Carreras, B. H. (2026). Combining automated writing evaluation with metalinguistic reflection to develop language skills. *Intersection: A Journal at the Intersection of Assessment and Learning, 7*(1), 139-157.

Combining Automated Writing Evaluation with Metalinguistic Reflection to Develop Language Skills

Helena Borrell Carreras, Ph.D.

Author Note

Helena Borrell Carreras, ORCiD: https://orcid.org/0000-0001-8346-4369

"I have no conflicts of interest to disclose."

Intersection: A Journal at the Intersection of Assessment and Learning
Volume 7, Issue 1

Abstract: This paper investigates the potential of grammar checkers as automated writing evaluation (AWE) tools to promote effective feedback amongst undergraduate students who are learning grammar and writing skills. By incorporating online grammar checkers as a part of their formative assessment combined with metalinguistic reflection, students can develop the capacity for self-correction and become more aware of their language choices. A study has been conducted with undergraduate students of Catalan language in the context of a Translation and Interpreting degree in Barcelona with the aim to demonstrate that integrating this technology with metalinguistic reflection can enhance students' feedback engagement and their ability to recognize and address issues in writing. The results show that most students value the tool's support in learning grammar; however, it often fails to address issues of cohesion and style, requiring the assistance of the instructor to resolve such problems.

Keywords: *technology-mediated feedback; feedback literacy; style awareness; language learning; L1*

Introduction

Automated writing evaluation (AWE) tools are assistive technologies that offer feedback on grammar, spelling, and writing style throughout the writing and revision process. Traditional automated scoring systems have been replaced with instructional tools for in-class use —i.e. sample essays, graphic organizers, dictionaries, and tools for checking spelling and grammar—and they can now help students improve their writing in classroom practice (Karatay & Karatay, 2024). Based on the experiences and perceptions shared by both teachers and students, systems with automated feedback enhance writing performance (Limpo et al., 2020). However, research on how AWE technology is integrated into day-to-day teaching practice and can be used as feedback tools remains scarce (Giessler, 2023).

Grammar checkers, in particular, have become essential tools for reviewing and improving texts in the field of translation and revision. The increasing dependence on these tools has improved their development in recent years, particularly with the integration of AI. The focus of research until now has largely been on evaluating their effectiveness in identifying and correcting errors, in particular, with the use of Grammarly (Abu Qub'a et al., 2024; Alshayban, 2024; Fitria, 2021; Ghufron & Rosyida,, 2018). Most studies on technology-assisted grammar learning have been directed toward second

language acquisition (Bahari & Gholami, 2022), as learners face challenges in acquiring and internalizing the rules of a non-native language. A significant body of research has been carried out on the use of grammar checkers to improve grammar and academic writing skills in English as foreign language (Lipalam et al., 2023). Nevertheless, the use of this technology as a tool to learn grammar and writing skills in the students' L1 has not been sufficiently addressed.

This study aims to fill this gap by investigating the implications for the integration of grammar checkers in teaching grammar and writing skills to future translators into Catalan, looking at their use as tools that can contribute to more autonomous, reflective, and effective learning processes. By using them as a form of feedback together with metalinguistic reflection, the aim is to determine whether grammar checkers can be beneficial to feedback engagement and ultimately facilitate grammar knowledge and style awareness and improve students' confidence in writing.

This study will address the following research questions:

1. How can online grammar checkers be incorporated as AWE feedback tools to help students improve their writing skills?
2. How effective are grammar checkers, when combined with metalinguistic reflection, in improving students' writing skills?
3. How do students perceive the usefulness of these tools after interacting with them and reflecting on their own errors?

Literature Review

Theoretical Framework

Vygotsky's socio-cultural theory of learning and the broader constructivist framework have had a strong influence on language learning (Pathan et al., 2018). This theory posits that learners actively build knowledge through interaction with their environment and social context. The concept of Zone of Proximal Development (ZPD) (Vygotsky, 1978) offers useful insights for teaching grammar and writing skills in higher education. This notion refers to the space between what an individual can achieve on their own and what they can do with the help of someone —or something— with more capabilities, suggesting that in this "zone" learning occurs. Thus, by helping students to learn in their ZPD, instructors can promote powerful learning. The concept of scaffolding, on the other hand, which builds on socio-cultural theory of learning, refers to the process by which teachers —or other more capable peers or, in this study, machines— can provide temporary support to learners to help them accomplish tasks they could not complete independently (Bruner, 1978).

Principles of constructivism often include learning personalization, reflective thinking, problem-solving, relevance to daily life, collaborative learning, discussion and debates, and teacher scaffolding (Mishra, 2023). In constructivism, the teacher's role shifts from a knowledge transmitter to a facilitator and pedagogical models promote active learning. In this context, Constructivist Learning Environments (CLEs) are designed to foster problem-solving and conceptual development, where learners support each other and construct knowledge using resources and tools to solve a real-life problem. For instance, in a BA Thesis Writing course implemented at a Chinese university, a technology-enhanced

CLE using the online management system Moodle successfully supported problem-based learning (Gu et al., 2020).

A key principle of social constructivism is learning through providing and receiving feedback. While the sources extensively detail feedback provided by peers and instructors (scaffolding), the function of a grammar checker directly aligns with the resulting reflective and revision process expected within a CLE (Barak, 2017). Technology is not a replacement for teachers, but grammar checkers can offer immediate feedback that students can use to strengthen their skills. In the context of Vygotsky's ZPD and Bruner's idea of scaffolding, grammar checkers can help bridge the gap between students' grammar and style knowledge and ability, and their potential, by offering suggestions that guide learners toward correct usage as an active, participatory and student-centered guidance. This can be especially helpful for students working at different levels, ensuring that each student obtains appropriate support according to their learning needs.

