Hindawi

Research Article

AI in the Foreign Language Classroom: A Pedagogical Overview of Automated Writing Assistance Tools

Wael Alharbi ⓘD

Yanbu English Language and Preparatory Year Institute, Yanbu Education Division,
Royal Commission for Jubail and Yanbu, Yanbu, Saudi Arabia

Correspondence should be addressed to Wael Alharbi; alharbiwa@rcyci.edu.sa

Recent technological advances in artificial intelligence (AI) have paved the way for improved and in many cases the creation of entirely new and innovative, electronic writing tools. These writing support systems assist during and after the writing process making them indispensable to many writers in general and to students in particular who can get human-like sentence completion suggestions and text generation. Although the wide adoption of these tools by students has been faced with a steady growth of scientific publications in the field, the results of these studies are often contradictory and their validity may be questioned. To gain a deeper understanding of the validity of AI-powered writing assistance tools, we conducted a systematic review of the recent empirical AI-powered writing assistance studies. The purpose of this review is twofold. First, we wanted to explore the recent scholarly publications that evaluated the use of AI-powered writing assistance tools in the classroom in terms of their types, uses, limits, and potential for improving students' writing skills. Second, the review also sought to explore the perceptions of educators and researchers about learners' use of AI-powered writing tools and review their recommendations on how to best ingrate these tools into the contemporary and future classroom. Using the Scopus research database, a total of 104 peer-reviewed papers were identified and analyzed. The findings indicate that students are increasingly using a variety of AI-powered writing assistance tools for improving their writing. The tools they are using can be categorized into four main groups: (1) automated writing evaluation tools, (2) tools that provide automated writing corrective feedback, (3) AI-powered machine translators, and (4) GPT-3 automatic text generators. The analysis also highlighted the scholars' recommendations regarding dealing with learners' use of AI-powered writing assistance tools and grouped the recommendations into two groups for researchers and educators.

1. Introduction

The recent developments in technology in general and in artificial intelligence (AI) in particular have impacted every aspect of our life including education. These advances in AI technology have had a profound impact on language learning and teaching by changing the way that we produce or perceive language. In this study, we discuss how AI technology has been disruptive to the ways writing is produced, taught, learned, evaluated, and edited. Authoring tools such as automated writing evaluation (AWE) or automated essay scoring, which were originally designed to assist writing teachers in assessing their students' assignments, have completely changed with AI technology where they shifted from the conventional checking of grammar and spelling to offering extensive

support in identifying writing problems and offering suggestions for improving the writing quality. Over the past few years, corrective feedback (CF) has become synchronous and immediate either as part of the available cloud-based word processor suites or as standalone apps or software suites making it possible to produce more accurate writing [1–3]. According to Dale and Viethen [4], the greatest development that AI has brought to writing was the AI-based sentence and phrase autocompletion and alternative wording suggestion features. All these advances have been possible and will continue to develop, thanks to AI applications and systems which collect large sets of data and then process them by utilizing artificial neural networks and machine learning technologies. All have resulted in momentous improvements and breakthroughs in turning texts into structured data and extracting

meaning from them by utilizing AI-based natural language processing (NLP) and natural language understanding (NLU).

Despite the widespread use of digital authoring and writing tools in different day-to-today and professional environments, incorporating these tools into the language classroom has been controversial. As AI-powered digital writing assistance goes beyond vocabulary and grammar to more sophisticated and "human-like" help, then language educators and researchers may have reservations about the authenticity of students' submitted writing. Such concerns are legitimate since these intelligent tools provide writers with near-human translations, rephrased sentences, and large chunks of text at a click of a button allowing learners to copy and paste the intelligently authored suggestions into their written work with little or no learning taking place.

While many language professionals do not mind the presence of some AI-powered language-proofing tools in their classrooms (such as AWE tools), they hold a very strong stance against the use of machine translation (MT). They make this distinction between AWE tools based on the depth and breadth of linguistic help students get from these tools assuming that MT tools offer more language output and require no or minimum cognitive processing. Although the distinction between these tools is clear, it is not that simple to separate the functions and be selective. According to Dale and Viethen [4], AI-powered writing assistance systems are typically built on massive linguistic models, which promote and offer a whole range of language assistance services as a package starting from MT to sentence and text generation. Eaton et al. [5] argued that what makes these generated texts unique and worrisome at the same time is the fact that they are very difficult to be detected by antiplagiarism tools.

No matter what language educators think or feel about students' use of AI-powered writing assistance tools for doing their writing tasks, it might be time to take a more realistic approach to the issue and treat it as an inescapable fact that they need to accept and live with. Instead of banning these tools in the classroom and discouraging consulting them at home to do their writing assignments without having any control or system in place to monitor students' use of external resources, educators may alternatively have to try to find ways to utilize these tools into the classroom in a way that helps students learn by providing appropriate guidance [6–8].

It is indisputable that these digital tools have a range of strengths and weaknesses which all can be discovered and explored by students using them in exploratory environments mediated by experienced and knowledgeable teachers [9]. Undoubtedly, AI-powered writing assistance tools have a great potential in enhancing the teaching and learning processes in the language classroom. However, to unlock their potentials, the impact of these tools on the learning process should be critically analyzed. Moreover, understanding the limitations of these tools in understanding the pragmatic and contextual complexity of human language can help us gain the linguistic insights needed for the right integration in the writing classroom [10]. A complex and informative learning environment can be created, and hence be broadly understood, by allowing students to interact with the AI-powered tools and the interaction of software, which all are mediated through the teacher. The scrutiny of the interaction of this mix will help researchers understand it from a broader and ecological perspective, which is missing or scarcely investigated in research studies [7]. Just as any technology integrated into the classroom, AI-powered writing assistance tools can play an important role in transforming the students' learning process and enhancing their writing skills. However, these tools need to support their learning experience [8]. A more thoughtful approach that considers the ecology of implementation is probably the best option for educators. While coexistence with these tools sets the tone for smoother implementation, it is still not well established in many educational settings how this ecological perspective toward AI-powered writing assistance tools should break the ice and forge links between people, technology, and organizations [9].

As the world has been recently experiencing an unprecedented boom in AI-powered technologies that have become easily accessible and available to learners around the globe, our understanding of the teaching and learning processes is being challenged everyday. Although researchers and education practitioners raced to test these technologies to measure their impact on the instructional environment, the knowledge gap between what we know and the learners' actual use of these technologies is widening as students are consulting these tools outside the classroom and without the consent of their teachers. The relationship between the increasing number of AI-powered writing assistance tools that students use and the educators' awareness of these tools is noncorrelational as some educators are not as technology savvy as their students, which sometimes results in passing students who do not deserve to pass. In the literature, the views on the use of AI-powered writing assistance tools are mixed with some researchers to see its use as a form of cheating and academic dishonesty while others find a great potential in them as contributors to language learning knowledge and as text improvers. Integrating AI technology into most of educational systems and applications is relatively new and its impact on the learning process is yet to be empirically verified. The available literature that investigated these tools is either revealing conflicting results or treating each writing assistance tool individually. With students increasingly using these tools to generate or improve their L2 texts and assessed work, educators need to know about these tools and their strengths and weaknesses. They also need to be informed about the best ways to deal with this new reality and whether or not they need to change or update the way they teach and assess their students.

Currently, there is a lack of comprehensive reviews on the available AI-powered writing assistance tools and their pedagogical implications. Existing reviews that are related to the use of writing assistance tools have focused either on the use of individual tools in their early versions prior to AI integration, or on a specific type of writing assistance tools. Thus, an in-depth overview is needed on recently developed AI-powered writing assistance tools, including information on their types, strengths, weaknesses, their impact on students writing quality and the researchers' recommendations regarding the use of these tools as in the classroom. The

TABLE 1: Initial search items/strings.

Topic	Search items/strings
AI-powered writing assistance technologies/tools	"automated writing" or "automated writing scoring" or "artificial intelligence" or "AI" or "artificial intelligent writing assistance tools" or "artificial intelligent writing assistance technologies" "Google Translate" or "Grammarly" "machine translation" or "artificial intelligent writing tools"/"Artificial Intelligence-powered writing systems/AI-powered writing tools/AI-powered writing assistance tools/Automated Writing Corrective Feedback Tools/Automatic Text Generation and Deep Learning Technology" or "text editors" or "synchronous feedback"
Educational level	"higher education" or "higher ED" or "university" or "college" or "undergraduate" or "graduate" or "postsecondary" or "post-secondary" or "tertiary"
Learning Setting	"student" or "learn" or "learner" or "classroom" or "EFL classroom" or "ESL classroom" or "L2 learner"

purpose of this study is to explore the recent scholarly publications that evaluated the use of AI-powered writing assistance tools in the classroom by shedding light on these tools in terms of their types, uses, limits, and the potential for improving students' writing skills. This study also seeks to explore the perceptions of educators and researchers about learners' use of AI-powered writing tools and review their recommendations on how to best ingrate these tools into the contemporary and future classroom. To guide our inquiry and selection of research articles, we formulated the following three research questions:

(RQ1) What state-of-the-art AI-powered writing assistance technologies are in use by students and teachers in tertiary education, and what are they used for?

(RQ2) What are the strengths and limitations of these technologies, and how do they impact students' writing?

(RQ3) How do researchers and higher education practitioners view the use of AI-powered writing assistance technologies, and what are their recommendations?

With this information generated from the present systematic review, educators may gain a deeper awareness of available AI-powered tools which will enable them to facilitate the use of these tools effectively and appropriately. In the next section, we describe our methodological approach, the research questions, and the systematic review guidelines. Then, we present our findings based on our analysis of the relevant literature. Finally, we conclude by discussing recommendations for educators and plans for future research.

