Cambridge University Press.
- Lee, I. (2020). Utility of focused/comprehensive written corrective feedback research for authentic L2 writing classrooms. *Journal of Second Language Writing*, 49, 1–7. <https://doi.org/10.1016/j.jslw.2020.100734>
- Liao, H. (2015). Using automated writing evaluation to reduce grammar errors in writing. *ELT Journal*, 70(3), 308–319.
- Lim, F. V., & Phua, J. (2019). Teaching writing with language feedback technology. *Computers and Composition*, 54, 102518. <https://doi.org/10.1016/j.compc.2019.102518>
- Li, J., Link, S., & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing*, 27, 1–18.
- Nova, M. (2018). Utilizing Grammarly in evaluating academic writing A narrative research on EFL students' experience. *Premise: Journal of English Education and Applied Linguistics*, 7(1), 80–96.
- O'Neill, R., & Russell, A. (2019). Stop! Grammar time: University students' perceptions of the automated feedback program Grammarly. *Australasian Journal of Educational Technology*, 35(1), 42–56. <https://doi.org/10.14742/ajet.3795>
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. Rowley, MA: Newbury House.
- Ranalli, J. (2018). Automated written corrective feedback: How well can students make use of it? *Computer Assisted Language Learning*. DOI: 10.1080/09588221.2018.1428994
- Rezaei, A.R., & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, 15, 18–39.
- Salteh, M. A., & Sadeghi, K. (2010). Teachers' and students' attitudes toward error correction in L2 writing. *The Journal of Asia TEFL*, 12(3), 1–31. <http://dx.doi.org/10.18823/asiateftl.2015.12.3.1.1>
- Shermis, M. D., Mzumara, H. R., Olson, J., & Harrington S. (2001). Online grading of student essays: PEG goes on the World Wide Web. *Assessment & Evaluation in Higher Education*, 26(3), 247–259.
- Stevenson, M., & Phakiti, A. (2019). Automated feedback and second language writing. In K. Hyland, & F. Hyland (Eds.). *Feedback in second language writing: Contexts and issues* (pp. 125–142). (2nd ed.). New York, NY: Cambridge University Press.
- Qassemzadeh, A., & Soleimani, H. (2016). The impact of feedback provision by Grammarly software and teachers on learning passive structures by Iranian EFL learners. *Theory and Practice in Language Studies*, 6(9), 1884. <http://doi.org/10.17507/tpls.0609.23>
- Ventayen, R. J. M., Orlanda-ventayen, C. C. (2018). Graduate students' perspective on the usability of Grammarly® in one ASEAN state university. *Asian ESP Journal*, 14(7.2), 9–30.
- 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.

- Ware, P. D., & Warschauer, M. (2006). Electronic feedback and second language writing. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Context and issues* (pp. 104–122). New York: Cambridge University Press.
- Wilson, J., & Andrada, G. N. (2016). Using automated feedback to improve writing quality: opportunities and challenges. In Rosen, Y., Ferrara, S., Mosharraf, M. (eds). (2016). *Handbook of Research on Technology Tools for Real-World Skill Development* (pp. 678-703). IGI Global: US. DOI: 10.4018/978-1-4666-9441-5.ch026
- Yang, Y.F. (2010). Students' reflection on online self-correction and peer review to improve writing. *Computers & Education*, 55, 1202–1210.
- Zhang, Z. (2017). Student Engagement with computer-generated feedback: A case study. *ELT Journal*, 71(3), 317–328.
- Zhang, Z., & Hyland, K. (2018). Student engagement with the teacher and automated feedback on L2 writing. *Assessing Writing*, 36, 90–102. <https://doi.org/10.1016/j.asw.2018.02.004>.
- Zheng, Y., & Yu, S. (2018). Student engagement with teacher-written corrective feedback in EFL writing: A case study of Chinese lower-proficiency students. *Assessing Writing*, 37, 13–24.