Grammar checkers can be harnessed to promote constructivist learning in education via different methods, such as creating tailored experiences for students, enabling real-time adjustments for each student's needs, and providing adaptive scaffolding to support students. By using this technology in a constructivist approach, students can engage better in learning activities and instructors can adapt their classes to address each student's level (Gilakjani et al., 2013), rather than only using technology to present information in a more sophisticated manner.

AWE and technology-mediated feedback

There is strong evidence that frequent feedback brings about learning benefits (Black & Wiliam, 1998). When this feedback is combined with effective instruction, it can be very powerful in enhancing learning (Hattie, 2012). In the field of language learning, research shows that providing feedback on grammar can improve writing outcomes (Cavaleri & Dianati, 2016). However, providing regular and personalized feedback on grammar and style is time-consuming, especially in large classrooms. Teachers may struggle to offer timely feedback to every student, especially if they require detailed attention and there are various levels in the classroom. Furthermore, when delivered in groups, the feedback messages might be influenced by how relevant they are perceived to be to oneself or to other group members (Hattie & Timperley, 2007).

AWE tools can help overcome these problems. A major advantage is the immediacy and interactivity of AWE feedback, which can help overcome the time lag shortcomings of traditional teacher feedback (Giessler, 2023). On the one hand, these tools provide real-time feedback and do not require immediate instructor action; on the other, feedback is personalized to each student's needs. Timely feedback is crucial for enhancing student learning and engagement and research suggests that technology-assisted feedback can help students' monitor their progress, obtain timely continuous support, promote their engagement in feedback and develop self-regulated learning strategies (Carless & Winstone, 2023; Gikandi et al., 2011). Grammar checkers, in particular, offer color-coded feedback according to error type, which has been found to be useful for increasing awareness, as it directs learners' focus toward specific errors (Hamid et al., 2018).

However, using this technology on its own does not necessarily enable students to engage in a dialogue that demonstrates their uptake of feedback. Current research emphasizes the need to find strategies that support the long-term integration of AWE feedback into students' writing development (Karatay & Karatay, 2024), as simply receiving automated feedback may not be sufficient to ensure the internalization of corrections. Additionally, studies indicate that drawing on diverse sources of feedback enhances students' writing and learning in a comprehensive manner (Xu & Zhang, 2023). Technology-mediated feedback and teacher feedback can complement each other and enhance students' feedback literacy. The role of the teacher is, thus, to design feedback activities that promote dialogue and critical engagement in order to complement technology.

Metalinguistic reflection

From a social constructivist view, feedback is seen as a process where understanding is formed through dialogue and collaborative construction between participants (Carless & Boud, 2018). The ability of students to read, interpret, and utilize feedback effectively is a crucial skill for improving student learning outcomes. Responding to feedback information involves students engaging with their teachers to discuss how they have understood and utilized the feedback received (de Kleijn, 2023). Metalinguistic reflection, as the process of consciously analyzing and understanding language, including its grammar rules and usage, can help interpret feedback and utilize it in a critical manner. As it involves examining one's language use and the linguistic features contained in a text, it enhances awareness and can foster learning. Actions that take place in this process include reflecting on various aspects of linguistics such as syntax and morphology, verbalizing language rules that allow generalizations, and elicitation via different activities such as journal writing and discussions (Allison, 1998; Simard, 2004). Strategies for implementing reflection can include facilitating metatalk, designing authentic and purposeful activities, scaffolding metalanguage gradually and linking grammar to writing (Myhill et al., 2020).

Research shows that metalinguistic reflection can promote learning and, although it cannot be claimed that metalinguistic awareness leads to mastery, it is believed that it helps students to notice aspects they might not have otherwise identified (Simard et al., 2007). Therefore, such reflection can help students interpret feedback by engaging in a dialogue with the instructor which turns this interaction into a learning opportunity. In practice, instead of just passively accepting a suggestion, it makes students reflect actively on why something might not be correct and discuss and think through the linguistic choices they make.

Grammar checkers as tools for reviewing texts

Grammar checkers have become an essential tool for revising and improving written texts. The increasing reliance on these tools has stimulated research and led to significant advances in their development in recent years, particularly with the integration of AI, which raises questions about the continuing need for human input in the writing and revision process (Moneus & Sahari, 2024; Zanaty, 2024). Despite these advances, grammar checkers still present limitations that prevent them from being fully reliable. As a result, human reviewing and proofreading remain essential at present.

First introduced in the 1980s, these systems have progressed significantly since then. In the 1990s, grammar checkers were introduced in commercial packages, like Microsoft Word. The traditional approach to those grammar checkers involved manually creating grammatical rules, and those integrated into word processors at the time were relatively basic, focusing mainly on spotting spelling errors rather than addressing more complex grammatical or stylistic issues. Rule-based grammar checkers work well for short, well-structured texts, but they struggle with more complex or semantically ambiguous content due to their rigid models (Soni & Thakur, 2018).

Language checking is highly complex due to the numerous variables involved. It is not just about following grammatical rules, but one must also consider factors like register, dialects, and style. Advances in natural language processing (NLP) have led to the development of grammar checkers that improve these aspects. However, unlike rule-based systems, it is difficult to explain the errors that result from these systems, monitor the result, and make changes if necessary. Therefore, hybrid models, which combine rules and machine learning, are believed to be the most suitable for improving system performance (Soni & Thakur, 2018).

