# CHAPTER II

## LITERATURE REVIEW

The purpose of this study is to examine Indonesian EFL students' preference in peer evaluation or automated writing evaluation to evaluate their English writing essay. This chapter discusses several areas in the literature including peer evaluation and automated writing evaluation as part of writing process, utilizing *Grammarly* as an automated writing evaluation tool, peer evaluation versus *Grammarly,* previous studies related to the students' preference on receiving feedback from peer evaluation and *Grammarly* and the theoretical framework on which this study is grounded.

### A. Peer Evaluation in EFL Writing Process

There are a few different names for what we call "peer evaluation," including "peer editing," "peer critiquing," and "peer feedback." Each classification represents a unique viewpoint on the input, most notably in regard to the position along the continuum at which this feedback is provided. Peer evaluation is defined by Hyland and Hyland (2006) as any verbal or written remarks provided by classmates or group members that proposes different benefits to incorporating peer input in any format. These comments can be made either verbally or in writing. It is intended to save teachers time, which will enable them to devote more of their attention to providing instruction that is of greater value. It is generally agreed upon that feedback from one's peers is more pertinent to the level of language progress achieved by the student. Students are able to improve their audience consciousness (i.e., their awareness of readers other than the teacher) when they have numerous readers. In addition, the reader can improve their own writing skills by reading and critiquing the work of others (Keh, 1990).

The theory of cognitive growth proposed by Vygotsky is put into effect through the use of peer review. The fundamental premise of this theory is that growth is a social process. According to this theory, knowledge is built through the interactions of individuals within society, and learning is the process of

internalizing social interaction. According to Vygotsky (1978), in order for a kid's growth to take place, it is necessary for the youngster to engage in social interaction with a more capable member of society from the very beginning of their lives. However, in order for such interaction to be successful, the aid that is provided by the member who is more capable (the expert) needs to be attuned to the requirements of the youngster (for example, see the research conducted by Ellis and Rogoff in 1986, McLane in 1987, Wertsch in 1979, and Wertsch and Hickmann in 1987). "Scaffolding" is a term that has been used to describe the graded and contingent character of the aid that is supplied by the expert (Wood, Bruner, & Ross, 1976). The most essential thing to remember about the metaphor of scaffolding is that it not only assists the novice in doing the work at hand, but it also supplies information that, if ingested by the youngster, makes it possible for the child to complete the task on their own (Greenfield, 1984).

While Vygotsky's primary focus was on the cognitive development of children, scholars such as Tharp and Gallimore (1988) and van Lier (1996) have argued that his theory can be extended to encompass all forms of learning. This includes both formal and informal instructional settings, as well as asymmetrical (i.e., expert-novice) and symmetrical (i.e., equal ability) groupings. Furthermore, it is suggested that Vygotsky's theory is applicable to individuals of all ages, including both children and adults. Indeed, several studies conducted in the field of mainstream education (e.g., Forman & Cazden, 1985) as well as L2 classrooms (e.g., Donato, 1988, 1994; Kowal & Swain, 1994; Ohta, 1995; Swain & Lapkin, 1998) have provided evidence that scaffolding can occur within the framework of peer interaction. Conversely, peer engagement may entail a unique form of scaffolding within its dynamics. For example, Ohta (1995) conducted a research and Kowal and Swain (1994) conducted another study, both of which found that the role of the expert in pair work might be dynamic, with learners alternating in assuming the expert role. Ohta (1995) demonstrated that the association yielded advantages for those with varying levels of expertise, including seasoned practitioners and those with limited experience. According to van Lier (1996), it has been shown that pupils have the ability to acquire knowledge through the act

of instructing others. According to Allwright (1984), second-language learners can improve the coherence and clarity of their own representation of second-language information by practicing teaching or explaining concepts to other people.

Some researchers have attempted to determine how scaffolding of peer assistance works by analyzing peer interactions. The scholars in question have directed their attention towards various elements, including the functions and purposes of the interactions, the perspectives adopted by the reader, and the dynamics within the group (Mangelsdorf & Schlumberger, 1992; Nelson & Murphy, 1992; Guerrero & Villamil, 1994; Lockhart & Ng, 1995; Villamil & Guerrero, 1996; Guerrero & Villamil, 2000). Several research have indicated that peer group dynamics, despite the increased concerns surrounding them, can really yield advantageous outcomes. Guerrero and Villamil (2000) conducted a study examining dyadic interactions involving the modification of a narrative text authored by one of the partners, serving as an illustrative example. The researchers made the finding that around 80 percent of interactions were characterized by collaboration. Additionally, they noted that successful pairs of peers demonstrated a concern for safeguarding each other's emotional well-being.

