



GenAI in Language Teaching, Learning, and Assessment: Stakeholders Insights from Two Undergraduate Language Programs¹

Daniel Murcia 

The Pennsylvania State University
Pennsylvania, USA

María Fernanda Jaime-Osorio 

Universidad Surcolombiana
Neiva, Colombia

Luis Felipe Jaramillo-Calderón 

Universidad Tecnológica de Pereira
Pereira, Colombia

Yimmy Alexander Hoyos Pipicano 

Secretaría de Educación del Huila
Algeciras, Colombia

1/29

Abstract

This qualitative exploratory study delves into stakeholders' uses, ethical considerations, and perceived challenges and opportunities about generative technologies powered by artificial intelligence for teaching, learning, and assessment processes. To do this, we draw on insights from 195 students, 35 professors and the two chairs from two Colombian undergraduate English language teaching programs. We collected data through open-ended questionnaires and focus groups, which were analyzed through content analysis. Findings show a growing usage and acceptance of generative technologies powered by AI among the academy members, with emerging strategies to foster learning, teaching, and assessment processes. Yet, stakeholders pose a number of threats and ethical concerns regarding the potential consequences of these technologies in the long run, including the loss of fundamental abilities, trust, and authenticity

¹ Research article

among stakeholders. Thus, they make a call to embrace a critical/collaborative approach and to design democratic policies that may enlighten the path to integrate GenAI in language education and assessment.

Keywords: Higher Education; GenAI; ELT; Language Assessment; Artificial intelligence.

Resumen

La IAGen en la enseñanza, el aprendizaje y la evaluación de lenguas: Percepciones de grupos de interés en dos programas de pregrado en idiomas

Este estudio exploratorio cualitativo profundiza en los usos, consideraciones éticas, retos y oportunidades percibidos por los actores involucrados sobre las tecnologías generativas impulsadas por inteligencia artificial para los procesos de enseñanza, aprendizaje y evaluación de lenguas. Por medio de cuestionarios abiertos y grupos focales, los cuales fueron analizados mediante análisis de contenido, se compiló una muestra de 195 estudiantes, 35 profesores y dos directores de dos programas de pregrado en enseñanza del inglés. Los hallazgos muestran un creciente uso y aceptación de estas tecnologías entre los miembros de la academia, con estrategias emergentes para fomentar los procesos de aprendizaje, enseñanza y evaluación. Sin embargo, se plantean amenazas y preocupaciones éticas como la pérdida de habilidades fundamentales, confianza y autenticidad entre los actores involucrados. Se hace un llamado a adoptar un enfoque crítico y colaborativo, así como a diseñar políticas democráticas para integrar la IA generativa en la educación y evaluación de idiomas.

Palabras Clave: Educación superior; IAGen; ELT; Evaluación de lenguas; Inteligencia artificial.

Résumé

L'IA Générative dans l'enseignement, l'apprentissage et l'évaluation des langues :

Perspectives des parties prenantes de deux programmes de premier cycle de langues

Cette étude de cas qualitative et exploratoire approfondit les usages, les considérations éthiques, les défis et opportunités perçus par des acteurs impliqués concernant les technologies génératives alimentées par l'intelligence artificielle dans les processus d'enseignement, d'apprentissage et d'évaluation des langues. Au moyen de questionnaires ouverts et de groupes de discussion, qui ont été analysés par une analyse de contenu, nous avons constitué un échantillon de 195 étudiants, 35 professeurs et deux directeurs de deux programmes de premier cycle en enseignement de l'anglais. Les résultats montrent une utilisation croissante et une acceptation de ces technologies au sein de la communauté académique, avec des stratégies émergentes pour favoriser les processus d'apprentissage, d'enseignement et d'évaluation. Toutefois, des menaces et des préoccupations éthiques sont soulevées, telles que la perte de compétences fondamentales, de la confiance et de l'authenticité entre les acteurs impliqués. Les acteurs appellent à l'adoption d'une approche critique et collaborative, ainsi qu'à l'élaboration des politiques démocratiques pour intégrer l'IA générative dans l'éducation et l'évaluation des langues.

Mots-clés : Enseignement supérieur ; IAGén ; ELT ; évaluation de langues ; Intelligence artificielle.

ABOUT THE AUTHORS

Daniel Murcia

He is a Fulbright MinCiencias PhD scholar in the Applied Linguistics program of the College of Liberal Arts, The Pennsylvania State University. His academic work and research interests include Language Assessment, Automated Writing Assessment, and Discourse Analysis within the field of Artificial Intelligence.

Email: danielmurcia@psu.edu

María Fernanda Jaime Osorio

She holds an M.A. in Teaching English as a Foreign Language from Universidad Internacional Iberoamericana. Her research interests revolve around the pedagogical practices of in-service teachers and the use of ICTs in ELT. She is currently an associate professor and the vice provost for academic affairs at Universidad Surcolombiana.

Email: mariafernanda.jaime@usco.edu.co

Luis Felipe Jaramillo Calderón

He is a professor and researcher from Universidad Tecnológica de Pereira. He holds an M.A. in Bilingual Education from the same university. He is a member of the Poliglosia research group. His research interests are Bilingual Education, Language Assessment, emerging technologies in education and literacy development.

Email: luisfe00798@utp.edu.co

Yimmy Alexander Hoyos Pipicano

He holds an M.A. in English Didactics from Universidad Surcolombiana and a B.A. in English Language Teaching from Universidad Tecnológica de Pereira. He is currently a full-time teacher at a public school in Huila and is interested in exploring education for social justice in ELT. He is part of the ILESEARCH research group ascribed to Universidad Surcolombiana.