One popular online grammar checker, Grammarly, has been found to be a very useful assistance for university students, providing suggestions for corrective vocabulary, sentence rephrasing, and identifying possible misspellings (Daroina et al., 2022). The use of this tool has also been proven to boost students' writing and understanding of grammar rules (Cavaleri & Dianati, 2016), and its role in providing effective feedback has been highlighted (Daroina et al., 2022). Research on systems like Grammarly revealed that students who consistently used this technology performed better (Chang et al., 2021; Tambunan et al., 2022).

The two available Catalan language online checkers, a version of LanguageTool and Softcatalà's language checker, are mostly rule-based with over 8000 rules entered manually (Jaume Ortolà , personal communication, January 9, 2025). AI has recently been introduced in order to detect punctuation issues and improve style by offering correction recommendations. LanguageTool can be installed as an extension in the browser and offers real-time suggestions as long as the text under review is online. It also has a personalized dictionary to which users can add terms if they are registered with an account. The tool spots spelling, grammar and style errors (marked with three different colors) and gives recommendations with a message and a link to an online linguistic resource. On the other hand, Softcatalà, a non-profit organization focused on promoting Catalan-language technologies, offers an online checker that operates similarly to LanguageTool, but can only be used within their website by copy-pasting a text.

Grammar checkers as a supportive tool to enhance students' development

Online grammar checkers, when they are embedded within socio-cultural and constructivist approaches to learning, have shown to have pedagogical potential. Viewed as AWE feedback tools, grammar checkers can provide timely, individualized scaffolding that supports students' writing development and helps address practical challenges in higher education contexts. However, research consistently indicates that automated feedback alone is insufficient to guarantee meaningful learning or sustained improvement in writing skills.

Studies highlight the importance of combining technology-mediated feedback with metalinguistic reflection, dialogue, and teacher guidance to promote active engagement, feedback literacy, and self-regulated learning. While recent advances in AI have enhanced the accuracy and scope of grammar checkers, their limitations underline the continuing need for human input in the writing and revision process. Consequently, grammar checkers should be understood not as replacements for instructors, but as complementary tools whose effectiveness depends on how they are pedagogically integrated. Hence the need for further research into how students engage with grammar checkers alongside reflective practices, and how such integration influences their writing development and confidence.

Methodology

The study was conducted at the Universitat Autònoma de Barcelona, employing purposive sampling which involved a selected group of advanced language learners of Catalan pertaining to a single class (n = 30). Participation in the study was voluntary and anonymous, and students were assured that the information collected was totally confidential and used only for research purposes.

Course context

The aim of the course, which takes place over the first year of the degree in Translation and Interpreting, is to consolidate the students' grammatical competence in Catalan. It enables them to produce and understand texts of a certain complexity to prepare them for direct translation and reverse translation. Once the module is successfully completed, the students are expected to have a proficient level of Catalan (C2 of the Common European Framework of Reference). The group of students present differences regarding their knowledge of the Catalan language when they start. Most of them are native speakers but have different experiences in regard to learning the language based upon their origin and other individual factors. All the students have been through the Catalan education system, in which a C1 level certificate is granted at the end of secondary education. Despite this, it is well known that students arrive at university with various gaps and shortcomings in their knowledge and mastery of the standard Catalan language (Bach & Bernal, 2015). While the aim of this group of students is to achieve a level C2, there is a significant gap between their real knowledge and skills and the goal of the course. Therefore, it is clear that effective strategies are needed to close this discrepancy within a single academic year.

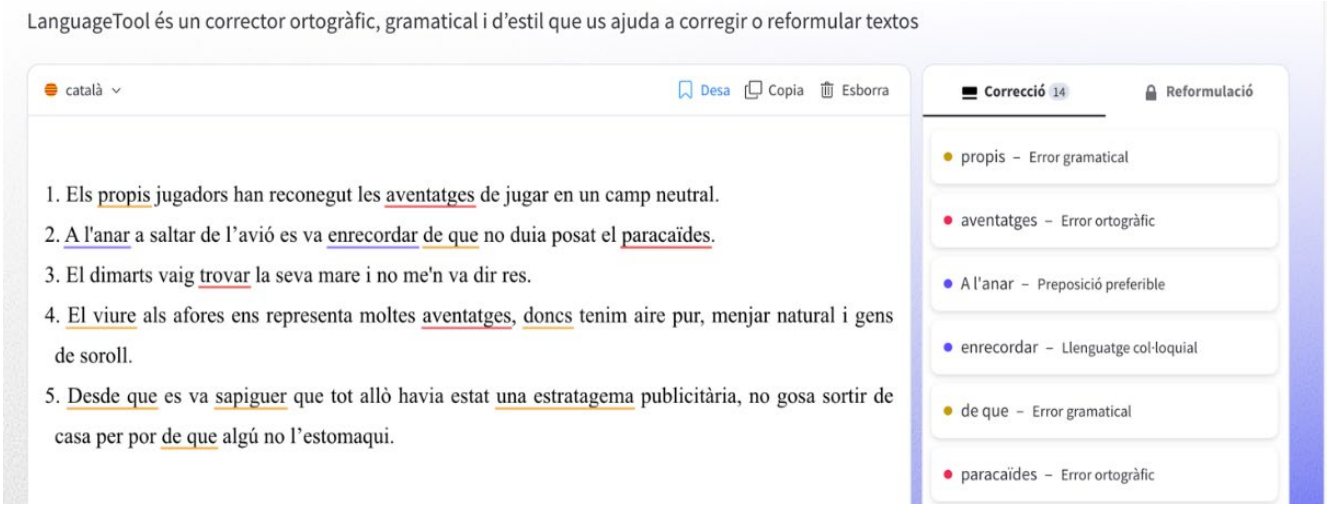
Data Collection

The research used three primary data sources: students' work (document analysis), class observation, and a survey featuring both Likert scale and open-ended questions. Data collection was structured into two main stages (Phase 1 and Phase 2), followed by the final survey distribution.