Villamil and Guerrero (2006) observe that collaborative interventions were the predominant form of activity in the examined sample. However, it is important to consider that the participants belonged to a culturally and linguistically homogeneous group of children, which may limit the generalizability of their experiences to other classroom contexts. Indeed, it has been recognized that peer responders who operate in their second language (L2) may exhibit deficiencies in the domains of communication and pragmatics, which are essential for achieving successful engagement. Furthermore, it is important to consider that these students may come from diverse cultural backgrounds, each with their own distinct expectations surrounding social interactions. This aspect can significantly influence the dynamics and outcomes of these interactions. The connection between culture and feedback, particularly in the context of peer feedback, has been the subject of a great deal of thought and debate in recent years.

When it comes to the organization of ESL peer response groups, cross-cultural concerns are considered to be one of the most significant factors to take into account (Carson & Nelson 1994, 1996; Nelson & Carson 1995; Nelson 1997). In multi-cultural collaborative peer response groups, 'conflict, or at the very least, high degrees of discomfort may arise,' as stated by Allaei and Connor (1990:24), and a number of research have investigated the likelihood of this happening. Nelson and Murphy (1992, 1993) conducted a study in which they found that students from diverse cultural backgrounds exhibited distinct expectations pertaining to factors such as the responsibilities of group members, the operational dynamics of the group, and various strategies of politeness. Nelson and Carson (1998) claim that individuals with a "collectivist" cultural orientation, as observed in countries like China and Japan, often prioritize the attainment of group consensus and actively avoid jeopardizing group cohesion by refraining from advocating alterations to their peers' written works. According to Nelson and Carson (2006), the presence of individuals who share the same language and cultural backgrounds within a group may facilitate the preservation of face. This is because such individuals are more likely to comprehend the subtleties embedded within each other's messages, thereby aiding in the maintenance of group harmony. The authors further contend that linguistic and cultural homogeneity can play a significant role in fostering successful interaction within peer response groups.

Furthermore, socioculturalism enhances the concept of social constructionism by asserting that reality is not solely a product of interpretive construction, but rather it is fundamentally shaped by collective and social forces, as well as being appropriated and modified by individuals (Ortega, 2009). From this standpoint, the individual engages in an ongoing process of social contact and connection with others. According to Wenger (1998, p. 102), individuals within communities of practice engage in mutual interaction, collaborative activities, the construction of shared understandings, and the acquisition of knowledge from one another. Lave and Wenger (1991) believe that the process of learning, as conceptualized through increased participation in communities of practice,

encompasses the holistic engagement of individuals in their respective environments. This approach is rooted in the social-cultural framework. In accordance with van Lier (2000), it is imperative to situate language within various semiotic systems and within the broader scope of the world. The execution of this task is crucial for establishing a connection between cognitive processes and social actions.

Researchers and educators of second languages have shown a growing interest in comprehending the process of second-language development within educational settings, such as classrooms, through the examination of language, interaction, and artifacts (Ohta, 2000). The reason for this phenomenon is that the process of learning is influenced by the use of language (Ohta, 2000; Swain, 2000; Wells, 1999, 2000). Numbers studies have been conducted to investigate whether or not language or conversation has any bearing on the process of learning. According to Wells (2000), who references Franklin (1996), knowledge is formed and re-created as a result of the conversation that occurs between individuals who engage in activities together. Collaboration in the classroom of a foreign language allows for the construction of knowledge in both the content domain and the language itself. Learners are able to develop alternate messages as they work to complete a job because it forces them to concentrate their attention on what they are saying and gives them the opportunity to do so. As a direct consequence of this conversation, the students' collectively created performance exceeds the sum of their individual competences (Swain, 2000). The literature also acknowledges the necessity to classify the discourse in question as a combination of internal and external conversations that a student engages in with multiple voices, including their own (Weisberg, 2008). This matter has been brought to the attention of individuals.

## B. Automated Writing Evaluation in EFL Writing Process

Activity Theory (AT) offers a framework for examining the impact of socio-historical contexts on the utilization of learning tools by second language (L2) learners. This framework also considers the role of socially organized and object-directed resources in mediating these activities (Engestrom, 2001; Lantolf

& Thorne, 2006). Automated Writing Evaluation puts this theory into practice, making it an essential component of sociocultural theories. The concept of appropriation is fundamental to the field of applied linguistics. Grossman et al. (1999) propose that appropriation refers to the cognitive process through which individuals assimilate and accept the pedagogical tools that are accessible within specific social contexts, such as educational institutions. This process of appropriation facilitates the internalization of cognitive frameworks and patterns of thought that are intrinsic to distinct cultural practices.