Email: yimmy.hoyos@sedhuila.edu.co

HOW TO CITE THIS ARTICLE

Murcia, D., Jaime-Osorio, M. F., Jaramillo-Calderón, L. F. y Hoyos-Pipicano, Y. A. (2025). GenAI in Language Teaching, Learning, and Assessment: Stakeholders Insights from Two Undergraduate Language Programs. *Lenguaje*, 53(1S), e20314387. <https://doi.org/10.25100/lenguaje.v53i1S.14387>

INTRODUCTION

The use of Generative Artificial Intelligence (GenAI) technologies is a current and concerning topic in education and the language assessment field. GenAI in higher education has transformed language practices and teaching routines by offering alternative modes of interaction and language production (Bond et al., 2024; Salas-Pilco & Yang, 2022). This has displayed challenges and opportunities to explore within the language assessment area (Guio et al., 2021; Strasser, 2023). In this regard, very early in the popularization of GenAI tools like ChatGPT (OpenAI), diverse reactions emerged from the field. Some stakeholders were eagerly open to the possibilities, while others worried about the potential risk of such technology tools.

For instance, Bhutoria (2022) highlights the benefits of artificial intelligence to enhance collaboration, personalize learning, and foster technology use. However, there are also potential threats related to academic integrity, assessment issues, and data protection (Li et al., 2024). While both sides have strong arguments for their position, it is still unclear how teachers, students, and administrative staff of undergraduate language programs (Henceforth: stakeholders) of Colombian public universities use, perceive, or feel about GenAI for teaching, learning, and assessment (Salas-Pilco & Yang, 2022; Sued, 2024). There is a need to explore stakeholders' approaches to GenAI technologies, especially in developing countries like Colombia, in which grounded policies and adequate teaching environments remain as obstacles for technology normalization (Hoyos-Pipicano & Jaime-Osorio, in press). Most researchers in the field suggest the development of more situated research, and that call-to-action becomes a rationale for this study. Therefore, delving into such practices, attitudes, and insights is urgent for informing and updating language program curricula and language policies.

This need is especially important for language assessment processes. Although the increasing use of GenAI by students poses significant challenges for traditional ways to assess students, innovative uses of these technologies may offer alternative and brand-new ways not only to teach and learn languages but also to assess language learning. This article grounds the inclusion of learning and teaching processes in a special issue on language assessment since 1) empirical data from stakeholders, when inquired about GenAI, would frequently relate to learning, teaching and assessment practices in tandem, 2) assessment is viewed as a holistic/ongoing process that is present in education not only in formal-explicit ways, 3) it is also viewed as something that shapes and is shaped by teaching and learning processes; they have a symbiotic relationship (Panizzon, 2019).

Bearing in mind the previous panorama, we carried out an exploratory case study with students, professors and directors from two undergraduate language programs to answer the following research question: How do university stakeholders from two Colombian undergraduate language programs approach GenAI technologies in the context of English language learning, teaching and assessment? To answer this question,

we proposed three research objectives: (a) to describe stakeholders' perceptions and uses for learning, teaching, and assessment processes; (b) to identify their ethical considerations regarding language education and originality issues; (c) to determine stakeholders' reflections on potential challenges and opportunities.

In the following section, we elaborate on the main constructs guiding this study.

Disruptive Technologies

In 1995, Joseph Bower and Clayton Christensen coined the term disruptive technologies as those that cause a relevant change and abruptly interrupt the way in which industries, companies, and consumers operate by providing a simpler, cheaper, or more convenient alternative to existing products or services. Some examples are computers, the worldwide web or smartphones, which have been incorporated into the ELT field (Juárez-Díaz & Perales, 2019; Rokhila et al., 2023; Serafimovska, 2023). Consequently, these products and services have given an entrance to new approaches to education, such as computer-assisted language learning (CALL), mobile-assisted language learning (MALL), blended learning (BL), virtual learning (VL), remote teaching (RT), among many others. It is worth noting that technologies that once were disruptive tend to follow a path of normalization amongst users, merging into their everyday lives (Bax, 2003).

The literature shows that the integration of technology into English Language Teaching (ELT) has yielded predominantly positive outcomes in most cases. Serafimovska (2023) studied the effects of blended learning and smart technologies for ELT in Macedonia, reporting that the majority of teachers (19=95%) integrated technologies despite lacking technological infrastructure and that such integration enhanced teacher-student interaction and material diversification. In Colombia, while Gómez-Orjuela (2021) found that blended learning failed to enhance learners' communicative skills of fifty EFL learners from a university, Juárez and Perales' (2019) concluded that using technologies fosters English language learning, drawing on 21 empirical studies conducted in Japan, Turkey, and Colombia, to mention some.

In a similar vein, the literature suggests that the COVID-19 pandemic enhanced the use of disruptive technologies around the world. Cinganotto (2023) found that teachers in Italy successfully motivated and engaged students through online learning. In Indonesia, teachers resorted to WhatsApp, YouTube, and Google forms (Kusuma, 2022), and in Colombia, teachers learned to use video conferencing software and educational apps (Castañeda-Trujillo, 2022; España, 2022). However, in the post-pandemic era, there are different positions regarding the role that technologies play. While some authors argue that "it's clear that technology will continue to play a pivotal role in shaping how we learn and teach languages" (Rokhila et al., 2023, p. 999), others contend that technology use was a spur of the moment and that macro infrastructure and policy issues hinder technology normalization (Juárez-Díaz & Perales, 2021; Villegas-

Torres, 2022). All in all, stakeholders keep incorporating technologies for education; thus, the latest disruptive technology that has caught their attention is artificial intelligence (AI) that is capable of generating texts, images, or other data, in other words, GenAI.

GenAI in Education

Chiu et al. (2023) state that AI refers to “the ability of a digital machine to perform tasks commonly associated with intelligent beings... such as computer vision, speech, machine learning, big data, and natural language processing” (p. 1). Regarding the latter task, Pelletier et al. (2023) highlight that GenAI can “create text, sounds, and images that mimic human creations with considerable adaptability and accessibility” (p. 18). Examples of GenAI applications are ChatGPT, Gemini, PhotoRoom, Perplexity, Canva, Beta, Pixlr, and Gamma. The number of applications keeps growing without precedent. In turn, GenAI in education (AIEd) represents a paradigm shift in the field through the possibilities of personalizing content delivery and engaging learners (Singha et al., 2024). Additionally, the authors highlight that these technological advancements support assessment by offering immediate feedback, catering to diverse learning styles, and enabling highly personalized experiences. Then, the authors conclude that AI enhances language assessment through adaptive testing, instant data analysis, and automated scoring, contributing to more accurate and insightful assessments.