Phase 1 was designed to familiarize themselves with the tool. At this stage, the students completed a number of error detection exercises. Students practiced over five one-hour sessions, reviewing approximately 25–30 sentences per session that contained grammar, spelling, and style errors. The students reviewed the sentences individually and corrected the errors they spotted independently. Following this, the sentences were reviewed with the assistance of a grammar checker (Figure 1), and false positives and false negatives were identified.

Figure 1

Example of an error detection exercise carried out in class using Language Tool.



Following this, students identified the type of error detected in each case, according to the colors and the suggestions made by the tool, whether the error fell into the category of grammar, spelling or style (grammar = yellow; spelling = red; style = blue).

Students raised some questions during the use of the grammar checker, which allowed for clarifying grammar issues that had been taught in previous sessions, and new grammar points were dealt with arising from other questions. The process of metalinguistic reflection consisted of categorizing the errors according to type, discussing options to implement changes (how one option or the other affects the text) and eliciting rules that apply to each case. No questions about spelling or style were formulated at this stage.

Phase 2 took place at the end of the first semester (December 2024), following five class practice sessions conducted over five consecutive weeks, and aimed to assess whether using the language checker improves learning. Data was collected via a piece of writing: students wrote an essay (approximately 300 words) as part of their first midterm assessment. The test was handwritten and completed independently in one hour. The piece of writing was assessed by the instructor using color-coded feedback, omitting the correct word or phrase, which allowed students to identify and find the right solution for each case. This form of corrective feedback can be referred to as 'metalinguistic feedback', where the teacher provides information or questions related to the correct form without explicitly giving it (Klimova & Pikhart, 2022).

The students received color-coded feedback manually by the teacher, using the same colors as the grammar checker according to error type. An extra color was assigned to flag up Castilian interference (i.e. grammatical and lexical loans from this language, as it is a well-known issue due to language contact) and lexical accuracy (interference and lexical accuracy = green). The instructor collected common errors from the tests and listed them by error type, using this list as a warm-up activity in a feedback session. This activity was similar to the exercises in Phase 1, except that the sentences were taken from the students' own work. After receiving feedback, students worked independently to review and address their mistakes. The teacher moved around the classroom to solve any issues that may have arisen with the feedback and the students completed their corrections independently. While the students where working on their text, the teacher prompted the students who struggled with the corrections to explain the reasons behind their language choices, which encouraged discussion of grammar, word choice, or text structure. Most grammar and spelling issues were addressed without extra help, but some students required the assistance of the instructor for issues relating to cohesion, coherence and style. The papers were returned to the instructor, who reviewed the results and tallied corrected errors, uncorrected errors, and incorrectly corrected errors to identify which error types persisted.

At the end of the experiment, a survey was conducted to gauge the students' opinions towards the use of the tool. The survey was distributed to the students to find out their perception of the use of the grammar checker to help them identify errors and self-correct their work. A total of 26 out of 30 students answered the survey.

Instruments

LanguageTool was employed as the primary instrument for error detection practice in Phase 1. It was selected due to the fact that it is the only freely available Catalan grammar checker that supports online error detection through a browser extension and provides color-coded feedback. Essays, on the other hand, were used as the assessment tool in Phase 2, accompanied by metalinguistic reflection. The instructor provided information or prompts about the correct form without explicitly supplying the word or phrase, which enabled students to identify and determine the appropriate solution themselves.

The final survey utilized a combination of Likert scale questions and one open-ended question. The Likert scale questions presented participants with four statements, and responses contained nominal values in a 5-point ordinal scale ranging from "Strongly disagree" to "Strongly agree"). The open-ended question enabled us to understand the reasons why the tool was useful to the students. Data was analyzed calculating the median and mode (Cohen et al., 2018), which allowed us to identify the most typical attitude towards the grammar checker.

Data Analysis

After counting the number of errors identified in the writing assessments by type, the results show a clear pattern. Spelling and grammar errors have mostly been addressed successfully, with a high rate of correct corrections. In contrast, the majority of errors related to coherence, style, and cohesion remain unaddressed or are corrected incorrectly. Finally, errors associated with Castilian interference

and lack of lexical accuracy present more balanced outcomes, with similar numbers of errors being addressed and left unresolved. The table below presents the number of errors in each assessed exam paper, indicating how many were corrected correctly, corrected incorrectly, or left uncorrected, categorized by type.

Table 1

The Number of Errors in Each Assessed Exam

	S	S	S	G	G	G	C/C/S	C/C/S	C/C/S	I/A	I/A	I/A	TOTAL CORRECTED	TOTAL ERRORS
Exam paper 1				2			3	1					5	6
Exam paper 2	6	1		1	1	2	2	1		1	1		10	16
Exam paper 3	1			2		1	1			1			5	6
Exam paper 4				4				5		4	1		8	14
Exam paper 5		5			3			12					0	20
Exam paper 6		4		1			3	1					4	9
Exam paper 7	3	1		4			1		1	1			9	11
Exam paper 8	2	1				1	1			1		1	4	7
Exam paper 9	13	1			4			4			3		13	25
Exam paper 10	2				1			2					2	5
Exam paper 11	1	1		2			4		1				7	9
Exam paper 12	12			1	5			8		1	1		14	28
Exam paper 13	3			1				4			1		4	9