Hence, appropriation can be regarded as a prototypical mechanism of internalization, denoting a transformative process wherein inter-psychological functioning, manifested through social interactions involving socially constructed artifacts, is internalized and undergoes a metamorphosis into intra-psychological functioning (Vygotsky, 1978). Appropriation is a phenomenon characterized by the act of an individual acquiring and assimilating something that is not originally their possession, thereby asserting ownership and making it an integral part of their identity or creative expression. The process of internalization involves various sub-processes, including imitation, observation, goal-setting, selection, evaluation, and adaptation. These sub-processes are employed to transform externally mediated materials into psychological artifacts that facilitate higher mental activities, such as writing and language learning (Lei, 2008). Internalization refers to the cognitive process by which externally provided resources are transformed into psychological artifacts that play a role in facilitating higher mental functions.

Within the context of goal-directed processes, Vygotsky's notion of mediation assumes a central role, with particular emphasis placed on the significance of artifacts, such as physical instruments and symbolic signals, as essential means for learners to facilitate their appropriation. These objects can be seen as a fusion of tangible instruments and symbolic representations. Engestrom (2001) proposes that the mediating resources within his Activity Theory framework are characterized by their object-oriented nature. These resources have the potential to be expanded upon and transformed into four interconnected

mediators. The elements under consideration are the artifact, the community (a collective of individuals with a common aim), the rules (prescribed regulations and criteria), and the division of labor (allocation of social roles). Previous studies in the field of sociocultural research have examined the many forms of mediating resources and strategies employed by students in the context of general writing (Lei, 2008) or peer feedback (Yu & Lee, 2016). The existing literature on this topic has identified several mediating resources, including the internet, languages, and writing criteria. However, there is a lack of comprehensive analysis regarding the differences in the utilization and internalization of these resources among students with different levels of English proficiency, specifically in relation to their incorporation of automated feedback in second language (L2) writing.

According to Grossman et al. (1999), appropriation can take place on any one of five distinct levels, each of which is affected by a unique mediator and guided by a distinct set of internalization processes and objects. Lack of appropriation is the first level, and it is the lowest level. This occurs when an individual either does not use the mediational tools or refuses to use them for various reasons, such as not having an adequate conceptual grounding. The second step is called "appropriating a label," and it involves simply learning the name of a tool without any knowledge of the instrument's functions. When a person learns some features of a tool but is unable to grasp the whole concept and how it might be applied, they have moved on to the next level, which is called appropriating superficial features. A person has reached the fourth stage when they have appropriated conceptual understandings, which indicates that they have an understanding of the theoretical underpinning that informs and justifies the usage of a tool. The very last step is to reach the degree of mastery, which is the greatest level and requires not only complete knowledge of the tool but also the capacity to use it effectively in actual situations.

These five levels imply that appropriation is a process that evolves over time and, as a result, provide a solid foundation for comprehending the myriad ways in which pupils make use of automated input. According to Grossman et al. (1999) and Yuan (2017), the five stages imply that the success of appropriation is

typically dependent upon the preceding experiences, objects/goals, skills, and particular social situations within which the appropriation occurs. In recent years, the concept of appropriation has been shown to be a useful paradigm for examining either instructors' use of curriculum standards (Yuan, 2017) or students' use of technology in education (Laffey, 2004). This has been the case both in the classroom and in research settings. In light of the fact that the usage of AWE tools is becoming increasingly prevalent in L2 settings (Li et al., 2019), it is very necessary to broaden the scope of the discussion to include the application of AWE by student writers from the viewpoints of appropriation and AT.