2. Methodology

To find relevant literature for this study, the Scopus research database was used for the literature search. The reason why Scopus was chosen over other scholarly databases was that it is considered the largest and most comprehensive database for peer-reviewed abstract and citation literature [11]. Based on relevant AI-powered writing assistance literature [12, 13], we identified a number of search keywords (Table 1).

The Scopus search produced a preliminary unfiltered dataset of 379 research papers (last retrieved September 2022). To make sure that all the retrieved studies were relevant to the research questions of this study, a further filtering process was conducted. The filtering process was based on three inclusion and three exclusion criteria. The first inclusion criterion was that studies were supposed to be based on the empirical methodology for data collection and are published in journals that are peer-reviewed [2]. The second criterion was that the type of writing assistance tools under question in those studies should be dependent on AI as a backbone of their operation. Therefore, studies that investigated writing assistance tools and those tools were not based on AI or machine learning were excluded.

To narrow down the collection of studies and only include the most recent ones, the third criterion was only papers published between 2017 and 2022 which included in the selection. This process was important for the validity of this study for three reasons. First, the use of technology in the classroom has significantly changed in response to COVID-19 when students used a variety of writing assistance tools to study and/or do their assignments when classes were suspended for almost 2 years. Second, technology evolves and develops very rapidly, so to give a clear description of the current situation and make the right predictions in regard to the use of AI-powered writing assistance in the classroom, focus should be on the most recent studies. Third, Google Translate, which is one of the most highly consulted writing assistance tools by students [3–5], started using AI (or what is known as neural machine translation) in its system in the year 2017. Hence, any studies published before 2017 would be irrelevant and consequently were excluded from the dataset. The screening process of the selected studies was done by following the guidelines of the preferred reporting items for systematic reviews and meta-analyses (PRISMA) flowchart [6]. Figure 1 shows the screening process started with 379 studies and ended with 104 papers.

The literature screening process was carried out in two steps. The initial step started with reading titles and abstracts to verify face eligibility. The studies that passed the first step were then read fully for review and analysis. After the review process, two main themes were identified. From each theme, several subthemes emerged. Table 2 shows the screening process of the papers yielded two main themes and six subthemes.

Theme (1): Current and emerging AI-powered writing assistance technologies. This theme has four subthemes:
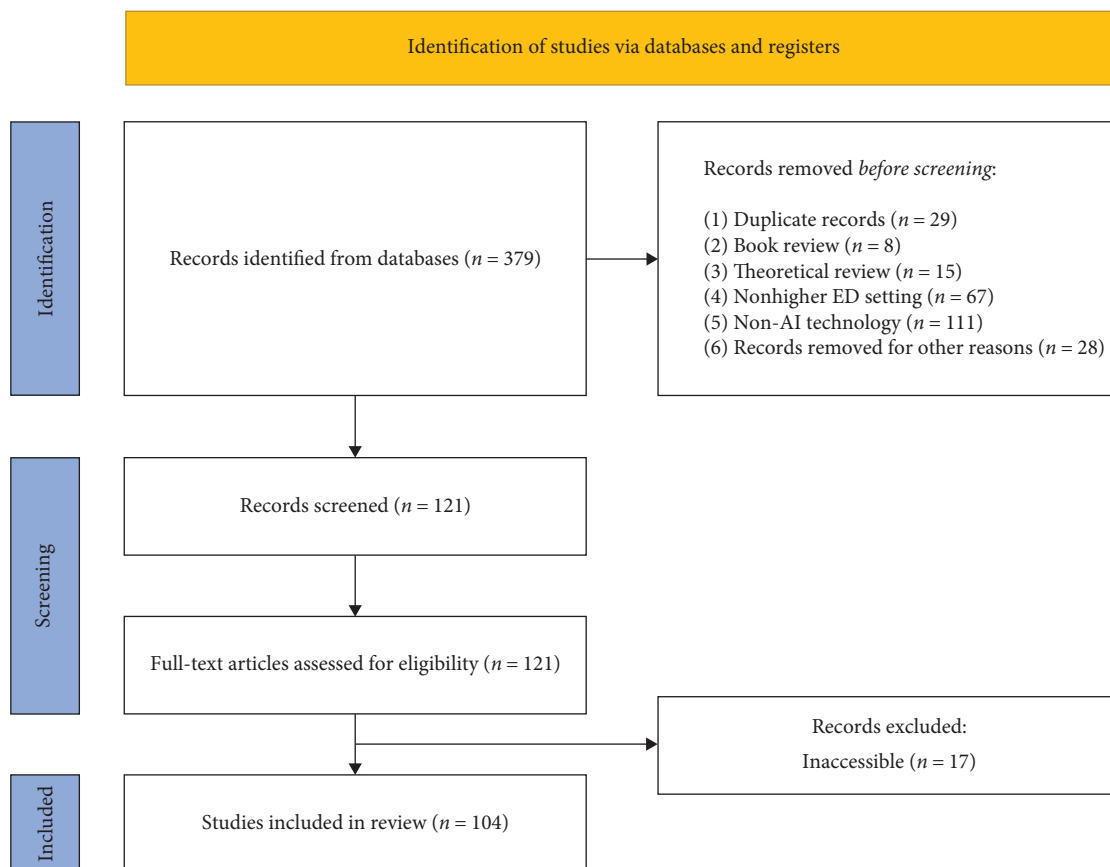
(a) Automated writing evaluation (AWE)

FIGURE 1: PRISMA flowchart for the selection process of the articles.

(b) Automated writing corrective feedback (AWCF)

(c) AI-enabled machine translation

(d) Automatic text generation (GPT-3)

Theme (2): Recommendations by scholars for researchers and educators on how to deal with students' use of AI-powered writing assistance tools. This theme has two subthemes:

(a) Classroom integration (coexistence with the tools)

(b) Adopting ecological perspectives toward these tools

In the following section, each theme and its subthemes are discussed against the study's research questions.

3. Results and Discussion

Three research questions guide the inquiry of this review study. The answers to the first and second research questions are discussed under Theme (1), which gives an overview of the current and emerging AI-powered writing assistance technologies that are in use in instructed learning environments in terms of what they are used for, their strengths and weaknesses and how they impact student's writing.

3.1. Theme (1): Current and Emerging AI-Powered Writing Assistance Technologies. AWE tools, such as Criterion, MY Access!, or WriteToLearn, have been incorporated in some educational settings for sometime now which is promising in terms of the availability of a body of research about them. AI-powered synchronous text editors are more recent than asynchronous ones. Examples of those include Grammarly, ProWritingAid, and Writing Mentor applications which have been gaining popularity in educational, professional, and personal settings. These tools intelligently provide users with automated written corrective feedback (AWCF). According to Ranalli and Yamashita [3], AWCF has been used as a descriptor in emerging research that is investigating and exploring the use of these tools. Over the past few years, instant online translators such as Google Translate have come a long way and become accessible on a variety of devices and in different formats, thanks to the huge leaps in mobile technology and AI technology. The latest addition and improvement to intelligent writing assistance tools are systems that can generate texts instantaneously and autonomously with a single prompt. Regardless of their grammatical accuracy, these text generators, such as Google Compose, can offer linguistically acceptable, and sometimes human-like, word choice suggestions and improvements. More sophisticated systems such as GPT-3 go further and suggest complete texts that need only a topic or prompt to operate. In the following subsections, we will shed light on each type of writing assistance system.

3.1.1. Electronic Feedback through Automated Writing Evaluation Systems. AWE systems are now broadly used in both first- and second-language teaching contexts and at all

TABLE 2: The themes of the reviewed papers and the studies related to each theme.

Main themes	Subthemes	The reviewed studies		
Current and emerging AI-powered writing assistance technologies	Automated writing evaluation	Burstein et al. [14]; Ortega [15]; Bridgeman and Ramineni [16]; Ranalli et al. [17]; Godwin-Jones [18]; Hibert [19]	Hussein et al. [20]; Warschauer et al. [21]; Saricaoglu [22]; Link et al. [23]; Peng et al. [24]; Palermo and Wilson [25]	Nunes et al. [26]; Camacho et al. [27]; Godwin-Jones [28]; Huang and Wilson [29]; Ranalli [30]
	Automated writing corrective feedback (AWCF) tools (text editors supplying synchronous feedback)	Ellis [31]; Dembsey [32]; Arroyo and Yilmaz [33]; Zheng and Yu [34]; Ranalli [1]	Nova [35]; Conijn et al. [36]; O'Neill and Russell [37]; Ghufron [38]; Huang et al. [39]; John and Woll [40]; O'Neill and Russell [41]; Dale and Viethen [4]	Dodigovic and Tovmasyan [42]; Zomer and Frankenberg-Garcia [43]; Dizon and Gayed [44]; Ranalli and Yamashita [3]
	AI-enabled machine translation	O'Neill [45]; Lewis-Kraus [46]; Ellis [31]; Godwin-Jones [47]; Crossley [48]; O'Neill [45]; Fredholm [49]	Godwin-Jones [50]; Hussein et al. [20]; Enriquez Raiído and Saínchez-Torrón [51]; Lee [52]; Vinall and Hellmich [53]; Hellmich and Vinall [7]; Dale and Viethen [4]	Urlaub and Dessein [54]; Zhang and Torres-Hostench [55]; Klekovkina and Denié-Higney [56]; Jolley and Maimone [57]; Ryu et al. [58]; Vinall and Hellmich [10]; Pellet and Myers [9]
	Automatic text generation	Ruder [59]; Dale [60]; Floridi and Chiriatti [61]; Ferrone and Zanzotto [62]; Dale [63]	Dale and Viethen [4]; Dizon and Gayed [44]; Eaton et al. [5]; Godwin-Jones [47]; Zhang and Li [64]	Schmalz and Brutti [65]; Anson [66]; Anson and Straume [67]
Recommendations by scholars for researchers and educators	Classroom integration (Coexistence)	Jiang et al. [68]; John and Woll [40]; Koltovskaia [2]; Woodworth and Barkaoui [69]; Hellmich and Vinall [7]	Hellmich and Vinall [7]; Kessler [70]; Ling et al. [71]; Ranalli [30]; Li [72]	Li [72]; Knowles [73]; Pellet and Myers [9]; Sumakul et al. [74]; Fyfe [75]
	Adopting ecological perspectives	Larsen-Freeman [76]; Zhang and Hyland [77]; Hockly [78]; Patout and Cordy [79]; Jiang et al. [68]; Koltovskaia [2]; Link et al. [23]	Lee [52]; Woodworth and Barkaoui [69]; Zhang [80]; Hellmich and Vinall [7]; Nunes et al. [26]; Ranalli [30]	Huang and Wilson [29]; Li [72]; Lieshout and Cardoso [81]; Pellet and Myers [9]; Sun and Fan [82]