6/29

Similarly, this shift reflects on increasing uses of GenAI in higher education (Strasser, 2023), with the main repercussions for assessment processes. Kurni et al. (2023) point out that GenAI uses in education are related to feedback, questioning, and example generation. To illustrate, Ricart (2024) found that ChatGPT was useful, when supervised by professors, to provide immediate writing feedback to 98 university students from Spain. In the same line, Bekou et al. (2024) found that ChatGPT implied “personalized learning [also suggested by Kurni et al., 2023], teachers’ professional development enhancement, and access to authentic language resources” (p. 104), as seen by 62 ELT Moroccan teachers. Bekou et al. also reported challenges regarding “concerns about correctness, cultural appropriateness, and over-reliance on technology” (p.104). Similarly, Makeleni et al. (2023) identified academic dishonesty and the laziness of teachers and students as challenges. Therefore, the international literature shows that stakeholders are aware of both opportunities and challenges regarding GenAI, particularly those related to assessment.

In Colombia, however, AI in education was reported to be scarce as the incorporation of these technologies into education is still recent and challenging (Galaczi, 2023; Ministerio de Ciencias, Tecnología e Innovación [Minciencias], 2024), especially in developing countries (Makeleni et al., 2023). Claro et al. (2024) studied AIEd and gamification with 45 EFL teachers from different educational levels. The authors reported benefits regarding enhancement in teachers’ didactic strategies and assessment practices,

and students' confidence and engagement. Nevertheless, they pointed out the importance of training and professional development for AIED to successfully take place. In this regard, Jamshed et al. (2024) found that EFL learners (258) from Saudi Arabia held positive attitudes towards GenAI despite finding it challenging to use this tool. Hence, teachers and students require specialized preparation to address the challenges of even "stronger" AI.

GenAI and Assessment

An essential GenAI effect relates to assessment. In this study, "assessment in the classroom is not merely a measure of performance and achievement towards learning outcomes, but it should also promote and facilitate the learning process" (Trace, 2021, p. 317). In other words, assessment implies an ongoing process to identify learning opportunities (Carless, 2023). The way teachers assess is a complex phenomenon that shapes and is shaped by their views about language teaching and learning and by different cultural, political and ideological forces, highlighting that assessment practices are not merely technical decisions but a much more complex process (Inbar-Lourie & Donitsa-Schmidt, 2009). In turn, this formative view broadens the possibilities to brand-new assessment processes inspired by the use of GenAI. As Trace (2021) synthesizes, assessment as learning can have multiple forms, such as traditional tools, longitudinal and personalized assessment, alternative and democratic assessments, co-developed rubrics, self- and peer-assessment, and individualized, process-based and dynamic assessments, all of which emphasize learning and formative feedback. These varied forms of assessment may be developed or executed with the help of GenAI tools, for instance, automated essay scoring, speech recognition, natural language processing, and adaptive testing (Hidayah et al., 2023).

In this line, Edmett et al. (2023) explored the views of 1348 teachers from 118 countries (18% from the Americas). Findings showed that teachers used AI for automated grading (22%) and felt that this practice did not necessarily change the nature of assessment. Likewise, teachers acknowledged AIED opportunities to diagnose learners' English language proficiency so that they would receive remedial action. However, they also warned about potential damage from this feature, as learners could be classified as winners or losers. In a similar vein, Bewersdorff et al. (2023) analyzed the potential of large language models (LLM) to identify students' errors in protocols of 65 students. The authors found that AI successfully identified simpler tasks, yet it failed to provide feedback regarding complex ones. Hence, teachers anticipate positive outcomes from the integration of GenAI in assessment practices, but make a case not to use these technologies to segregate students or to assign AI tasks that may be too complex.

Critical Language Assessment in the Age of GenAI

GenAI has been so disruptive to common education and assessment practices that many stakeholders and scholars have identified the need to rethink assessment itself before current practices render ineffective. These changes should go beyond discursive strategies and embrace structural changes (Corbin et al., 2025) and even the very nature and purpose of assessment itself (Bozkurt et al., 2024). On the one hand, while discursive strategies, such as adding AI-usage acknowledgements, or specifying whether to use or not AI, are important, they are all reliant on students' compliance. This is why Corbin et al. (2025) suggest rather emphasizing assessment procedures where teachers have control over actual AI-usage or actual student learning stage. Additionally, some of GenAI's misuses, such as using it to reduce cognitive engagement in a writing task (an easy way out), stem from education's overemphasis on grades of end products over actual learning processes. This is why it is worthwhile to explore assessment practices that shift the emphasis to the process, such as with computerized dynamic assessment procedures (Poehner et al., 2015), or the actual effort of students, such as labor-based grading contracts (Inoue, 2023).

8/29

METHODOLOGY

This exploratory case study addresses how university stakeholders from two Colombian undergraduate language programs approach GenAI technologies for English language learning, teaching and assessment. Accordingly, case studies enable in-depth explorations of complex phenomena (Stake, 2006), like GenAI in ELT teacher education. Additionally, this study has a delimited *unit of analysis*: students, professors and directors from the two undergraduate language programs, aligning with a case study in that “the unit of analysis, not the topic of investigation, characterizes a case study” (Merriam & Tisdell, 2016, p. 38). Therefore, we are specifically interested in what stakeholders think and do with GenAI in regards to their roles as pre-service English language teachers, teacher educators, and directors of ELT undergraduate programs. Next, we explain why the two sites belong to a single case, as well as data collection and analysis methods.

Context and Population

We conducted the present study in contexts that share similar socioeconomic backgrounds and institutional horizons. To illustrate, they are public Colombian universities located in medium-sized cities of the Andean region that are usually overlooked in the national panorama, as much of the attention focuses on larger higher education institutions. Furthermore, they serve mainly a low social strata population from urban and rural areas. Thus, the programs' missions emphasize students'

development of critical, humanistic, integral and research capabilities to address educational issues from the ELT perspective. Put simply, this case study offers insights about GenAI from stakeholders of two underexplored higher education institutions.