The Automated Writing Evaluation (AWE) system offers prompt and personalized feedback to improve the writing abilities of students (Li, 2021; Link et al., 2020). Automatic Writing Evaluation (AWE) is a comprehensive software that integrates various components to enhance language learning. It incorporates Automatic Written Corrective Feedback (AWCF) to facilitate error identification, an Automated Essay Scoring (AES) algorithm to assess writing proficiency, a management system to offer learners multiple opportunities for drafting, and a repository of writing resources including a dictionary, thesaurus, Writers' Handbook, and other self-access materials (Chen & Cheng, 2008; Stevenson & Phakiti, 2014). Several studies have examined the correlation between automated and human grading of student writing (Bridgeman et al., 2012). Additionally, the effects of automated analysis on student writing performance have been investigated by Jiang et al. (2020) and Ranalli (2021). Bridgeman et al. (2012) conducted a study to examine the relationship between automated and human grading of student writing. One of the key advantages of Automated Writing Evaluation (AWE) lies in its efficacy, since it facilitates prompt feedback provision. The acceleration of the practice-feedback loop is crucial in fostering the development of metacognitive abilities in the realm of writing (Kellogg et al., 2010).

The assessment of essays is presently facilitated by several Automated Writing Evaluation (AWE) systems, like My Access, Criterion, and Grammarly, which are extensively employed by English writers. In the study conducted by Li

(2021), Criterion error reports were utilized to assess the quantity of errors identified and classified by Criterion. The findings of the study revealed a notable enhancement in linguistic precision from the initial to the last iterations of the document. In a study conducted by Lai (2010), an examination was undertaken to assess the extent to which the utilization of MY Access improved the writing proficiency of college students who are English as a Foreign Language (EFL) learners. The AWE technique received a predominantly positive response from the majority of students, who expressed their appreciation for the prompt feedback provided. However, certain portions of the input were perceived by them as predetermined, reiterated, and open to several interpretations. According to Barrot (2020), Grammarly is a highly effective tool for enhancing writing proficiency, particularly in the areas of revision and editing.

### C. Utilizing *Grammarly* as an Automated Writing Evaluation Tool

Grammarly is a digital writing tool that may provide automated essay feedback by finding problems in grammar, vocabulary, mechanics, and style. It does this by analyzing the text to determine where corrections need to be made. It works on computers running Mac OS X, Windows, Android, and iOS and can be accessed using well-known web browsers including Chrome, Safari, and Firefox (Barrot, 2020). It has recently undergone a revamp to make its user interface accessible as a web application, as an extension to a web browser, as an extension to Microsoft Word, and as a native desktop application. All of these options are currently available. Grammarly requires users to register with the website at https://www.grammarly.com/signup before they may create an account. Once the user's account has been created, Grammarly will send them to a clean and user-friendly dashboard that operates in accordance with how English texts are read and aggregates associated metrics. This dashboard runs according to how English texts are read. The user profile and any other applications that have been connected with it are displayed on this dashboard (see Figure 1).
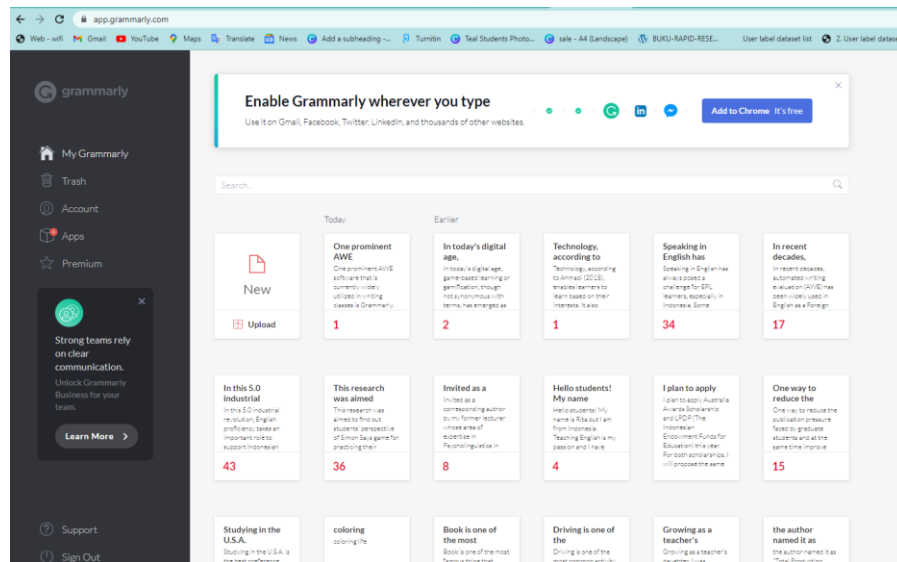
Figure 1. *Grammarly's* Dashboard

Grammarly is beneficial for writing lessons, particularly during the stages of revising and editing the work that students have produced. During the stage where revisions are being made, it can be utilized to get rid of any duplicate text and to acknowledge sources. The Plagiarism Detection tool of Grammarly highlights specific passages of the text that have been plagiarized, gives users with the reference information they need to correctly acknowledge the original source, and provides an overall score for the originality of the writing. Because it just displays the percentage of text matches based on the number of similar fragments found in other sources, the originality score report that Grammarly provides does not require any specialized knowledge or abilities to understand. Take note that this functionality is not included in any free versions of the software.