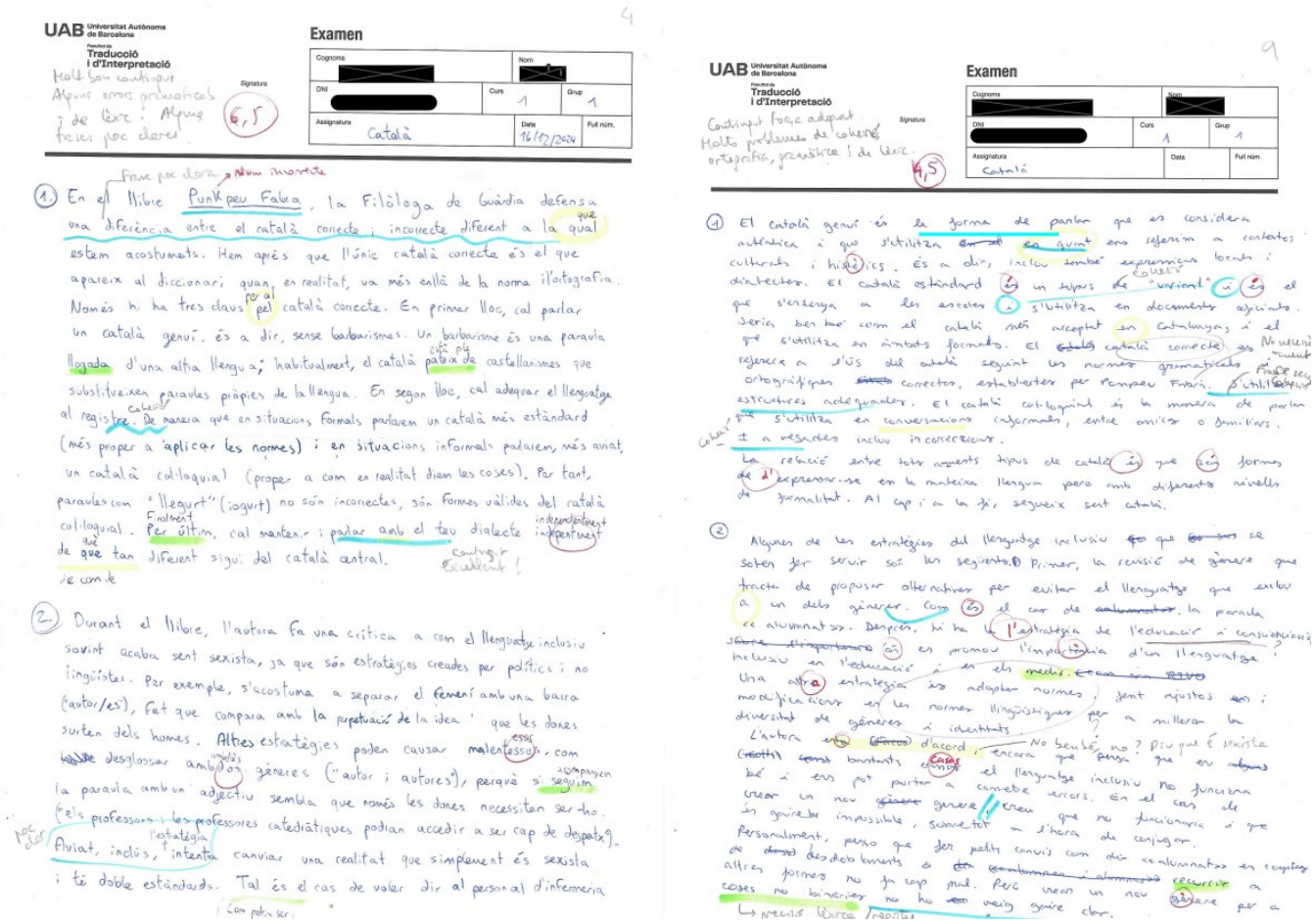
Exam paper 14				1			2			3			6	6
Exam paper 15	10			3		1				1	3		14	18
Exam paper 16	1	1		1			4	1			1		6	9
Exam paper 17		1		4				1					4	6
Exam paper 18	8			1				3		1	2		10	15
Exam paper 19	3			6			2						11	11
Exam paper 20		1		1	1		1			1	2		3	7
Exam paper 21	4			1			3		1		1		8	10
Exam paper 22				3		2	2		1				5	8
Exam paper 23		1		1			4		1	1			6	8
Exam paper 24				1			1	1		1			3	4
Exam paper 25	1			3	1			4					4	9
Exam paper 26	3							6		2			5	11
Exam paper 27	2	3		2		1	2	1	2	1		1	8	15
Exam paper 28				1	1			2	1	1	1		2	7
Exam paper 29	4	1		1		1		2					5	9
Exam paper 30	1			2					2	1			3	6
TOTAL	80	22	0	50	17	9	36	59	10	22	17	2	188	324

Key:
S: spelling (corrected) / S: spelling (uncorrected) / S: spelling (incorrectly corrected)
G: grammar (corrected) / G: grammar (uncorrected) / G: grammar (incorrectly corrected)
C/C/S: coherence / cohesion / style (corrected) / C/C/S: coherence / cohesion / style (uncorrected) / C/C/S: coherence / cohesion / style (incorrectly corrected)
I/A: interference / lexical accuracy (corrected) / I/A: interference / lexical accuracy (uncorrected) / I/A: interference / lexical accuracy (incorrectly corrected)

After the students' intervention, the instructor moved around the classroom and observed how some style and punctuation issues still presented problems, or were simply not addressed, mostly because the alternative required rephrasing the whole sentence (not just a word replacing another) and, in most cases, the students did not know how to correct it without assistance (see examples in Figure 2). By the end of the session, most students had not addressed all the issues in their papers, and only a few showed interest in rewriting the text to improve its style.