In regard to participants, we invited a comprehensive sample from both programs who became a single unit of analysis, given the above general similarities. In alignment with a case study, we resorted to purposeful sampling as it is useful “to find people with certain traits that are important to the study” (Rahman, 2023, p. 49), that is, stakeholders from the two teaching education programs. Ultimately, the sample included (n=195) students who might endure socioeconomic challenges and came from central and peripheral areas. Also, it involved (n=35) professors and (n=2) program chairs with more than fifteen years of teaching experience with master’s and PhD degrees in TEFL.

Data Collection

We collected data through three open-ended questionnaires and two focus group interviews (Creswell & Creswell, 2018). A sample is provided in Appendix A. The ethical considerations of this study included obtaining informed consent from participants for voluntary participation, ensuring confidentiality, and anonymity. For instance, the open-ended questionnaires included an overall description of the study and displayed a link to a detailed consent form, by which they confirmed voluntary participation. The student questionnaire was sent from the program director's mail, ensuring student anonymity. While some students could potentially be part of the researchers’ classrooms, participation in the study did not affect their academic process.

The first data collection instrument was open-ended questionnaires that enquired about different themes. The student questionnaire (five questions) inquired about autonomous and guided use of GenAI in their education process and perceived possibilities and challenges. The teacher questionnaire (eight questions) focused on instructional practices and possibilities to enhance or limit learning, teaching and assessment processes. The chair questionnaires (four questions) inquired about language policy and curriculum development. Overall, we collected 232 questionnaires. Although all stakeholders could provide insights into learning, teaching and assessment, we focused (but not limited) the questionnaires on participants’ most immediate roles. For example, students were likely to use GenAI to support their *learning* process; nevertheless, their questionnaire included open questions like: Have you used generative artificial intelligence (GenAI) technologies in your university courses? If so, in *what way*? Or how do you think artificial intelligence will transform your *educational* process?

The second data collection instrument was focus groups. We designed an interview protocol of five questions to further enquire about GenAI uses and practices, institutional policies, and future opportunities in light of language assessment processes. Then, we invited four professors from each program to participate in two separate focus

groups via Google Meet. These lasted approximately one hour and were recorded for further analysis. Finally, we highlight that all instruments were piloted with individuals who resemble the target populations to make them clear, pertinent, and concise.

Data Analysis

The qualitative data from the questionnaires and focus groups were analyzed through content analysis, a “qualitative-data reduction and sense-making effort that ... attempts to identify core consistencies and meanings” (Patton, 2002, p. 453). We used the software ATLAS.ti 24 WEB to code data using its integrated AI codifier feature. Then, we provided the research questions and objectives and the following potential themes: perspectives, attitudes, ethical considerations, opportunities, threats, practices/uses and needs. With this information, the AI-coding beta system carried out an initial analysis that ended up in over 200 codes, many of which lacked organization or relevance. Therefore, after agreeing on the definitions of the themes previously mentioned, we teamed up in pairs to explore the relevance of codes and generate new codes when needed. We synthesized the number of codes to 55.

Inter-Coder Agreement (ICA) was initially calculated and yielded a coefficient of 0.58. Upon closer examination, discrepancies were primarily observed in excerpts categorized as “opportunities” or “perspectives.” Subsequent discussions clarified that these disagreements stemmed from semantic differences in label application rather than divergent interpretations of the underlying constructs. To enhance reliability, we resumed coding and recalculated the ICA, resulting in a coefficient of 0.90. The coding framework was subsequently refined through iterative processes, with minimal reliance on AI-generated codes, to inform the ensuing analysis.

10/29

RESULTS

The revision of the overall coding led to three key thematic categories: (a) GenAI uses for English language learning, teaching, and assessment; (b) Stakeholder’s ethical concerns in language education and assessment; (c) Need for GenAI Literacy and ethical uses and practices for GenAI Integration; and (d) Challenges and opportunities to integrate GenAI as part of language education and assessment dynamics.

GenAI Uses for English Language Learning, Teaching, and Assessment: Critically-Oriented Practices of Assessment

Participants reported positive perspectives (341=51%)² and attitudes (181=27%) towards uses of GenAI for English language learning, teaching and assessment. To begin with,

² These percentages refer to the total coding occurrences per category.

GenAI was reported to be mostly used as a consultant (135 =22%) to clarify, summarize, extend, and collect information by students. For instance, student #50 mentions that: "I have used them in seminars such as sociolinguistics or language acquisition to summarize readings, ask for concept definitions that I do not have very clear and need a profound explanation."³

Additionally, GenAI is greatly used to enhance academic writing (113=15.1%) processes. In this regard, student #125 mentions that they have used Chat GPT to help them: "edit the text in such a way that it complies with grammar rules and has coherence and cohesion or ask for ideas to start writing a particular task". Similarly, professor #13 wrote that: "Some activities involve going to one of these artificial intelligences and giving it specific instructions in front of the students; also, opening to dialogue about appropriate and ethical use should be a pillar in education supported by AI." Thus, these participants describe potential uses of GenAI technologies for *language learning*, as they help participants analyze content and refine written productions in English.

It is relevant to mention that for *language teaching and assessment*, professors recognize the use of GenAI as a time saver (135=18%). For example, professor #20 argues that "they [GenAI technologies] can help professors to reduce time in material design and be of great support for brainstorming and designing better educational practices". Also, professor #11 mentions that GenAI technologies are useful to save time in assessment-related tasks: "for teachers, being able to be more productive by reducing the time spent on repetitive and time-consuming tasks, such as generating assessment rubrics." Hence, GenAI has become part of the participants' regular teaching-learning-assessing processes. Nevertheless, they have also gone beyond such frequent uses.

11/29

Participants mentioned both ordinary and extraordinary GenAI-based assessment tasks. To illustrate, professor #18 has used Gen AI "as a helper in providing feedback on student work with students so they can receive feedback from AIs themselves, and to create assessment materials." In regards to material design, professor #20 uses GenAI "for the generation of rubrics, and the result for me has been very satisfactory. I have had to make some adjustments with respect to the level of the task that is required." Additionally, the same professor added: "I include a text in my tests and instruct it [GenAI], generate a text with these characteristics. In my case, where my subject is advanced grammar, it has given me about 50% effectiveness." Notably, these professors are cautious when using GenAI, so they make adaptations when implementing it for average assessment tasks.