To this day, it has the ability to identify duplicate content on more than 16 billion online pages and academic documents. Given these affordances, it is possible that students will be forced to turn in their amended papers together with the Grammarly plagiarism report before they can move on to receiving feedback from their peers and teachers. Students who use Grammarly's Check My Writing feature are much less likely to plagiarize, whether on purpose or by accident. The tool that checks for plagiarism can be helpful for writers in any sector, despite the fact that it was initially developed with students in mind. The editing affordances

provided by Grammarly can be utilized throughout the final stage of the writing process, which is referred to as the editing stage. As a tool for editing language, it offers real-time feedback by highlighting faults and incorrect linguistic usage that can be detected in the text (see Figure 2 for an example of this).
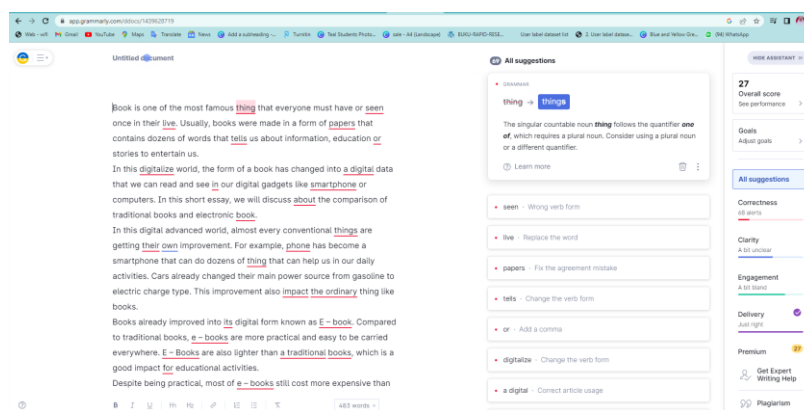


Figure 2. Editing Textbox

The use of red underlining to indicate a spelling, grammatical, or punctuation fault is contrasted with the use of blue underlining to indicate concerns with conciseness and clarity. The tone, formality, and politeness of the phrase are shown by the purple underlining, while the ideas that can make the remark more engaging are indicated by the green underlining. This writing aid is driven by artificial intelligence, and one of its most notable features is that it provides corrections along with their associated explanations. It also provides the total text score, which ranges from one to one hundred and is based on the various suggestions that are made throughout the page. The better the score of the document, the less suggestions there are to improve it. Students can use a plagiarism checker to further delete any duplicate information that they could have used throughout the revision step, even if the primary focus of the final stage is on refining the language.

The utilization of Grammarly as a tool for language acquisition was found to be supported by the most recent empirical findings. For example, Koltovskaia (2020) discovered that using Grammarly encouraged students' utilization of their cognitive and metacognitive functions through the practice of observing. She went on to explain that it may be utilized as a helpful resource for writing, particularly

in situations when pupils were actively involved. In a similar vein, O'Neill and Russell (2019) stated that Grammarly was effective in providing grammatical guidance in a range of learning scenarios, regardless of whether the students were local or international and regardless of whether the instruction was delivered online or in-person.Moreover, a growing body of empirical research has provided evidence of the clear utility of this tool in facilitating students and scholars to write with exceptional precision and efficiency (Gao & Ma, 2020; Guo et al., 2021; Hassanzadeh & Fotoohnejad, 2021; Jiang et al., 2020; Li, 2021; Link et al., 2020; Waer, 2021). According to Waer (2021), Barrot (2020) asserts that the utilization of Grammarly is highly advantageous in the context of writing pedagogy, namely in the stages of revision and editing within the writing process. By utilizing this tool to detect instances of plagiarism during the revision phase of the writing process, students can effectively eliminate duplicate content and appropriately attribute sources.

In spite of the adaptable affordances it provides, I discovered some flaws in Grammarly that demand immediate attention from the company. To begin, it will often make suggestions for improvements that could cause the assertion to be false. For instance, it produces false positives when it attempts to correct proper nouns like "Dayz Hotel" or technical or unusual phrases like "generalizability" and "ebook." When it comes to identifying instances of possible plagiarism, Grammarly has a tendency to highlight certain common phrases as items that could be plagiarized. 'Therefore, it is vital to observe the following...' and 'I agree with what you said regarding the matter' are two instances of phrases that illustrate this point. The users will need to perform additional filtering to remove these unneeded corrections.