education levels, from elementary school to university. Writing is considered a complex process that combines low-level skills such as mechanics, spelling, and higher-level skills pertaining to logical sequence, organization of content, and stylistic register appropriateness. Second-language writing is inherently difficult as it poses its own unique set of challenges pertaining to potential deficiencies and gaps in syntactic, pragmatic, lexical, and/or rhetorical knowledge. As far as giving useful CF to writers is concerned, it is consequently a difficult, demanding, and arduous task for many L2 teachers. How to give useful CF to students has always been a controversial topic in second-language writing research [83]. Despite a few researchers who believe otherwise, it seems that there is a consensus that CF may be very useful when it is proven and used properly [23, 31, 84]. However, it is not easy to make broad generalizations about the usefulness of CF for students as there are several contextual variables at play [31, 85].

Teachers often find it extremely time-consuming and tedious to provide feedback on student writing. Depending on class size, providing students with individual feedback that is tailored to their needs and inaccuracies in their writing may be challenging, daunting, and demoralizing. When compared to human readers and raters, AWE systems and applications have great potential for providing quick and consistent CF. Compared to instructor-provided feedback, AWE systems can sometimes offer far more detailed feedback owing to the additional writing resources integrated into these tools [20, 86].

Although research on the efficacy of AWE shows mixed and inconsistent findings [23], there is an increasing agreement among researchers that student writing quality can significantly improve as a result of using AWE systems when implemented in a context-appropriate manner [26, 53]. However, Camacho et al. [27] argued that making such generalizations can be dangerous since most of these studies also show that several variables are involved, such as the nature and amount of the provided instructional support, teachers' beliefs, practices, and attitudes toward the presence and use of automated writing assistance in language classes, and how

practice is provided to students. Variables related to students when discussing the effectiveness of using AWE are also important and maybe more telling. Personal characteristics of students including their proficiency level, their beliefs, and attitudes about the usefulness and validity of AWE, and the stage at which the AWE system is used during the writing and editing process are all paramount variables that need to be added to the teachers' variables [25, 86].

Several studies have shown that the earlier the stage during which AWE is provided, the more useful AWE feedback will be [17, 29]. These studies have also shown that among the many types of revisions made by students using AWE systems, lexical appropriateness and grammar accuracy were the most frequent as opposed to revisions for content or structure [14, 16]. Furthermore, although significant improvement may be seen in individual texts due to the use of AWE systems, these studies hardly ever showed any long-term improvement or even were able to prove that learning did take place [29]. Some researchers, such as Ranalli [30], went further and strongly criticized the use of AWE tools in the language classroom and pointed out that these tools "did not live up to expectations" (p. 2). He argued that instead of enhancing their writing skills and contributing toward developing their second language, most students use them merely for proofreading with no or little cognitive processing.

Despite the recent significant advances in AI-powered AWE systems where quick, synchronous, and varied automated support is provided to writers, there are areas that AWE systems that fall short of offering assistance. Organization, coherence, and argumentation strength are all examples of areas that AWE systems may not be of great help. This may be attributed to the complexity of human language that makes it very difficult for AI systems to fully understand the richness and complexity of human language and its contextual pragmatic and contextual aspects of a language [29, 86].

Among the reasons why AWE studies were blamed is the fact that many researchers are associated or affiliated with the companies that sell these systems [86]. Their contribution to the early body of literature guided the research that came after in which they emphasized the reliability of these systems and how closely aligned they were with the feedback that is normally produced by human raters [22, 87]. This is also reflected in the early use of AWE systems in formal assessment contexts where the required type is strictly defined [15, 18]. The use of these AWE systems in formal assessment contexts is reflected in those early studies in which the kinds of writing tasks were specifically and precisely defined [16, 88]. Instead of taking AWE's published research and brochures for granted, teachers, researchers, and scholars are encouraged to make these companies accountable by validating their claims and conducting systematic and critical studies that could improve AWE research [30, 85]. Conflicting research claims and findings, a lack of details about the possible uses of such systems, and the lack of control groups in those studies make it hard to reach solid conclusions about the ultimate use of AWE tools. Hibert [19] criticized the nature of AWE research as being "theoretically and methodologically fragmented" (p. 209). Other researchers such as Ranalli and Yamashita [3]

called for more independent research studies to combat the inadequacy of methodological information regarding the way their AI systems are configured. Hibert [19], in his systematic review of AWE literature, found it surprising to see many AWE studies generally failing to benefit from the amount of multi-layered data that are automatically generated and collected by these systems through their data collection capabilities during the interaction between the users and their systems on the computer. To help draw a full picture of the effectiveness of automated feedback evaluation systems, both contextual and individual variables that are likely to influence the efficacy of AWE systems should be identified. To do so, clustering techniques and methods implemented in data mining research could be effectively used [18, 21, 24, 30].

3.1.2. Automated Writing Corrective Feedback (AWCF) Tools. Another underexplored area in computer-assisted language learning is the use of editing tools similar to AWE tools, which provide instant real-time AWCF [3]. While AWE tools provide feedback and suggestions for already written texts, AWCF tools such as Grammarly can continuously and simultaneously provide corrections and suggestions while writers compose the text. Other well-known AWCF tools besides Grammarly include ProWritingAid and Ginger [4]. AWCF tools mainly focus on lower-level writing errors, such as lexical and grammatical, leaving structural, and organizational errors untreated. Another difference between AWE and AWCF systems is accessibility. While access to AWE tools is provided via web portals, AWCF tools are available on various platforms. Grammarly, for instance, is available as an independent tool or embedded with some writing systems and text editors and processors-like Google Docs or Microsoft Word. Moreover, Grammarly has recently become available as an extension to be added to web browsers. Grammarly has gained popularity all over the world over the last few years with very strong marketing campaigns accompanying its evolving popularity. However, further research and studies must be conducted owing to its value in the field that represents a new advanced AI-powered technology that supports writers in the digital era we live in today [3, 4].

Using Grammarly in EFL educational settings has been examined in several studies in the literature. Many of these studies have found that the feedback generated by Grammarly was mostly accurate. However, other studies found that Grammarly was unable to flag errors accurately by either over-flagging (otherwise known as false positives), or by missed-flagging (otherwise known as false negatives) [32, 40, 42]. As with the feedback generated by AWE, Grammarly's feedback has also been criticized for being either too long or overly repetitive [32, 41]. Moreover, the way Grammarly worded its feedback was also a concern. In an attempt to be easily understood, Grammarly is programmed to avoid providing explanations that are too difficult to be understood by nonspecialized users. This avoidance of providing very technical feedback sometimes results in endangering the best utilization of the feedback provided by oversimplifying it [42]. In contrast, other studies in the literature described Grammarly's feedback as being sometimes too technical and hence very difficult to

understand [32, 37]. It is often very complex to explain how writers process the automated CF they receive from such systems, but two of the most important factors are grammatical terminology knowledge and the proficiency level of the learner. According to Zheng and Yu [34], limited linguistic knowledge can prevent students from adequately processing feedback, preventing them from taking advantage of further revision opportunities.

It has been difficult to find consistent recommendations for how Grammarly should be used in studies. While some researchers recommend it for low-proficiency language learners or beginners [35], others think otherwise and recommend its use among advanced English language learners. According to Koltovskaia [2], AWCF may not be fully understood by students who lack the required linguistic competence and therefore may not be able to use Grammarly effectively. Despite the obvious concerns around the nature and accuracy of the feedback provided by AWCF, there is almost unanimous agreement amongst researchers who investigated the use of Grammarly in language classrooms. These researchers recommend that if Grammarly is to be introduced in the language classroom, it is advisable to use it as a starting point coupled with teacher feedback and not as a stand-alone tool [41, 42].

There are several positive aspects to Grammarly that have been reported in the reviewed literature. The positives include its speed in giving feedback, versatility in access platforms, and its availability in two versions; free and paid with adequate features in the free version [35, 39]. The findings of numerous studies suggest that using Grammarly indeed improves writing quality [38]. Moreover, it was found that its use resulted in lexical diversity gains [44]. One of the top features of Grammarly researchers found useful was error categorization [41]. Unlike human raters who may have difficulty categorizing the exact nature of all the errors in L2 texts, the algorithmic analysis capabilities of Grammarly provide personalized and targeted feedback based on the nature of the error [41]. Furthermore, Grammarly's ability to identify textual borrowing was praised by many scholars as it helped students avoid plagiarism [42]. O'Neill and Russell [41] argued that Grammarly allows learners to correct their writing before final submission, in addition to helping students develop self-regulation skills given its ease of use and availability on different platforms.