Figure 2

Two exam papers with corrected answers by students



Results

The experiment shows that the grammar checker was unable to identify most issues related to cohesion, coherence, and style, which are typically more context-dependent. As a result, students often required additional guidance from the instructor to recognize these problems and to reflect on why certain sentences were inappropriate or unclear. This limitation highlights the difficulty automated tools face in addressing higher-level writing features that depend on textual organization and stylistic adequacy.

In contrast, grammar and spelling errors were largely corrected successfully. In many cases, students were able to identify and resolve these issues independently, while in others they relied on LanguageTool or supplementary resources such as the Catalan grammar and dictionary available online. This suggests that rule-based and surface-level errors are more readily detected and corrected by automated tools, particularly when learners are already familiar with the underlying grammatical principles.

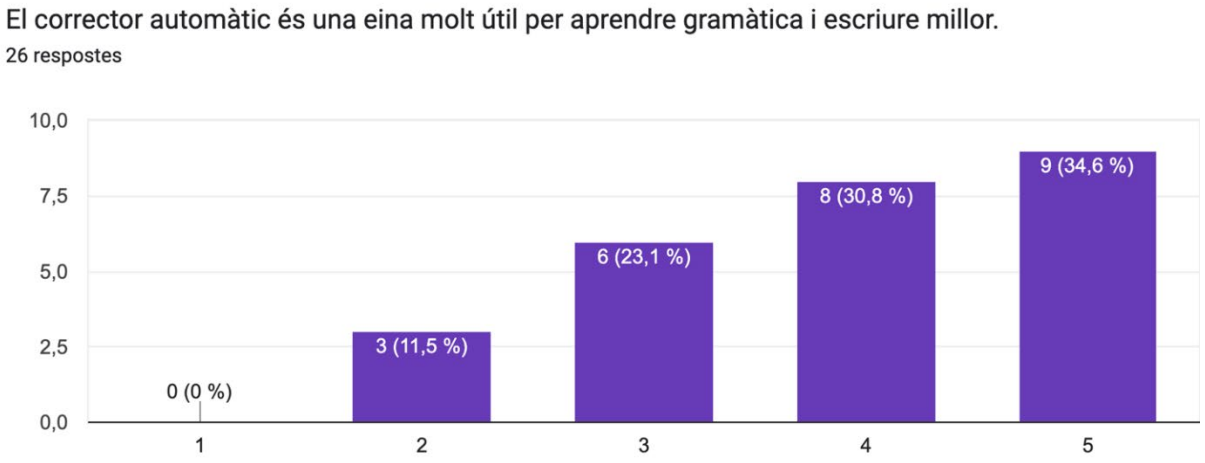
The combination of AWE and metalinguistic reflection appeared to foster a degree of learner autonomy. Some students were able to engage critically with the feedback provided, evaluate their own linguistic choices, and implement appropriate corrections without consulting external sources. Nevertheless, the findings also indicate that automated tools alone are insufficient for developing advanced writing skills, and that instructor-mediated feedback remains essential, especially for addressing complex aspects of written discourse.

The survey, distributed at the end of the session, provided some insights regarding the students' opinion on the tool. When asked about how useful the grammar checker was for identifying errors and understanding the nature of such an error, the central tendency was that the students agreed that it was useful (median = 4, mode = 4). When asked about the use of colored marks to identify the type of error, the same result was obtained (median = 4, mode = 4). However, a significant number of students (19.2%) gave a score of 1 and 2 to color-coding, indicating that it may not be an effective strategy for certain learners (see Appendix A).

One last Likert scale question (Figure 3) was designed to gather the students' opinions on the effectiveness of the grammar checker in helping them learn grammar and improve their writing skills. The central tendency was that the students agreed that the tool was very useful (median = 5, mode = 4). While the lowest score had zero answers, an 11.5% gave a score of 2 to this statement, indicating that the tool does not work for everyone.

Figure 3

Answer distribution results for the following statement: "The grammar checker is a very useful tool to learn grammar and writing skills"



In an open-ended question, students were asked to explain their response, providing a range of perspectives on the reasons they found the tool useful to their learning. On the one hand, most comments referred to the tool's ability to identify errors that students would not have noticed otherwise; on the other, the students highlighted additional benefits of the grammar checker: saving time and practicing (promoting autonomy and providing timely feedback), identifying error types, understanding the reasons behind errors, and visually retaining grammar rules (using metalinguistic reflection). One response was negative, mentioning their own inability to determine whether the suggestions provided by the tool were correct, and three more were positive but commented on the tool's incapacity to detect many errors.

Conclusion

This small-scale study explored the pedagogical potential of combining automated writing evaluation tools with metalinguistic reflection to support the development of grammar and writing skills among undergraduate students of Catalan in a Translation and Interpreting degree. Grounded in a socio-constructivist framework, the research examined how grammar checkers can function as feedback tools when combined with reflective and instructor-mediated learning practices. The findings suggest that grammar checkers, when integrated into formative assessment and paired with guided metalinguistic reflection, can contribute positively to students' feedback engagement and foster a degree of learner autonomy. In particular, the tool proved effective in helping students identify and correct spelling and grammar errors, which are more surface-level and rule-governed. Many students were able to address such errors independently, often without consulting additional resources, which suggests that automated feedback can support self-correction and reinforce existing linguistic

knowledge. At the same time, the study shows the limitations of grammar checkers as AWE tools. Issues related to cohesion, coherence and style were frequently overlooked by the tool and posed challenges for students. These findings show the difficulty automated systems face in dealing with writing features that are context-dependent and require discourse-level awareness. Consequently, instructor-mediated feedback and classroom dialogue remained essential for helping students interpret feedback, reflect on their linguistic choices, and develop more advanced writing skills.