³ Data extracts quoted directly from the data collection instruments during 2024. Participants are identified by numbers, not by name, to protect their anonymity. All excerpts were translated from Spanish by the Authors.

By the same token, professors are aware that GenAI tools make mistakes, so they approach them critically. For instance, professor #23 argues that “ChatPDF [sic]⁴ sometimes invents the answers, or its explanations are less clear than those that I could offer as a teacher”. In the same line, professor #30 considers that “we have to understand and define what specific activities we can entrust to AI, mainly tasks that do not require depth of thought”. Hence, professors value their human cognition, which encourages them to double-check GenAI’s work; in other words, they use their knowledge to test and assess GenAI.

In regard to extraordinary tasks, GenAI technologies are used as inspiration sources or self-reflection triggers. That is why student #102 mentioned that:

The transformation will depend solely on how these tools are used because some use them as tools that enable development, while others rely on them entirely. In the latter case, the transformation would be entirely negative, it limits the process of creating and learning. In my case, it transforms my process in terms of making me more aware of what I am doing incorrectly and giving me tools to extend my knowledge and explore other areas. So, ChatGPT in this case acts as a stimulus that alters the results of my educational process.

12/29

Hence, we noticed three critical oriented assessment incidents here. First, the participant resorts to ChatGPT to assess its impact on her learning process. Second, she foresees negative outcomes if GenAI is used uncritically, as creative and critical thinking would fail to develop. Third, she urges others to approach GenAI tools consciously. Therefore, these are extraordinary examples in that GenAI might contribute to complex academic and personal assessment processes.

Stakeholders’ Ethical Concerns in English Language Education and Assessment

The second finding describes underlying threats that have raised ethical concerns and considerations among participants. The most frequent concerns were: AI dependence (116 =32.9%), lack of critical and collaborative use of GenAI (91=25.8%), GenAI trustworthiness (47=13.3%) and inappropriate use (33=9.3%) (see Table 2). Hence, we synthesized these codes into two subfindings: the originality dilemma in assessment and uncritical and dependent use of AI for assessment demands. Before moving on, we should briefly address other GenAI threats (see Table 1) identified by stakeholders.

⁴ ChatPDF was mistakenly used as a reference to ChatGPT as confirmed by the participant.

Table 1. Tabulation of Codes Related to GenAI Threats

GenAI Threats		
Codes	# of quotes per category	% out of the total coding occurrences per category
AI Dependence	116	32.9%
Lack of critical/collaborative use of GenAI	91	25.8%
GenAI Trustworthiness	47	13.3 %
Inappropriate use	33	9.3%
Plagiarism	31	8.8%
Ability loss	25	7.1%
Authenticity loss	9	2.6%

For instance, student #8 reported that “it could be a limiting or hindering factor to use information provided by AIs, such as better ways to say things, since it could create a lack of confidence”. This insight suggests that easy access to AI-generated writing suggestions could erode confidence in students’ abilities, particularly when they perform without the support of an AI-assisted writing tool. Also, regarding learning, stakeholders consider that AI misuse could lead to the loss of productive skills in the first and target languages. For example, director #2 pointed out that “it [GenAI] negatively affects the program in terms of the scope of our training purpose. The first is the extent to which the use of GenAI is affecting the development of students' C1 competence in English and Spanish.” This idea resembles the need for responsible GenAI use regarding the composition of texts in the target or first language. Accordingly, professor #19 expressed the need of “reinforcing political and ethical judgment in the [GenAI] possible uses, first in teachers and then in students.” Thus, the professor reflects on the importance of AI literacy. Now, we shall elaborate on the most common concerns mentioned by participants.

The Originality Dilemma in Assessment

Some concerns have to do with authenticity, academic honesty, trustworthiness and difficulty in telling human-made tasks apart from AI-assisted or AI-generated ones. To illustrate, student #92 said: “I want to be a teacher, but I feel that it is going to be complicated to recognize my students’ own work from one done by artificial intelligence.” Similarly, professor #20 stated that if learners “do not use them as support but rather as a replacement for creating texts and answering quizzes, teachers will enter into dilemmas of trust towards the ethics and honesty of students.” Thus, participants describe how the misuse of these technologies and the difficulty in telling them apart may erode the trust between academic members and bring into question the authenticity of every work.

The authenticity challenges professors face in assessment processes have prompted various responses to adapt to or overcome the identified threats. On the one hand, professor #32 argues that: “for summative written production activities, students carry out the activity in the classroom, and I request it by hand and restrict the use of the cell phone. However, I have implemented the use of ChatGPT for training activities.” Hence, the professor has fostered GenAI use for some formative assessment writing activities but opted for no technology use when it comes to summative assessment procedures.

14/29

On the other hand, professor #25 considers that: “they [GenAI tools] could be used to assess class assignments. I think that the evaluation must change in what is going to be evaluated: not the content, but the structure and display of the test.” In other words, the professor urges a reimagination of assessment strategies in terms of structure and delivery of the test.

Another response to ensure authenticity involves AI detectors, even though some stakeholders consider them unreliable. To exemplify, student #126 said “I am concerned about professors who trust AI detectors that may not be so reliable. In the best cases, they make you repeat the writing. I don't think it's so fair when there are people who are simply unlucky.” Similarly, professor #13 commented that: “When I started reading the essay and the feedback, the things that the AI said were not true. AI does not have the capacity to identify this metaphor [or] simile. The judgment must continue to be human.” Therefore, these testimonies show that stakeholders face challenges whether they use GenAI for assessment tasks or not.

Finally, participants highlighted another issue related to the originality dilemma. That is, professor #20 found that: “They used it as a corrector. People say I want to present a well-done work, but I have grammatical and linguistic errors. The problem is that when they copy and paste the final product, it does not appear to be theirs.” Thus, the situation problematizes the extent to which an AI-enhanced written product is still original or authentic.