Grammarly and similar tools do not distinguish between texts written by a native speaker and an L2 learner, which could be problematic sometimes. From a performance point of view and when compared to texts written by native speakers, texts written by language learners typically tend to incorporate unpredictable and more complex errors as a result of language interference [33, 36]. The more complicated errors a text may have, the longer it will take AWCF systems to process the text and give feedback. This delay may be also attributed to the fact that both output and parsing processes are all done in the cloud [3]. None of these tools have settings that allow for different feedback based on user characteristics (e.g., if you are an L2 learner). If available, this can increase the speed with which automated feedback is processed for L1

users. On the other hand, L2 users may also benefit from such a setting as they can instruct the system to employ a hybrid approach in which the generated feedback comes from the learner corpus as well as the stored structured data [43, 89]. Hence, adding this feature to AWCF systems is likely to add a much-needed option of differentiation in the specificity and nature of the CF allowing these systems to accommodate more varied CF depending on the writing task and the characteristics of the writer [1].

The feedback generated by Grammarly is not generic but specific to the type of error identified in the text. Although continuous access to CF is more effective for revising, it may also reinforce students' "low-level focus" on grammar and spelling instead of meaning [3, p. 14]. The effectiveness and usefulness of the CF could also be dependent on the type of writing task at hand [1]. Hence, it can be argued that although specific and explicit CF can improve writing quality, implicit feedback can lead to long-term L2 gains. A meta-analysis of CF by Li [90] showed that the generic and implicit feedback was found to be more effective at improving long-term learning, which was confirmed by posttests that were administered long after the study. For implicit CF to leave a positive effect on long-term L2 gains, it needs time to see positive outcomes [31]. More longitudinal research is needed, however, to see if automated corrected feedback generated by computers would prove to be more effective in learning gains rather than teacher-supplied CF.

3.1.3. AI-Powered Automated Translation Tools. Bringing MT to the foreign language classroom has been very controversial. Vinall and Hellmich [91] believed that when comparing MT systems to the other available digital tools, the former stands out as particularly polemical. Generally, language teachers tend to forbid its use in their classes [56]. Crossley [48] argued the reasons why many language teachers discourage the use of MT in their classrooms are either they consider using it as cheating, or they fear it could lead to an end to the demand for FL instructors. The widespread notion that the way students use Google Translate for completing their assignments is just by copying and pasting without engaging with the target language is not accurate and is an overly simplistic point of view [10, 56, 57]. Second-language learners tend to use MT to look up individual words or phrases rather than translating whole texts from their first language to the target language. Research studies surveyed the use of MT in the language classroom and asked participants why they used Google Translate revealed that students used it for its convenience and speed [20, 92]; in addition to the fact, it is freely available and accessible through many platforms and mobile devices [46, 93]. The majority of students surveyed in recent studies [7, 51] used Google Translate for completing learning tasks in various language education settings. It seems that both Grammarly and Google Translate have become ubiquitously indispensable tools for students writing in a second language.

Several studies in the literature have explored the use of MT in second language acquisition, with special attention to using it in writing tasks [4, 57]. Many studies entail allowing

students to use MT for completing their first drafts or comparing their first drafts with machine-translated versions. Many studies have demonstrated significant improvements in the writing quality of students' writing when MT was integrated into the learning tasks [45, 49, 92, 94, 95]. Nevertheless, as it is the case with AWE research, research on the use of MT in the L2 writing classroom seems to mainly focus on examining the quality of the writing samples produced using MT rather than trying to answer the one-million-dollar question as to whether or not there was evidence of any gains in lexical or grammatical knowledge or any long-term transfer to general writing ability. Again, and as it is the case with AWE research, several MT studies have reported significant improvement when teachers mediated the learning process and provided training on the use of MT [49, 52, 95]. In a recent study that examined the impact of providing training on editing texts produced by MT, Zhang and Torres-Hostench [55] found that students could successfully correct raw MT output and gained insight into MT limitations. It takes both focused attention and advanced reading ability to master postediting skills, which are essential both in professional translation and language learning.

According to Hellmich and Vinall [7], among the possible drawbacks of promoting the use of Google Translate in the foreign language classroom are that it could lead both learners and teachers to form a wrong reductionist perception of language. In other words, it can promote the idea that human languages are merely discrete and unique codes that can be easily re-encoded from one language to another based upon a one-for-one transfer from one language to another [58]. As a result of mistakenly forming a simplistic and instrumentalist view of language Hellmich and Vinall [7] warned that some learners might think of Google Translate as an answer key to their language problems and, therefore, fail to accept the complexity and richness of human interaction [54]. While MT might be able to capture the semantic aspect of language, it is likely to miss the nuance and context-dependent properties behind human communication and interaction. Students may see accuracy in language use as the primary goal of language learning (just as offered by the CF generated by AWE or MT). That is, at least, how language learners have been conditioned by most L2 classroom practices and formal assessments.

One of the potential benefits of introducing MT into a second language classroom is making use of students' first language and contrasting the patterns of its uses with the learning of second language [10]. As such, this is consistent with SLA's multilingual turn research that enables learners to work with multiple languages independently and comfortably [96, 97]. Exploring the different roles of lexicogrammatical structures in different languages moves students away from traditional grammatical and lexical distinctions to usage-based models of language that emphasize patterns and collocational use of language [56, 98]. Thus, examining machine-translated texts may challenge the theory that languages are regulated and based on grammatical rules [99]. Working with MT can offer students insight into the epigamic characteristics of the language through its emphasis on

statistical probability provided that the MT consultation and the whole learning process are mediated by the teacher [48]. Following the common use of MT exclusively for lexical assistance is not likely to guarantee these insights [9, 50]. Likewise, a useful approach to developing a usage-based understanding of language would be to work with corpora [47, 99].

3.1.4. Automatic Text Generation and Deep Learning Technology. The performance of MT and automated writing and writing evaluation tools has significantly improved over the last few years thanks to the evolving strength of large language models (LLM) which make up the basis and foundation of NLU in the field of language technology nowadays. With the appearance of advanced generations of language models, AI systems have been increasingly becoming more able to create texts on their own using predictive text technology. Simply put, language modeling is associated with predicting what word should come next considering what word preceded it [100]. Large language models can be defined as AI systems that are based on large datasets which can be analyzed through machine learning and can lead to the capacity to interact with human language efficiently. The language model optimized and implemented in neuro-linguistic programming is built on mathematical modeling of big linguistic data, not on language grammatical knowledge. In order to understand how large language models work and how they made it possible for computers to generate or author texts autonomously, it is necessary to understand that these models are mainly AI systems that are made from huge data libraries that are analyzed by machine learning. These advances in technology have led to the ability to process human language in efficient ways. Basing language models on complicated analysis of statistical data are not a new concept as it can be traced back to the 1940s when N-Gram models made their first appearance in the field of computational linguistics and probability [64].

Despite its huge steps forward in NLP, GPT-3 follows trends that are already underway in AI-powered writing assistance [62, 65]. With advances in language modeling, writing tools have moved toward automatic text generation. The emergence of intelligent text generators may mark the "biggest change in writing since the invention of the word processor" [61, p. 691]. It is indisputable that artificial intelligent powered writing assistance tools are now providing assistance that was never available in the near past. In 2018, Google added a feature called smart compose to its search and writing products. In addition to providing autocompletion suggestions, it can also be customized to match the context of the sentence being typed. For example, Google's email client not only suggests wording and offers autocompletion options based on the text you just typed in the email but also tailors the suggestions by taking into consideration the sender's message to which you are replying. Microsoft Office too improved its well-known grammar and spelling checkers by incorporating GPT models into its Microsoft Editor where all Microsoft products can now offer text prediction and paraphrasing capabilities in addition to spelling and grammar checking

features. As far as Grammarly is concerned, it has been enhanced with predictive text features in addition to its grammar-checking capabilities [44, 59]. As it is reported by Dale [63], autocompletion in writing assistance tools can be regarded as an essential feature rather than an optional one.

Not long ago, autocompletion capabilities were limited to words and phrases, which meant that the suggestions were more likely to be correct. Dale [60, p. 485] refered to this type of text generation as "short leash." Nevertheless, GPT-3 models, and their anticipated stronger successors, have changed and will definitely continue to change the game. Godwin-Jones [28] demonstrated that writing tools that are based on GPT-3 models can generate significantly longer texts across a wide range of genres. As far as textual coherence and flow are concerned, generated texts often closely resemble those written by humans. Eaton et al. [5] described the texts generated by GPT-3 model as dreadfully convincing. To generate texts based on GPT-3 model, the system does not require any training as all it needs to function properly is a few prompts. In order to generate text, one must describe the writing task briefly or provide an example. GPT-3 generates texts in a variety of languages despite the overwhelming majority of the data being in English. In an attempt to see if OpenAI's GPT-3 model is able to generate creative writing, the system successfully completed the poem using the right sonnet format and with stanzas in Italian [61]. Dale and Viethen [4] suggested that the GPT-3 model could even write computer codes, compose poetry, translate texts, summarize texts, correct grammar, and even power chatbots. Rather than assisting writers with writing, it writes on their behalf.

The use of such text generators in educational settings raises a whole set of issues. Eaton et al. [5] argued that it is likely that intelligent text generators will be used by many students across all disciplines once widely available. In this case, authenticity, creativity, and attribution are at stake. In essence, humans and machines cocreate texts and hence share authorship. The assessment of such written work will present a challenge for language educators as they must find creative ways to assign credit fairly and consistently. It will be necessary for writing teachers to find tasks that blend automatic text generation with student effort, just as they did with MT [28].

The aforementioned discussion serves as answers to research questions 1 and 2. For research question 3, which is related to the recommendations put forward by the scholars in the reviewed studies, the following sections discuss these recommendations in detail.