Students' perceptions further reinforce these conclusions. Overall, participants valued the immediacy, clarity, and time-saving nature of the grammar checker, particularly its color-coded feedback and its role in making errors more visible. However, the variability in responses also suggests that such tools do not benefit all learners equally and that their effectiveness depends on students' ability to critically evaluate suggestions and engage in reflection. This reveals the importance of providing scaffolding to support students in using automated feedback effectively.

All in all, the results suggest that grammar checkers should not be viewed as replacements for teacher feedback, but rather as complementary tools that can extend learning opportunities when integrated into pedagogical practice. Despite the small sample size of the study, findings contribute to a body of research on technology-mediated feedback in L1 contexts and offer practical insights for instructors seeking to integrate automated tools into teaching reflective and student-centered writing.

Challenges and Future Directions

The study also revealed some challenges associated with the use of grammar checkers. While these can be very helpful, there is a risk of students becoming overly reliant on the tool. If students consistently depend on grammar checkers without developing their own critical thinking skills or understanding of language rules, they might miss out on the deeper learning process. As already pointed out by the students themselves, grammar checkers might not catch every mistake in a text, or other nuances such as register, which could lead them to overlook these aspects in their writing. Therefore, it is necessary to take into account the potential implications of using grammar checkers to avoid hindering the development of independent writing skills.

While the grammar checker tested is useful, there is room for improvement in its development as an AI-powered tool. Therefore, a potential future research direction could explore how advancements in this area might enhance students' written expression skills. This is particularly relevant as the study's findings show that many students continue to struggle with self-correction when it comes to cohesion, coherence and style.

References

Abu Qub'a, A., Yousef Abu Guba, M.N., & Fareh, S.I. (2024). Exploring the use of grammarly in assessing English academic writing. *Heliyon, 10*. https://doi.org/10.1016/j.heliyon.2024.e34893

Allison, D. (1998). Investigating learners' course diaries as explorations of language. *Language Teaching Research, 2* (1), 24-47. https://doi.org/10.1177/136216889800200103

Alshayban, A. (2024). A comparative study of the error-detection accuracy of Grammarly and Microsoft Word Editor in formal English writing. *World Journal of English Language,14*(5), 535-544. https://doi.org/10.5430/wjel.v14n5p535

Bach, C., & Bernal, E. (2015). Percepció i realitat del (des)coneixement de la normativa de la llengua catalana en arribar a la universitat. Estudi sobre els alumnes de primer de traducció i de llengües aplicades a la UPF. *Revista de Llengua i Dret, 64*, 156-170.

Bahari, A., & Gholami, J. (2022). A systematic review of current research on affordances and challenges of technology-assisted grammar learning. *Computer Assisted Language Learning, 32*(1), 125-148.

Barak, M. (2017). Cloud pedagogy: Utilizing web-based technologies for the promotion of social constructivist learning in science teacher preparation courses. *Journal of Science Education and Technology*, *26*(5), 459-469.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Bruner, J. S. (1978). The role of dialogue in language acquisition. In A. Sinclair, R. J. Jarvelle & W. J. M. Levelt (Eds.), *The child's concept of language* (pp. 241-256). Springer-Verlag.

Cavaleri, M., & Dianati, S. (2016). You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students. *Journal of Academic Language and Learning, 10*(1), A223-A236.

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315-1325. https://doi.org/10.1080/02602938.2018.1463354

Carless, D., & Winstone, N. (2023). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in higher education*, *28*(1), 150-163. https://doi.org/10.1080/13562517.2020.1782372

Chang, T., Li, Y., Huang, H., & Whitfield, B. (2021). Exploring EFL students' writing performance and their acceptance of AI-based automated writing feedback. Paper presented at the ACM International Conference Proceeding Series, 31-35. https://doi.org/10.1145/3459043.3459065

Cohen, L., Manion, L. & Morrison, K. (2018). *Research Methods in Education*, 8th ed. Routledge.

Daroina, A., Febriani, W.E., Aulianisa, A., Fadlia, W.A., Zuhri, P.K., Zuhri, K.H., & Purwokerto (2022). Systematic literature review: Grammarly as a medium in analyzing grammar for university students. *Conference on English Language Teaching*.

de Kleijn, R. A. (2023). Supporting student and teacher feedback literacy: an instructional model for student feedback processes. *Assessment & Evaluation in Higher Education*, *48*(2), 186-200. https://doi.org/10.1080/02602938.2021.1967283

Fitria, T. N. (2021). Grammarly as AI-powered English writing assistant: Students' alternative for writing English. *Metathesis: Journal of English Language, Literature, and Teaching, 5*(1), 65-78. https://doi.org/10.31002/metathesis.v5i1.3519

Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & education*, *57*(4), 2333-2351. https://doi.org/10.1016/j.compedu.2011.06.004

Gilakjani, A., Leong, L., Ismail, H. M. (2013). Teachers' Use of Technology and Constructivism. *International Journal of Modern Education and Computer Science (IJMECS), 5*(4), 49-63. https://doi.org/10.5815/ijmecs.2013.04.07

Ghufron, M. A. & Rosyida, F. (2018). The role of Grammarly in assessing English as a Foreign Language (EFL) writing. *Lingua Cultura*, *12*(4), 395-403. https://doi.org/10.21512/lc.v12i4.4582