Uncritical and Dependent Use of AI for Assessment Demands

Stakeholders consider that GenAI dependence yields an uncritical use, which obscures students' actual competences. For instance, student #188 expressed that "excessive use can generate laziness in its users, leaving every task or job they have to artificial intelligence, without being able to internalize something or without achieving anything on their own." This was further problematized by director #1, who noted that: "if there is no critical stance on the part of the students, they become simply consumers of the information without further analysis or further interpretation." Hence, these participants consider that consumerist behaviors in education are exacerbated by the availability of GenAI tools when students use them uncritically.

In this line, Table 1 shows that GenAI dependence is the most pressing risk stakeholders have identified regarding the use of GenAI technologies when facing assessment requirements. Professor #35, when referring to the final proficiency test of the program, stated that: "when [students] are faced with a real scenario where they must show their abilities, there are students who have not been able to demonstrate that [C1] competence. Similarly, student #14 responded that GenAI dependence "completely suppresses or cancels my imagination for certain activities, which is why I think that they should not be used very often." Therefore, stakeholders show concern about GenAI overreliance and language- or creative-skills loss.

15/29

Need for GenAI Literacy and Ethical Uses and Practices for GenAI Integration

Living in the so-called globalized and digitized era reinforces the idea that technology integration occurs naturally, yet this subfinding reveals the need to reassess such a narrative. That is, the majority of professors (19=54%) did not use GenAI technologies in their teaching and assessment practices because of the following two reasons. First, professors ignored GenAI's functionality (4=11.4%), and expressed the need "to manage them and understand their specific characteristics to use them according to their purpose (professor #23)." Second, they (6=17%) opted not to integrate GenAIs despite having some acquaintance with them. For instance, professor #16 comments that "I have heard about their existence. However, I've never been interested in using them." Therefore, professors reveal literacy and reluctance issues that challenge the naturalistic view of GenAI integration.

Additionally, students, whose majority are acquainted with GenAIs (172=88%), address the need for ethical practices. Student #130 considers that "people don't understand that they harm themselves when they ask an AI to do work for them. They should be used in a way that supports education, not as the ultimate solution to mental laziness." In the same line, student #29 comments that challenges include "the lack of proper use and knowledge about the various AIs, as most people only know about

ChatGPT and use it to generate something they are too lazy to do." Hence, both students converge with the idea that AIs should be used responsibly to *support* their learning process. What is more, they argue that laziness is a potential underlying reason for irresponsible use that suggests validity and authenticity issues, given that assessment would revolve around IA rather than human work.

In closing, professors and directors also addressed the need for ethical GenAI uses when doing assessment tasks. Professor #46 considers that "teachers and students need to know when and how to use them. Just as it is unfair for students to provide AI-generated answers, teachers should be concerned about not always creating AI-generated questions." In this regard, director #2 adds that in relation to assessment design with AI assistance, "we do not have a professional development plan for teachers in this regard." Therefore, it is paramount to enhance GenAI literacy and ethical practices among stakeholders, especially for the design of their assessment processes and how these influence their teaching routines.

Challenges and Opportunities to Integrate GenAI: Professional Development and Autonomy as Beacons in the GenAI Future

16/29

Participants reflected on challenges hindering GenAI integration and potential opportunities arising from these disruptive technologies. The most common challenges were related to AI illiteracy (72=63.1%) and unethical uses and practices (21=18.4%), and the most prevailing opportunities included professional development (34=35.7%) and autonomy (29=30.5%). Table 2 shows other challenges and opportunities related to teaching, learning and assessment.

Table 2: Tabulation of Codes Related to Gen AI Challenges and Opportunities

<i>Challenges</i>			<i>Opportunities</i>		
Codes	# of quotes per category	% out of the total coding occurrences per category	Codes: Opportunities for	# of quotes per category	% out of the total coding occurrences per category
Lack of...					
AI literacy	72	63.1%	Professional development	34	38.6%
ethical AI uses and practices.	21	18.4%	Autonomy	34	38.6%
AI policy	13	11.4%	Personalized Assessment	20	22.8%
investment in tech-infrastructure	5	4.3%			
Time	3	2.6%			

Regarding the latter, participants anticipate opportunities for personalized assessment. Director #1 commented that “students have different styles and respond to assessments differently. These technologies could contribute to making those assessment materials or tests better meet the learning style.” Thus, the participant comments about the potential role of GenAI to build on the student's persona.

Also, stakeholders foreshadow a potential revolution in the assessment field. Professor #6 argues that “it [GenAI] challenges many of our conceptions of teaching and learning, questioning our role as educators, and the ways we plan, execute, and evaluate in education.” Likewise, student #193 considers that “it's necessary to innovate the way we evaluate and assess concepts because monotony or lack of creativity leads to the ease with which these intelligences solve tasks.” Therefore, stakeholders envision the necessity of rethinking their roles to thrive in the GenAI conjuncture.

Just as GenAI has brought significant challenges, it has also come with professional development and autonomy opportunities for education and assessment. To begin with, stakeholders were prone to update themselves to better respond and make use of GenAI technologies. For instance, professor #6 reflected that “for teachers, [GenAI] disrupted many of the teaching and learning conceptions, questioning our role as teachers, and above all, the ways of planning, executing and assessing.” Hence, the

teacher has analyzed the intersections between GenAI technologies and education opportunities, including assessment.

Also, stakeholders identified GenAI as an opportunity to lead global discussions, drawing from local efforts. That is why Director #2 mentioned the “possibility for students to be at the vanguard of global dynamics by developing their digital literacy.” Yet, they recognize the need and opportunity for varied and inclusive professional development processes. To illustrate, director #1 said that: “There are possibilities through training mechanisms that can take the form of workshops, conferences, seminars, courses. And it seems that collaborative work at the community level can be more effective, where teachers [and] students are involved.” Thus, while stakeholders acknowledge the need to grow professionally in response to GenAI needs, they highlight the value of bottom-up processes of collaborative exploration and reflection that involve all stakeholders.