3.2. Theme (2): Recommendations by Scholars for Researchers and Educators. The new realm of writing support presented by the advent of text generators and the widespread use of MT and AWE/AWCF tools highlights opportunities and challenges for L2 teachers in general and writing teachers in particular. It is both unrealistic and unacceptable to reject or ignore the use of advanced writing assistance tools after they have become so naturalized and widely available in a globalized and modern world [7]. In automated writing assistance literature, several studies have called for

boycotting these tools and banning their use in instructed learning environments, since these tools are believed to offer unethical help to students and threaten academic integrity. However, in the recent literature reviewed in this study, several other researchers have called for more realistic approaches that acknowledge the existence of AI-powered writing assistance tools in the classroom whether we like it or not. They also acknowledge that these tools can bring great benefits to the learning process if educators changed the way they view these tools and adopted a more holistic perspective toward them. The recommendations are discussed in the following section.

3.2.1. Artificial Intelligence-Powered Writing Assistance Tools in the Classroom: A Call for Integration. Writing assistance tools that are powered by AI technology need to be used and advocated with thoughtful, informed differentiation based on situated practices, goals, and expectations. In other words, these tools should be used according to their fit with pedagogical and curricular objectives, not based on their convenience [72]. Regrettably, administrative bureaucracies, institutional regulations, stakeholder pressure, and marketing hype might not give educational systems, foreign language programs, or even individual interested faculty the option of making their own decisions. And even if a specific writing assistance tool is mandated, there will be a variety of writing experience opportunities. The use of AI writing tools should be balanced by assigning writing tasks involving both the system and other means. The targeted reading audiences could go beyond the AI system and the teacher where possible [23]. Students should not be distracted from the communicative purpose of writing by AI writing tools. Their interaction with the tool should be part of a comprehensive language program that does not neglect the significance of communication [23].

This applies to both MT, AWE, and AWCF tools. It might mean integrating MT into everyday classroom communicative activities. Using context and word choice are important even in simple tasks. The same is true for registers, genres, and styles. The word "That's great!," for instance, could be interpreted in several different ways, depending on the context in which it occurred, for instance, if it was a reaction to a positive or negative situation. Automated translations are unlikely to reflect such pragmatic and contingent considerations. In their paper, Pellet and Myers [9] explained how common L2 learning tasks may be utilized to demonstrate the pragmatics of a given language and the limits of MT. Likewise, Ranalli [30] recommended that learners could review AWE feedback critically to determine its usefulness and effectiveness. Following such an activity, he suggested giving students a text that they are to proofread for which they have to identify errors and then correct them. It is beneficial to provide students with hands-on activities that will help them to become informed users of language-assistance resources.

The categorization feature in Grammarly can be used to target specific grammar points in focus-on-form activities, as recommended by John and Woll [40]. Moreover, they suggested that teachers may ask students to check their own writing for a specific type of error and get an AWCF tool to

check their writing to see if it flags that type of error. Another way to integrate writing assistance tools into the second language classroom is to find learning tasks that contain specific material previously or currently learnt and use them as prompts for learning. To do so, Knowles [73] suggested that teachers may ask their students to come up with checklists of specific vocabulary or grammar based on their encounters with the writing assistance tool. He suggested that the grading rubrics are to be based on Google Translate and include vocabulary and grammar identification. Additionally, Pellet and Myers [9] recommend an activity that utilizes Google Translate to encourage learners to connect recent study topics to recent experiences. In that exercise, teachers ask students to discuss the sociopragmatic aspects found in a text that is translated by Google Translate. Moreover, teachers may also ask students to record their experiences and encounters with the writing assistance tool in a diary to use it for future reflective practices. While practical classroom experiences and teacher mediation are all integral parts of any plans for introducing intelligent writing assistance tools into the classroom, several studies have also shown that explicitly directed instruction and systematic guidance are equally useful. Part of that process could be raising teachers' and students' awareness about how these intelligent writing assistance tools work, and the type of writing tasks and activities they are best fitted for and the limitations of their performance. By building familiarity as well as confidence, "calibrated trust" can be established in their use [30, p. 14]. It is important for students to develop realistic expectations about the utility of tools when choosing and using them. It is recommended to design a holistic writing strategies training course that combines conventional writing strategies with writing strategies using automated writing assistance tools [71].

A greater understanding of metalinguistics can be achieved through training and usage modeling of AI tools. That may result in turning these tasks into language-related episodes [101], where students explicitly talk about their language interactions and negotiate meaning. Koltovskaia's [2] study of Grammarly uses also discussed language-related episodes. There are several examples of MT provided by Pellet and Myers [9] as well as AWE examples provided by Woodworth and Barkaoui [69]. It is hoped that such experiences will assist in the appropriate use of advanced language tools in the future and may result in greater learner autonomy.

Students' attitudes and teachers' beliefs toward technology tools may have a definitive effect on the effectiveness of such tools [68]. External factors are just as important for teachers as they are for students. Both curricular and administrative factors likely influence teachers' use of technology. An important factor is teachers' comfort level with technical tools that may be relevant to teaching and support [70]. A complicating factor for writing assistance tools that are powered by AI is the speed of development and change these tools constantly go through. In the past, educators who disallowed Google Translate thinking of its translation as unreliable might not now have realized how much it has developed today. Likewise, automatic writing evaluation and AWCF systems are constantly improving and growing in functionality.

3.2.2. Adopting an Ecological Perspective Toward Automated Writing Assistance Tools.

Teachers who use AWE or MT as an instructional tool for improving writing and language development tend to use several other strategies as well to provide feedback. It has been shown that automatic writing evaluation and AWCF studies emphasize the importance of keeping using instructor feedback on student writing rather than relying only and exclusively on automated feedback [29, 52]. An ideal situation was proposed by Link et al. [23] in which they proposed a hybrid approach. In their proposed approach, the sentence-level problems are dealt with by the AWE tool while the higher-order writing issues are left to the teacher to provide feedback on. It is also possible to combine AWE with peer review [78], as well as MT [72]. Moreover, Pellet and Myers [9] suggested a more complicated hybrid approach in which they also described a three-step revising process, starting from AWE to peer evaluation to teacher CF. Peer review is now part of some AWE tools, such as MI Write and Criterion. Among the Google Translate learning activities that were proposed by Pellet and Myers [9], there are several activities that promote learner-to-learner interactions.

Most of the unfavorable perceptions about AI writing assistance tools may be attributed to the failed trials for integrating them into some local learning environments. As described by Cotos [102], contextual factors tend to be overlooked when discussing AWE's benefits. Grimes and Warschauer [86] provided illuminating examples of how the integration of such tools into the learning environment impacts the success of these tools. Embracing AWE tools wholeheartedly as a powerful writing instruction tool [103] is as erroneous as disallowing Google Translate. Although AI can be used to develop students' writing skills, Huang and Wilson [29] stated that they should play a supporting, not leading role. Cotos [102, p. 647] puted that the "ecology of implementation" of automated writing assistance tools requires deliberate and thoughtful use of these tools in contextually appropriate manners. Although the discussion of artificial intelligence tools often lacks that larger, ecological perspective, many studies in the literature pointed in that direction. Grimes and Warschauer [86] suggested the notion of "social informatics" as an approach to breaking the barriers between educational organizations, technology, and teachers. Through this approach, technologies, people, and organizations are treated as a "heterogeneous socio-technical network," in which none of them can be understood without the others (p. 10). This opinion contrasts with a "tool" focus, which undervalues the role of organizations and people. Based on the mediated learning experience theory [104], Jiang et al. [68] considered AWE systems to be sociocultural artifacts mediated by teachers and students. This perspective highlights the fact that the use of AWE systems impacts both student writing and teacher CF as well. To categorize the scaffolding process that takes place when the automated CF is complimented by the teacher's feedback, Nunes et al. [26] and Woodworth and Barkaoui [69] suggested following the sociocultural theory. To characterize the emerging outcomes from people and tools interactions across different institutional and individual scales, Hellmich and Vinall [7] proposed using an ecological approach in education. Their proposed approach sees language teaching

and learning as developing from multilayered complex relationships between different components in a given ecosystem. Time is another important aspect to take into consideration. The majority of research conducted on writing assistance tools is short-term. Long-term benefits can be tracked only in a small portion of longitudinal research, such as of Huang and Wilson [29]. According to Li [72], teachers' role in technology-enhanced classrooms is an under-researched area, with more attention placed on tools than on teachers. He explained how using an AWE tool may inevitably alter the ecology of the learning and teaching processes. To fully understand this dynamic, it is important to take into account the individual differences between teachers in addition to the characteristics of students [2].

Hellmich and Vinall [7] thought an AWE tool that comprehensively takes into consideration both local settings and other factors such as the users' knowledge and familiarity with the topic, genre conventions, the appropriate register for the assigned writing task, and lexical considerations based on the intended audience. When viewed from an ecological perspective, writing assistance in classrooms can be expected to have wildly different results in different environments. According to complexity theory, complex systems, such as those involving the interaction between individuals, institutions, and nonhuman entities, produce emergent outcomes that are likely to vary from one situation to another [76]. Several factors influence the outcomes, including initial conditions, evolving nonlinear, and shifting layered relationships among system components, as well as potentially unexpected processes/encounters [79].

In light of that variability, tracing individual case histories is imperative for illuminating what might contribute to the success or failure with artificial intelligence writing tools. This ties well into the person-centered perspective approach which is increasingly being applied in second language acquisition research [76]. There are widely varying patterns of emergence among individual students across studies where the use of AWE systems was tracked, such as Zhang and Hyland [77] and Zhang [80]. As part of their study of Google Translate as a tool for learning Dutch vocabulary, van Lieshout and Cardoso [81] used surveys to identify individual variables such as languages spoken, prior use, educational backgrounds, and autonomy experiences. In their study about students' use of Grammarly, Ranalli [30] argued that learning orientations are likely to influence students' use of digital writing tools. Those orientations were heavily influenced by student identities, including their self-image as students in addition to confidence in their language abilities. One of the interesting cases in Ranalli's study was a student who attributed his success in using Grammarly for improving his writing to his "process-related knowledge of Grammarly's workings" (p. 13), which he had gained as a premium user. Ranalli [30] concluded that learners' engagement in these systems may be "complex and multifaceted" (p. 13) and, hence, could widely vary from one case to another. Koltovskaia [2] also examined how several individual learners used Grammarly, looking for patterns of engagement and disengagement that impacted the effectiveness of the feedback provided by the tool. In order to determine if digital tools are being used effectively, qualitative research focusing on individual student learning pathways can be quite useful.