Giessler, R. (2023). EFL writers' cognitive engagement with AWE feedback. *Language Awareness*, *33*(2), 428-445. https://doi.org/10.1080/09658416.2023.2269088

Gu, P., Zhang, Y., & Gu, H. (2020). Creating a technology-enhanced constructivist learning environment for research ability development in a BA Thesis Writing course. *Computer Assisted Language Learning*, *33*(5-6), 538-566.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112. https://doi.org/10.3102/003465430298487

Hamid, H. A., Nasri, N. F., & Ghazali, N. (2018). Colours as a form of corrective feedback in EFL learners' writing. *GEMA Online Journal of Language Studies, 18*(4) 106-123. http://doi.org/10.17576/gema-2018-1804-08

Karatay, Y., & Karatay, L. (2024). Automated writing evaluation use in second language classrooms: A research synthesis. *System*, 103332. https://doi.org/10.1016/j.system.2024.103332

Klimova, B., & Pikhart, M. (2022). Application of corrective feedback using emerging technologies among L2 university students. *Cogent Education*, *9*(1). https://doi.org/10.1080/2331186X.2022.2132681

Limpo, T., Nunes, A., & Coelho, A. (2020). Introduction to the special issue on technology-based writing instruction: A collection of effective tools. *Journal of Writing Research*, *12*(1), 1-7. https://doi.org/10.17239/jowr-2020.12.01.01

Lipalam, A., Alivio, R. M. B., Sollano, J. Q., Cherish, R. M. T. T., Villarosa, A. O. T., & Tirol, S. L. (2023). Grammar checkers to boost students' academic writing proficiency in English. *International Journal of Multidisciplinary Research and Publications, 6*(3), 153-159.

Mishra, N. R. (2023). Constructivist approach to learning: An analysis of pedagogical models of social constructivist learning theory. *Journal of research and development*, *6*(01), 22-29.

Myhill, D. A., Watson, A., & Newman, R. (2020). Thinking differently about grammar and metalinguistic understanding in writing, *Bellaterra Journal of Teaching & Learning Language & Literature, 13*(2), e870. https://doi.org/10.5565/rev/jtl3.870

Moneus, A. M., & Sahari, Y. (2024). Artificial intelligence and human translation: A contrastive study based on legal texts, *Heliyon, 10*(6), 1-14. 10.1016/j.heliyon.2024.e28106

Pathan, H., Memon, R. A., Memon, S.A., Khoso, A. R., & Bux, I. (2018). A critical review of Vygotsky's socio-cultural theory in second language acquisition. *International Journal of English Linguistics, 8*, 232-236. 10.5539/ijel.v8n4p232

Simard, D. (2004). Using diaries to promote metalinguistic reflection among elementary school students. *Language Awareness, 13*(1), 34–48. https://doi.org/10.1080/09658410408667084

Simard, D., French, L., & Fortier, V. (2007). Elicited metalinguistic reflection and second language learning: Is there a link?. *System, 35*(4) 509–522. https://doi.org/10.1016/j.system.2007.06.004

Soni, M., & Thakur, J. S. (2018). A systematic review of automated grammar checking in English language. arXiv preprint arXiv:1804.00540.

Tambunan, A. R. S., Andayani, W., Sari, W. S., & Lubis, F. K. (2022). Investigating EFL students' linguistic problems using Grammarly as automated writing evaluation feedback. Indonesian Journal of Applied Linguistics, *12*(1), 16-27. https://doi.org/10.17509/IJAL.V12I1.46428

Vygotsky, L. S. (1978). *Mind in Society: Development of Higher Psychological Processes* (M. Cole, V. Jolm-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Xu, L., & Zhang, T. (2023). Engaging with multiple sources of feedback in academic writing: Postgraduate students' perspectives. *Assessment & Evaluation in Higher Education*, *48*(7), 995-1008. https://doi.org/10.1080/02602938.2022.2161089

Zanaty, D. (2024). The future of human translation in the artificial intelligence era. *Delta University Scientific Journal, 7*(2), 257-274. 10.21608/dusj.2024.320340.1089

About the Authors

Helena Borrell Carreras, Lecturer, Autonomous University of Barcelona, helena.borrell@uab.cat

Appendix A. Survey (Likert Scale Questions)

La pràctica amb el corrector m'ajuda a identificar els diferents tipus d'error que hi ha en un text (*Practicing with the grammar checker helps me identify the different types of errors that are in a text*)	Les marques de colors aplicades a la correcció del meu text m'han ajudat a identificar el tipus d'error i corregir-lo adequadament sense ajuda externa (*Color-coded marks applied to my text allowed me to identify the type of error and correct it properly without external help*)	Quan faig servir el corrector automàtic prenc decisions que m'ajuden a entendre per què hi ha un error (*When I use the grammar checker I make decisions that help me understand why there is an error*)	El corrector automàtic és una eina molt útil per aprendre gramàtica i escriure millor (*The grammar checker is a very useful tool to learn grammar and writing skills*)
4	3	5	5
4	2	4	5
5	4	4	5
4	3	3	3
3	3	4	2
5	3	4	4
5	3	4	5
4	5	4	5
4	5	4	4
4	5	5	4
5	5	2	2
4	5	4	3
4	4	4	4
5	4	4	4
5	4	4	3
3	4	4	4
5	4	4	5
4	1	4	3

3	3	3	4
5	5	5	5
4	4	4	4
3	5	5	5
3	1	3	5
4	2	4	3
3	4	4	3
1	1	2	2

Median: 4	Median: 4	Median: 4	Median: 4
Mode: 4	Mode: 4	Mode: 4	Mode: 5