To wrap up, participants highlighted GenAI as an opportunity to foster self-guided study and revisions for formal assessment procedures. As professor #26 commented: “These tools can generate new learning, consolidate learnings that are initially addressed in classrooms. It is much simpler, and it is no longer a necessity to have a human being on the other side of the screen.” Despite the risk of unethical uses presented previously, students commented on purposeful and autonomous uses of GenAI for assessment preparation. For instance, student #47 sees GenAI as “a teacher at hand, whom I frequently ask for exercises and to ask me questions when an exam is coming, so I can solve and study them”. Therefore, GenAI tools offer opportunities for students to autonomously prepare for assessment tasks by seeking GenAI-generated mediation and feedback.

18/29

DISCUSSION

The results of this study shed light on a debate about integrating GenAI into education. To begin with, the results show how GenAI has been identified as a disruptive technology (Bower & Christensen, 1995) that not only offers several possibilities into existing practices but also requires brand-new training, ongoing discussion and educational research. We could identify a growing openness (Jamshed et al., 2024) towards GenAI in education, indicating a possible normalization process (Bax, 2003, 2018) of these technologies among stakeholders. Thus, there is already a difference with Huang et al. 's (2023) extensive review of AI trends between 2000 and 2019, which showed that teachers' and students' reluctance was one of the biggest challenges for AI integration. In this study, the majority of professors and students showed willingness to ride the wave of AI in education, showing these technologies' disruption on stakeholders' worldviews and practices, and the ongoing normalization process.

As an illustration, participants reported GenAI's practicality to support multiple education processes. They mentioned autonomous learning (Bekou et al., 2024; Kurni et al., 2023), personalized feedback (Huang et al., 2023; Owan et al., 2023), and assessment instruments and material generation (Claro et al., 2024; Hidayah et al., 2023; Ricart, 2024). In regard to the latter, we confirm that the creation of language-teaching materials powered by GenAI fosters collaborative and interactive environments (Serafimovska, 2023; Singha et al., 2024), like the collaboration between GenAI and our participants. Additionally, these GenAI-supported processes suggest that the nature of assessment has been profoundly transformed, as students can receive feedback anytime and professors can generate countless assessment tools, to mention some examples. In general, stakeholders' perspectives align with studies that argue for the positive impact of technologies in ELT (Cinganotto, 2023; Juarez & Perales, 2019; Kusuma, 2022; Serafimovska, 2023).

By the same token, participants identified both opportunities and threats (Bewersdorff et al., 2023; Owan et al., 2023) from GenAI. On the one hand, they recognized opportunities like professional development (Bekou et al., 2024; Kurni et al., 2023), automated grading (Edmett et al., 2023; Huang et al., 2023; Owan et al., 2023), and autonomous learning and assessment preparation (Claro et al., 2024). Hence, this technology has positively transformed assessment processes, and inherently, English language learning (Juarez & Perales, 2019).

19/29

On the other hand, stakeholders indicated that GenAI might hinder their confidence as they will not trust their writing skills unless double checked by AIs. This idea contrasts with Claro et al.'s (2024) findings regarding students' confidence. The main difference is that students in this study elaborated mainly on GenAI autonomous use experiences while in the latter study they were purposefully guided by professors. Hence, this divergence reveals the importance of consciously bringing GenAI tools to ELT, as in the cases reported by Cinganotto (2023) or Kusuma (2022).

The increasing consumption of GenAI products as trendy systems that build students' deceiving confidence is considered by stakeholders as a consumerist behavior. They highlight how some students may resort to GenAI as a way to bypass reading and writing assessments that may require an actual cognitive effort. This uncritical attitude may be caused by an excessive focus on grades in education and the need to appear productive and effective, traits that are fostered in a system that values consumerist ethics (McLaren, 2014). Then, Bozkurt et al. (2024) insist that "there is a need to ensure that reliance on GenAI does not lead to a reduction in meaningful learning, as students may become passive consumers of GenAI-generated content" (p. 496). Additionally, there is a necessity not only to come up with more alternative and formative assessment practices, but also "rethink the very nature of [language] education," (Bozkurt et al., 2024, p. 507), particularly language assessment in light of GenAI. This, in turn, may reduce the student

authenticity and trustworthiness issues during assessment procedures that were identified in the findings, since a deep restructuring of the way they are assessed and the purposes of such assessment may incentivize more authentic, yet AI-supported if necessary, language productions (Corbin et al., 2025).

Also, stakeholders mentioned challenges such as dishonesty and laziness (Makeleni et al., 2023). In this line, the significant number of stakeholders, especially students, who reported the use of GenAI for academic writing processes or rubric generations hints that many of the texts we read and generate today in the classroom are AI generated or at least AI-Augmented (Castañeda-Trujillo, 2022; España, 2022). Because it is rarely admitted whether GenAI was used to generate or augment a text, loss of authenticity, academic honesty (Makeleni et al., 2023), and trust among stakeholders become significant challenges. Hence, the potential outcomes when integrating GenAI tools require a thoughtful review of conditions, informed by research and stakeholders voices, for successful implementation.

In light of the aforementioned challenges, stakeholders suggested some ideas to achieve a smoother transition. Similar to Bewersdorff et al. (2023), participants acknowledge the necessity of specialized training to accompany their willingness to integrate GenAI tools. Likewise, they highlight the importance of the human factor since AIs are prone to making mistakes, as mentioned in the latter study and Huang et al. (2023). Furthermore, stakeholders, especially professors and directors call for the definition of robust principles that may guide them and their students towards best practices (Galaczi, 2023). The author warns that the integration of GenAI to teaching and assessment should be responsible and ethical; that is, through “collaboration, research and informed insights ... to benefit teachers, learners and education institutions” (p 3.) Hence, there is a growing need for fostering learning opportunities towards the development of AI literacy.

To bridge the AI literacy gap, educational authorities worldwide must make judicious efforts. Romero et al. (2024) suggest the generation of “comprehensive training programmes and ongoing support structures that enhance AI literacy among faculty, staff, and students, while addressing the challenges and benefits of proficient AI use across the university” (pp. 133-134). These programs and structures should go beyond the mere knowledge of the technicalities, towards the development of critical skills and habits for its use. That way, stakeholders could better meet the new expectations of their given role and functions in academia and collaborate towards the overall updating of the learning, teaching, and assessment practices.