One of the many variables that influence the dynamics of the use of artificial intelligence tools is the human–machine relationship. Whenever technology is incorporated into instruction and instruction is implemented into practice, there are likely to be differing attitudes and reactions, ranging from enthusiastic acceptance to complete rejection [105, 106]. As a result, that can bring learner emotions to the scene and add them to the already complex equation, which eventually is very likely to impact technology use and the effectiveness of learning. The reviewed research shows that when using technology tools and emotions such as mistrust and anxiety may negatively affect students' motivation for using technology [82]. When it comes to the importance of students' trust in the digital tool, Ranalli [30] found that learners' acceptance of the feedback generated by the AWE tools was dependent on and conditioned by their trust in that tool. In order to understand the dynamics involved, Ranalli [30] proposed the use of human automation trust theory by Lee and See [107]. In their view, a trust may be key to the level of engagement that users have with technology tools. With L2 students, AWE poses a high degree of vulnerability, since these students are trying to solve problems in a language that they have not mastered yet and interacting with a tool they are unfamiliar with, which generates feedback they have to understand and act upon.

4. Conclusion

In light of the advancements in automated writing assistance, second language learners and writing instructors ought to be more aware of what artificial intelligence systems can offer in regard to writing assistance [8]. With all publicity surrounding the use of artificial intelligence in writing assistance tools, there is no doubt that students will likely use text generators and other emerging writing tools regardless of their effectiveness or ethics. Since this is likely to be the case, educators and researchers are responsible for finding ways to allow students to use the tools appropriately, and integrating their use into instruction whenever possible [6]. The benefits of training L2 learners on the best use of AI writing tools extend to even after their graduation as they are likely to use them for improving their texts in their future careers. The ability to use these technologies has grown to be a critical aspect of digital literacy in educational and professional settings.

It is to be hoped that the developers of the automated writing assistance tools designed for the educational community will incorporate the recommendations and suggestions from researchers by adding features that would make these tools more useful to both teachers and students. One of the main improvements would be the addition of flexibility of use. It would be helpful and educationally valued, for example, if errors were highlighted only without being labeled or even corrected [30]. Admittedly, giving users the option to toggle some of the tools' features on and off is pedagogically favorable in general. Frankenberg-Garcia [108], for example, suggested a useful case of a good writing assistance tool for helping second language writers use collocations correctly

and appropriately. The recommended automated writing assistance system (ColloCaid) allows writers to incrementally display information related to word collocations with the ability to show different levels of collocations, examples, and metalinguistic information. In addition to feedback, the system is also programmed "feed forward" by bringing to writers' attention collocations that they may not have remembered or are not aware they need to look up. For learners with various linguistic backgrounds and educational objectives, this scaffolding support can offer the flexibility required. As part of improving systems to be more responsive to learner context, systems must also support process writing, that is, the feedback provided by these tools is adjusted depending on the different drafts and stages of writing along with the revisions required [30]. It is also required for writing assistance tools to improve the way they generate CF by being responsive to different writing genres. It is generally recommended to be aware of how AWE systems operate in real practice and the body of research to enhance compatibility between tools and teaching/learning environments. Feedback in an L2 setting, for example, may differ in nature and formulation from feedback in an L1 setting. In a recent study by Wilken [109], it was found that adding L1 glosses to automated writing feedback was helpful. A feature like that would be particularly helpful for novice learners, especially if it can be displayed in either L1 or L2 [110]. Users can view a simultaneous translation underneath their writing in their native language via a reverse translation function.

An ecological perspective regarding the use of AI writing assistance tools in education calls for a wider and closer look that considers other important aspects such as society, equity, and learner agency [111–113]. Carvalho et al.'s [8] study addresses these aspects with a particular focus on designing for learning in an AI-driven environment. The authors note that with the increasing integration of AI into everyday life and education, significant disruptions and changes are likely to take place, bringing a heightened sense of uncertainty. According to the study, "in an AI world, both teachers and students must be engaged not only in the teaching and learning processes but also in co-designing for better learning" (p. 1). Moreover, they should work together to explore the goals, knowledge, and actions that might assist users in shaping the future of artificial scenarios [74, 75].

With AI playing an increasingly influential role in second language education, it is not unlikely that both learners and educators might be contributing to the systems and cocreating with algorithms [66, 67]. For this to be done fairly, designing for learning requires looking at the AI system from a broad sociological outlook taking into account the possible impact on the lives of the individuals [114]. Lütge et al. [115] presented a similar proposal in which they suggested teaching global citizenship through foreign language education in order to "empower educational actors to orient themselves in the face of unknowns" [116, p. 2]. As AI-enhanced writing tools become more available to second language learners, foreign language teachers will have to find different ways to reward creativity and value the freedom of the learners [12, 13]. They

will also have to make use of these tools in reducing their workload by getting these systems to highlight students' errors and writing problems. Envetually, this will leave more room for the much-needed individualized feedback by the teachers [117].

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author declares that there is no conflicts of interest.

References

[1] J. Ranalli, "Automated written corrective feedback: how well can students make use of it?" *Computer Assisted Language Learning*, vol. 31, no. 7, pp. 653–674, 2018.

[2] S. Koltovskaia, "Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: a multiple case study," *Assessing Writing*, vol. 44, Article ID 100450, 2020.

[3] J. Ranalli and T. Yamashita, "Automated written corrective feedback: error-correction performance and timing of delivery," *Language Learning & Technology*, vol. 26, no. 1, pp. 1–25, 2022.

[4] R. Dale and J. Viethen, "The automated writing assistance landscape in 2021," *Natural Language Engineering*, vol. 27, no. 4, pp. 511–518, 2021.

[5] S. E. Eaton, M. Mindzak, and R. Morrison, "Artificial intelligence, algorithmic writing & educational ethics," in *Canadian Society for the Study of Education [Société canadienne pour l'étude de l'éducation] (CSSE) 2021*, Edmonton, AB, Canada, 2021, May 29–June 3.

[6] G. J. Otsuki, "OK computer: to prevent students cheating with AI text-generators, we should bring them into the classroom. The Conversation," January 2020, https://theconversation.com/ok-computer-to-prevent-students-cheating-with-ai-text-generators-we-should-bring-them-into-the-classroom-129905.

[7] E. Hellmich and K. Vinall, "FL instructor beliefs about machine translation: ecological insights to guide research and practice," *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, vol. 11, no. 4, pp. 1–18, 2021.

[8] L. Carvalho, R. Martinez-Maldonado, Y.-S. Tsai, L. Markauskaite, and M. De Laat, "How can we design for learning in an AI world?" *Computers and Education: Artificial Intelligence*, vol. 3, Article ID 100053, 2022.

[9] S. Pellet and L. Myers, "What's wrong with "What is your name?" > "Quel est votre nom?": teaching responsible use of MT through discursive competence and metalanguage awareness," *L2 Journal*, vol. 14, no. 1, pp. 166–194, 2022.

[10] K. Vinall and E. A. Hellmich, "Do you speak translate? Reflections on the nature and role of translation," *L2 Journal*, vol. 14, no. 1, pp. 4–25, 2022.

[11] J. Baas, M. Schotten, A. Plume, G. Côté, and R. Karimi, "Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies," *Quantitative Science Studies*, vol. 1, no. 1, pp. 377–386, 2020.

[12] D. Ippolito, A. Yuan, A. Coenen, and S. Burnam, "Creative writing with an AI-powered writing assistant: perspectives from professional writers," 2022.

[13] D. Adams and K.-M. Chuah, "Artificial intelligence-based tools in research writing: current trends and future potentials," in *Artificial Intelligence in Higher Education*, pp. 169–184, CRC Press, 2022.

[14] J. Burstein, N. Elliot, and H. Molloy, "Informing automated writing evaluation using the lens of genre: two studies," *CALICO Journal*, vol. 33, no. 1, pp. 117–141, 2016.

[15] L. Ortega, "New CALL-SLA research interfaces for the 21st century: towards equitable multilingualism," *CALICO Journal*, vol. 34, no. 3, pp. 283–316, 2017.

[16] B. Bridgeman and C. Ramineni, "Design and evaluation of automated writing evaluation models: relationships with writing in naturalistic settings," *Assessing Writing*, vol. 34, pp. 62–71, 2017.

[17] J. Ranalli, S. Link, and E. Chukharev-Hudilainen, "Automated writing evaluation for formative assessment of second language writing: investigating the accuracy and usefulness of feedback as part of argument-based validation," *Educational Psychology*, vol. 37, no. 1, pp. 8–25, 2017.

[18] R. Godwin-Jones, "Second language writing online: an update," *Language Learning & Technology*, vol. 22, no. 1, pp. 1–15, 2018.

[19] A. I. Hibert, "Systematic literature review of automated writing evaluation as a formative learning tool," in *Transforming Learning with Meaningful Technologies*, M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, and J. Schneider, Eds., pp. 199–212, Springer, 2019.

[20] M. A. Hussein, H. Hassan, and M. Nassef, "Automated language essay scoring systems: a literature review," *PeerJ Computer Science*, vol. 5, Article ID e208, 2019.