## D. Research on Students' Perception of Using Peer Evaluation and Automated Writing Evaluation

According to the study conducted by Wang and Brown (2008), there are certain similarities between human and machine marking. Nevertheless, disparities between people and machines persisted. The capacity of even the most sophisticated computerized essay grading algorithms to overlook crucial inherent

characteristics is a possibility. The topic of physical education and its documentation in English writing for ESL and EFL students in various nations has received significant attention in the literature on feedback. In a study conducted by Levi Altstaedter (2018), the impact of peer feedback on the writing skills of English as a Foreign Language learners was examined. The findings revealed a significant enhancement in the overall quality of the students' final drafts subsequent to the receipt of feedback from their classmates. Huisman et al. (2019) conducted a study examining the influence of peer feedback on the academic writing skills of Australian students. The findings of their research indicated that students showed much greater improvement in their writing abilities when they engaged in either giving or receiving peer criticism, as opposed to not participating in these activities. In addition, Su and Huang (2021) discovered that the students in China regarded physical education (PE) to be enjoyable, particularly when they were given the task of providing feedback to their classmates.

Some scholars in Indonesia are looking into the success of physical education from the students' point of view as well as the students' writing performance in order to determine whether or not it is effective. For example, Fithriani (2019) conducted research with college English language education students to determine the advantages of providing written comments. According to the findings, having students get written comments from their peers not only helps them improve the quality and abilities of their writing, but it also stimulates critical reasoning and promotes learner autonomy. According to the findings of Dewi (2019), English Education students in Indonesia have a tendency to provide constructive comments as their feedback while evaluating the essays of their classmates. This research demonstrates that kids benefit in favorable ways from participating in physical education. In addition, Hentasmaka and Cahyono (2021) looked into the uptakes of peer feedback and the effects of that input based on the level of proficiency. They came to the conclusion that students' writing could be significantly improved by the use of peer feedback, independent of the students' current levels of ability, and that there was no significant difference in the number

of replies or outcomes. Iswandari and Jiang (2020) reviewed 16 different empirical studies on peer feedback in college EFL in order to determine how peer feedback has been explored throughout the course of the last ten years. As a consequence of this, they came to the conclusion that the researchers examined four distinct sorts of feedback criteria, with the majority of the study concentrating on the students' points of view about peer feedback.

Recent studies on Grammarly have emerged from various researchers (Barrot, 2020; Gao & Ma, 2020; Guo et al., 2021; Jiang et al., 2020; Link et al., 2020; Waer, 2021; Ariyanto et al., 2021; Fahmi & Cahyono, 2021; Ghufron, 2019; Karlina Ambarwati, 2021; Miranty et al., 2022). These studies have had a growing impact on the field of second language (L2) writing. For example, Barrot (2020) investigated how Grammarly may be utilized in the writing of ESL and EFL students. It provides effective grammatical help in a range of learning environments, whether they involve international or domestic students or are used online or face-to-face. It does this by promoting the use of students' cognitive and metacognitive operations through the process of noticing, and it does this by providing effective support. In order to evaluate the usefulness of Grammarly for English as a Foreign Language (EFL) authors, Guo et al. (2021) conducted an experiment with 36 students attending a top Chinese institution. They observed that when the students used the input from Grammarly to review their essay, there was a considerable reduction in the number of grammatical errors made by the students. In a similar vein, Fahmi and Cahyono (2021) investigated the viewpoints of 26 undergraduate students on the feedback provided by Grammarly and their teacher. The findings revealed that students with either poor or high English proficiency give good attitudes toward this form of feedback.

In English as a Foreign Language (EFL) composition classes, the usefulness of peer response groups and automated writing answers has been generally disregarded, particularly in Indonesia. This is the case despite the limits of earlier publications on the topic. To the best of our knowledge, the first study to compare AWE and peer evaluation as methods for grading students' essays was conducted by Lai (2010), who looked at the writing of 22 EFL students from

Taiwan. The researcher looked at it from three different angles: the product, the method, and the perspectives of the students. The findings of recent research conducted by Shang, (2022) indicate that pedagogical EFL writing implications can be explored in greater detail if peer evaluation and online peer feedback are compared and contrasted. As a result, there is a need for further elucidation regarding these two sorts of feedback that are encountered by students in Indonesia.