[21] M. Warschauer, S. Yim, H. Lee, and B. Zheng, "Recent contributions of data mining to language learning research," *Annual Review of Applied Linguistics*, vol. 39, pp. 93–112, 2019.

[22] A. Saricaoglu, "The impact of automated feedback on L2 learners' written causal explanations," *ReCALL*, vol. 31, no. 2, pp. 189–203, 2019.

[23] S. Link, M. Mehrzad, and M. Rahimi, "Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement," *Computer Assisted Language Learning*, vol. 35, no. 4, pp. 605–634, 2022.

[24] H. Peng, S. Jager, and W. Lowie, "A person-centred approach to L2 learners' informal mobile language learning," *Computer Assisted Language Learning*, vol. 35, no. 9, pp. 2148–2169, 2022.

[25] C. Palermo and J. Wilson, "Implementing automated writing evaluation in different instructional contexts: a mixed-methods study," *Journal of Writing Research*, vol. 12, no. 1, pp. 63–108, 2020.

[26] A. Nunes, C. Cordeiro, T. Limpo, and S. L. Castro, "Effectiveness of automated writing evaluation systems in school settings: a systematic review of studies from 2000 to 2020," *Journal of Computer Assisted Learning*, vol. 38, no. 2, pp. 599–620, 2022.

[27] A. Camacho, R. A. Alves, and P. Boscolo, "Writing motivation in school: a systematic review of empirical research in the early twenty-first century," *Educational Psychology Review*, vol. 33, pp. 213–247, 2021.

[28] R. Godwin-Jones, "Big data and language learning: opportunities and challenges," *Language Learning & Technology*, vol. 25, no. 1, pp. 4–19, 2021.

[29] Y. Huang and J. Wilson, "Using automated feedback to develop writing proficiency," *Computers and Composition*, vol. 62, Article ID 102675, 2021.

[30] J. Ranalli, "L2 student engagement with automated feedback on writing: potential for learning and issues of trust," *Journal of Second Language Writing*, vol. 52, Article ID 100816, 2021.

[31] R. Ellis, "Focus on form: a critical review," *Language Teaching Research*, vol. 20, no. 3, pp. 405–428, 2016.

[32] J. M. Dembsey, "Closing the Grammarly gaps: a study of claims and feedback from an online grammar program," *Writing Center Journal*, vol. 36, no. 1, Article ID 5, 2017.

[33] D. C. Arroyo and Y. Yilmaz, "An open for replication study: the role of feedback timing in synchronous computer-mediated communication," *Language Learning*, vol. 68, no. 4, pp. 942–972, 2018.

[34] Y. Zheng and S. Yu, "Student engagement with teacher written corrective feedback in EFL writing: a case study of Chinese lower-proficiency students," *Assessing Writing*, vol. 37, pp. 13–24, 2018.

[35] M. Nova, "Utilizing Grammarly in evaluating academic writing: a narrative research on EFL students' experience," *Premise: Journal of English Education and Applied Linguistics*, vol. 7, no. 1, pp. 80–96, 2018.

[36] R. Conijn, M. van Zaanen, and L. van Waes, "Don't wait until it is too late: the effect of timing of automated feedback on revision in ESL writing," in *Transforming Learning with Meaningful Technologies*, M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, and J. Schneider, Eds., pp. 577–581, Springer, 2019.

[37] R. O'Neill and A. Russell, "Stop! Grammar time: university students' perceptions of the automated feedback program Grammarly," *Australasian Journal of Educational Technology*, vol. 35, no. 1, pp. 42–56, 2019.

[38] M. A. Ghufron, "Exploring an automated feedback program 'Grammarly' and teacher corrective feedback in EFL writing assessment: modern vs. traditional assessment," in *Proceedings of the 3rd English Language and Literature International Conference*, ELLiC, Semarang, Indonesia, April 2019.

[39] H.-W. Huang, Z. Li, and L. Taylor, "The effectiveness of using Grammarly to improve students' writing skills," in *Proceedings of the 5th International Conference on Distance Education and Learning*, pp. 122–127, Association for Computing Machinery, May 2020.

[40] P. John and N. Woll, "Using grammar checkers in an ESL context: an investigation of automatic corrective feedback," *CALICO Journal*, vol. 37, no. 2, pp. 193–196, 2020.

[41] R. O'Neill and A. M. T. Russell, "Grammarly: help or hindrance? Academic learning advisors' perceptions of an online grammar checker," *Journal of Academic Language and Learning*, vol. 13, no. 1, pp. A88–A107, 2020.

[42] M. Dodigovic and A. Tovmasyan, "Automated writing evaluation: the accuracy of Grammarly's feedback on form," *International Journal of TESOL Studies*, vol. 3, no. 2, pp. 71–88, 2021.

[43] G. Zomer and A. Frankenberg-Garcia, "Beyond grammatical error correction: improving L1-influenced research writing in English using pre-trained encoder–decoder models," in *Findings of the Association for Computational Linguistics: EMNLP 2021*, pp. 2534–2540, Association for Computational Linguistics, 2021.

[44] G. Dizon and J. M. Gayed, "Examining the impact of Grammarly on the quality of mobile L2 writing," *The JALT CALL Journal*, vol. 17, no. 2, pp. 74–92, 2021.

[45] E. M. O'Neill, "Measuring the impact of online translation on FL writing scores," *IALLT Journal of Language Learning Technologies*, vol. 46, no. 2, pp. 1–39, 2016.

[46] G. Lewis-Kraus, "The great A.I. awakening. The New York Times," December 2016, https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html.

[47] R. Godwin-Jones, "Data-informed language learning," *Language Learning & Technology*, vol. 21, no. 3, pp. 9–27, 2017.

[48] S. A. Crossley, "Technological disruption in foreign language teaching: the rise of simultaneous machine translation," *Language Teaching*, vol. 51, no. 4, pp. 541–552, 2018.

[49] K. Fredholm, "Effects of Google Translate on lexical diversity: vocabulary development among learners of Spanish as a foreign language," *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, vol. 13, no. 26, pp. 98–117, 2019.

[50] R. Godwin-Jones, "In a world of SMART technology, why learn another language?" *Journal of Educational Technology & Society*, vol. 22, no. 2, pp. 4–13, 2019.

[51] V. E. Raído and M. S. Torrón, "Machine translation, language learning and the knowledge economy," in *Reimagining Communication: Action*, M. Filimowicz and V. Tzankova, Eds., pp. 155–171, Taylor and Francis/Routedge, 2020.

[52] S.-M. Lee, "The impact of using machine translation on EFL students' writing," *Computer Assisted Language Learning*, vol. 33, no. 3, pp. 157–175, 2020.

[53] L. Flower and J. R. Hayes, "A cognitive process theory of writing," *College Composition and Communication*, vol. 32, no. 4, pp. 365–387, 1981.

[54] P. Urlaub and E. Dessein, "From disrupted classrooms to human–machine collaboration? The pocket calculator, Google Translate, and the future of language education," *L2 Journal*, vol. 14, no. 1, pp. 45–59, 2022.

[55] H. Zhang and O. Torres-Hostench, "Training in machine translation post-editing for foreign language students," *Language Learning & Technology*, vol. 26, no. 1, pp. 1–17, 2022.

[56] V. Klekovkina and L. Denié-Higney, "Machine translation: friend or foe in the language classroom?" *L2 Journal*, vol. 14, no. 1, pp. 105–135, 2022.

[57] J. R. Jolley and L. Maimone, "Thirty years of machine translation in language teaching and learning: a review of the literature," *L2 Journal*, vol. 14, no. 1, pp. 26–44, 2022.

[58] J. Ryu, Y. A. Kim, S. Park, S. Eum, S. Chun, and S. Yang, "Exploring foreign language students' perceptions of the guided use of machine translation (GUMT) model for Korean writing," *L2 Journal*, vol. 14, no. 1, pp. 136–165, 2022.

[59] S. Ruder, "A review of the neural history of natural language processing. AYLIEN," October 2018, https://aylien.com/blog/a-review-of-the-recent-history-of-natural-language processing#nonneuralmilestones.

[60] R. Dale, "Natural language generation: the commercial state of the art in 2020," *Natural Language Engineering*, vol. 26, no. 4, pp. 481–487, 2020.

[61] L. Floridi and M. Chiriatti, "GPT-3: its nature, scope, limits, and consequences," *Minds and Machines*, vol. 30, pp. 681–694, 2020.

[62] L. Ferrone and F. M. Zanzotto, "Symbolic, distributed, and distributional representations for natural language processing in the era of deep learning: a survey," *Frontiers in Robotics and AI*, vol. 6, Article ID 153, 2020.

[63] R. Dale, "GPT-3: what's it good for?" *Natural Language Engineering*, vol. 27, no. 1, pp. 113–118, 2021.

[64] M. Zhang and J. Li, "A commentary of GPT-3 in *MIT Technology Review 2021*," *Fundamental Research*, vol. 1, no. 6, pp. 831–833, 2021.

[65] V. J. Schmalz and A. Brutti, "Automatic assessment of English CEFR levels using BERT embeddings," in *Italian Conference on Computational Linguistics 2021*, Bologna, Italy, http://ceur-ws.org/Vol-3033/paper14.pdf, July 2021.

[66] C. M. Anson, "AI-based text generation and the social construction of "Fraudulent Authorship": a revisitation," *Composition Studies*, vol. 50, no. 1, pp. 37–46, 2022.

[67] C. M. Anson and I. Straume, "Amazement and trepidation: implications of AI-based natural language production for the teaching of writing," *Journal of Academic Writing*, vol. 12, no. 1, pp. 1–9, 2022.

[68] L. Jiang, S. Yu, and C. Wang, "Second language writing instructors' feedback practice in response to automated writing evaluation: a sociocultural perspective," *System*, vol. 93, Article ID 102302, 2020.

[69] J. Woodworth and K. Barkaoui, "Perspectives on using automated writing evaluation systems to provide written corrective feedback in the ESL classroom," *TESL Canada Journal*, vol. 37, no. 2, pp. 234–247, 2020.

[70] G. Kessler, "Current realities and future challenges for CALL teacher preparation," *CALICO Journal*, vol. 38, no. 3, pp. i–xx, 2021.

[71] G. Ling, N. Elliot, J. C. Burstein, D. F. McCaffrey, C. A. MacArthur, and S. Holtzman, "Writing motivation: a validation study of self-judgment and performance," *Assessing Writing*, vol. 48, Article ID 100509, 2021.

[72] Z. Li, "Teachers in automated writing evaluation (AWE) system-supported ESL writing classes: perception, implementation, and influence," *System*, vol. 99, Article ID 102505, 2021.

[73] C. L. Knowles, "Using an ADAPT approach to integrate Google Translate into the second language classroom," *L2 Journal*, vol. 14, no. 1, pp. 195–236, 2022.

[74] D. T. Y. G. Sumakul, F. A. Hamied, and D. Sukyadi, "Artificial intelligence in EFL classrooms: friend or foe?" *LEARN Journal: Language Education and Acquisition Research Network*, vol. 15, no. 1, pp. 232–256, 2022.

[75] P. Fyfe, "How to cheat on your final paper: assigning AI for student writing," *AI & Society*, 2022.

[76] D. Larsen-Freeman, "Looking ahead: future directions in, and future research into, second language acquisition," *Foreign Language Annals*, vol. 51, no. 1, pp. 55–72, 2018.

[77] Z. V. Zhang and K. Hyland, "Student engagement with teacher and automated feedback on L2 writing," *Assessing Writing*, vol. 36, pp. 90–102, 2018.

[78] N. Hockly, "Automated writing evaluation," *ELT Journal*, vol. 73, no. 1, pp. 82–88, 2019.

[79] P. A. Patout and M. Cordy, "Towards context-aware automated writing evaluation systems," in *Proceedings of the 1st ACM SIGSOFT International Workshop on Education through Advanced Software Engineering and Artificial Intelligence*, B. Vanderose, Ed., pp. 17–20, Association for Computing Machinery, 2019.

[80] Z. V. Zhang, "Engaging with automated writing evaluation (AWE) feedback on L2 writing: student perceptions and revisions," *Assessing Writing*, vol. 43, Article ID 100439, 2020.

[81] C. van Lieshout and W. Cardoso, "Google Translate as a tool for self-directed language learning," *Language Learning & Technology*, vol. 26, no. 1, pp. 1–19, 2022.

[82] B. Sun and T. Fan, "The effects of an AWE-aided assessment approach on business English writing performance and writing anxiety: a contextual consideration," *Studies in Educational Evaluation*, vol. 72, Article ID 101123, 2022.

[83] J. Bitchener and D. R. Ferris, *Written Corrective Feedback in Second Language Acquisition and Writing*, Routledge, 2012.

[84] E. Y. Kang and Z. Han, "The efficacy of written corrective feedback in improving L2 written accuracy: a meta-analysis," *The Modern Language Journal*, vol. 99, no. 1, pp. 1–18, 2015.

[85] K. Hyland and F. Hyland, "Feedback on second language students' writing," *Language Teaching*, vol. 39, no. 2, pp. 83–101, 2006.

[86] D. Grimes and M. Warschauer, "Utility in a fallible tool: a multi-site case study of automated writing evaluation," *Journal of Technology, Language, and Assessment*, vol. 8, no. 6, pp. 1–43, 2010.

[87] V. Hegelheimer and J. Lee, "The role of technology in teaching and researching writing," in *Contemporary Computer-Assisted Language Learning*, M. Thomas, H. Reinders, and M. Warschauer, Eds., pp. 287–302, Bloomsbury, 2012.

[88] M. Stevenson, "A critical interpretative synthesis: the integration of automated writing evaluation into classroom writing instruction," *Computers and Composition*, vol. 42, pp. 1–16, 2016.

[89] C. Leacock, M. Chodorow, M. Gamon, and J. Tetreault, "Automated grammatical error detection for language learners," in *Synthesis Lectures on Human Language Technologies*, p. XV, 154, Springer, Cham, 2nd edition, 2014.

[90] S. Li, "The effectiveness of corrective feedback in SLA: a meta-analysis," *Language Learning*, vol. 60, no. 2, pp. 309–365, 2010.

[91] K. Vinall and E. A. Hellmich, "Down the rabbit hole: machine translation, metaphor, and instructor identity and agency," *Second Language Research & Practice*, vol. 2, no. 1, pp. 99–118, 2021.

[92] K. Fredholm, "Online translation use in Spanish as a foreign language essay writing: effects on fluency, complexity and accuracy," *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, vol. 18, pp. 1–18, 2015.

[93] A. Sukkhwan, "*Students' attitudes and behaviors towards the use of Google Translate*," Unpublished master's thesis, Prince of Songkla University, 2014.

[94] K. Fredholm, "Effects of online translation on morphosyntactic and lexical-pragmatic accuracy in essay writing in Spanish as a foreign language," in *CALL Design: Principles and Practice. Proceedings of the 2014 EUROCALL Conference. Groningen, The Netherlands*, S. Jager, L. Bradley, E. J. Meima, and S. Thouësny, Eds., pp. 96–101, Research-Publishing Net, 2014.

[95] E. M. O'Neill, "Training students to use online translators and dictionaries: the impact on second language writing scores," *International Journal of Research Studies in Language Learning*, vol. 8, no. 2, pp. 47–65, 2019.

[96] MLA Ad Hoc Committee on Foreign Languages, "Foreign languages and higher education: new structures for a changed world," *Profession*, vol. 2007, pp. 234–245, 2007.

[97] L. Ortega, "SLA for the 21st century: disciplinary progress, transdisciplinary relevance, and the bi/multilingual turn," *Language Learning*, vol. 63, no. s1, pp. 1–24, 2013.

[98] M. Tomasello, *Constructing a Language: A Usage-Based Theory of Language Acquisition*, Harvard University Press, 2003.

[99] N. C. Ellis, "Cognition, corpora, and computing: triangulating research in usage-based language learning," *Language Learning*, vol. 67, no. S1, pp. 40–65, 2017.

[100] A. Kasirzadeh and I. Gabriel, "In conversation with Artificial Intelligence: aligning language models with human values," 2022.

[101] M. Swain and S. Lapkin, "Problems in output and the cognitive processes they generate: a step towards second language learning," *Applied Linguistics*, vol. 16, no. 3, pp. 371–391, 1995.

[102] E. Cotos, "Genre-based automated writing evaluation," in *Research Questions in Language Education and Applied Linguistics*, pp. 645–650, Springer, Cham, 2021.

[103] M. Warschauer and P. Ware, "Automated writing evaluation: defining the classroom research agenda," *Language Teaching Research*, vol. 10, no. 2, pp. 157–180, 2006.

[104] I. Lee, "Revisiting teacher feedback in EFL writing from sociocultural perspectives," *TESOL Quarterly*, vol. 48, no. 1, pp. 201–213, 2014.

[105] W. Alharbi, "E-feedback as a scaffolding teaching strategy in the online language classroom," *Journal of Educational Technology Systems*, vol. 46, no. 2, pp. 239–251, 2017.

[106] W. Alharbi, "Students' perceptions and challenges in learning business English: understanding students' needs and job market requirements," *International Journal of Learning, Teaching and Educational Research*, vol. 21, no. 12, pp. 65–87, 2022.

[107] J. D. Lee and K. A. See, "Trust in automation: designing for appropriate reliance," *Human Factors*, vol. 46, no. 1, pp. 50–80, 2004.

[108] A. Frankenberg-Garcia, "Combining user needs, lexicographic data and digital writing environments," *Language Teaching*, vol. 53, no. 1, pp. 29–43, 2020.

[109] J. L. Wilken, "Perceptions of L1 glossed feedback in automated writing evaluation: a case study," *CALICO Journal*, vol. 35, no. 1, pp. 30–48, 2017.

[110] J. M. Gayed, M. K. J. Carlon, A. M. Oriola, and J. S. Cross, "Exploring an AI-based writing assistant's impact on English language learners," *Computers and Education: Artificial Intelligence*, vol. 3, Article ID 100055, 2022.

[111] T. N. Fitria, "Analysis on clarity and correctness of Google Translate in translating an Indonesian article into English," *International Journal of Humanity Studies*, vol. 4, no. 2, pp. 256–366, 2021.

[112] L. K. Wei, "The use of Google Translate in English language learning: how students view it," *International Journal of Advanced Research in Education and Society*, vol. 3, no. 1, pp. 47–53, 2021.

[113] J. E. Cansancio, L. I. C. Rillon, S. D. Sowaib, and C. A. G. Tuppal, "Perception of foreign language learners in the use of internet-based tools," *International Journal of Research Publications*, vol. 93, no. 1, pp. 208–222, 2022.

[114] W. H. Alharbi, "The affordances of augmented reality technology in the English for specific purposes classroom: it's impact on vocabulary learning and students motivation in a Saudi higher education institution," *Journal of Positive School Psychology*, vol. 6, no. 3, pp. 6588–6602, 2022.

[115] C. Lütge, T. Merse, and P. Rauschert, *Global Citizenship in Foreign Language Education: Concepts, Practices, Connections*, Routledge, 2002.

[116] C. Kramsch, "General editor's preface," *L2 Journal*, vol. 14, no. 1, pp. 1–3, 2022.

[117] C. A. Chapelle, E. Cotos, and J. Lee, "Validity arguments for diagnostic assessment using automated writing evaluation," *Language Testing*, vol. 32, no. 3, pp. 385–405, 2015.