CONCLUSIONS AND IMPLICATIONS

In this study, we set out to explore how stakeholders from two Colombian undergraduate language programs approach GenAI technologies in the context of English language

learning, teaching and assessment. In this sense, the question of this research was answered since several insights were gained about the perspectives and practices of stakeholders in GenAI in the context of language learning, teaching and assessment. Regarding the first objective, it could be evidenced that there is an overwhelming positive perspective and attitude towards GenAI and a growing use of these technologies, mainly for consulting purposes, time management, academic writing, and text simplification. Additionally, stakeholders recognize the importance of critically using these technologies in learning, teaching, and assessment processes, that is, not blindly normalizing them.

As the second objective intended, several threats and ethical considerations were identified. These threats include a likely dependency and loss of fundamental abilities such as critical thinking, creativity or language skills if these technologies are not properly used or are overused. Additionally, stakeholders identify a lack of trustworthiness and loss of authenticity, which results in a difficult dilemma of originality, posing significant challenges for language assessment. Thus, the study identified a number of challenges that at the same time drive stakeholders towards professional development and training opportunities that may help them better respond to the new contextual requirements.

Regardless of the aspect of the educational process (teaching, learning or assessment) stakeholders agree on the fact that the use of GenAI can lead to both positive and negative outcomes. That is why their call for critical reflection of GenAI is pertinent nowadays and invites us to “answer questions such as ‘What are the positions of the programs regarding GenAI use?’ ‘What is the incidence of GenAI in students’ thought, learning and production?’ ‘How can we integrate a critical use of GenAI to our lessons?’” (Directive #1). These questions should take us to a common speaking place that makes us reflect about the positive and negative perspectives, the ethical considerations, and challenges and opportunities regarding GenAI use.

These findings suggest that the way stakeholders address texts and produce them is changing, which calls for new alternatives in the classroom that require students not only to approach or produce text on their own, but also to collaborate with GenAI tools in doing so, or evaluate completely AI-generated products. This in turn implies rethinking the way educators and institutions assess, and the type of tasks they include in their assessments, so as to promote a healthy balance between AI-supported students' work and AI-restricted ones, so that student linguistic independence and language development be preserved.

Building on the insights gained regarding GenAI use, this study explored the critical role of structural changes in shaping institutional curricular structures (Corbin et al., 2025). Superficial or discursive policy shifts alone cannot adequately address the complex challenges posed by emerging AI technologies to language assessment. Instead, curriculum redesign that fundamentally reconsiders assessment formats, learning

sequences, and pedagogical approaches is essential. Such structural changes not only reinforce assessment validity but also foster authentic student engagement and equitable learning opportunities. Institutions must therefore prioritize the development of flexible, process-oriented language assessment in the curricula that integrate AI literacy and encourage critical, creative interactions with GenAI tools. After listening to the voices of stakeholders, this approach aligns with Corbin et al.'s (2025) call for embedding validity into the architecture of educational programs rather than relying on unenforceable rules or guidelines. Finally, advancing structural curricular change is imperative to prepare learners for a landscape where AI is an intrinsic part of knowledge production and communication.

Overall, these findings support the need for responsive education policies that address the professional development and AI-literacy needs and mitigate the threats that were identified, namely, loss of abilities, authenticity loss, critical and creative use, and academic honesty. It is important to notice that GenAI applications and, generally, technologies are rapidly growing, so governments and institutions are required to answer promptly to these needs with policies that are flexible and valid over time. To do so, countries should agree on new policies that, far from opening or deepening social gaps, strengthen human equity and guarantee quality education for all. Just as with assessment practices, language policies should go beyond rhetoric maneuvers and drive structural changes in language assessment practices, moving passed accountability-driven practices to the possibilities of process-oriented, formative approaches to language assessment.

22/29

Financing Data

The authors do not report having received institutional, governmental or private funding for this study.

Authors' contribution

Daniel Murcia: conceptualization, project administration, writing (original draft), software, formal analysis, data curation.

María Fernanda Jaime-Osorio: writing (review & editing), supervision, resources, methodology, formal analysis, data curation.

Luis Felipe Jaramillo-Calderón: writing (review & editing), software, supervision, resources, methodology, formal analysis, data curation.

Yimmy Alexander Hoyos-Pipicano: writing (review & editing), resources, methodology, formal analysis, data curation.

Conflicts of interest

The authors state that they do not have any financial, professional, or personal conflicts of interest that would interfere with the submission or publication of the manuscript.

Ethical implications

The author declares that the writing, research or publication of this article does not involve any ethical implications.

Statement using artificial intelligence

During the preparation of this work the authors used ATLAS.ti 24 -AI Beta Coding in order to establish the main codes to analyze their frequency. As for language use, GenAI was used to review sections of the abstract and the results. The editions made were fully conscious and human based.

Acknowledgments

Special acknowledgements to the participants and the institutions of the study who openly and generously shared their time and insights based on their academic experience in the current educational landscape.

23/29

REFERENCES

- Bax, S. (2003). CALL-past, present and future. *System*, 31(1), 13-28.
[https://doi.org/10.1016/S0346-251X\(02\)00071-4](https://doi.org/10.1016/S0346-251X(02)00071-4)
- Bax, S. (2018). MOOCs as a new technology: Approaches to normalising the MOOC experience for our learners; paper posthumously transcribed by Marina Orsini-Jones. In M. Orsini-Jones & S. Smith (Eds.), *Flipping the blend through MOOCs, MALL and OIL- new direction in CALL* (pp. 9-16). Research-publishing net.
<https://doi.org/10.14705/rpnet.2018.23.785>
- Bekou, A., Ben Mhamed, M., & Assissou, K. (2024). Exploring opportunities and challenges of using ChatGPT in English language teaching (ELT) in Morocco. *Focus on ELT Journal*, 6(1), 87–106. <https://doi.org/10.14744/felt.6.1.7>
- Bewersdorff, A., Seßler, K., Baur, A., Kasneci, E., & Nerdel, C. (2023). Assessing student errors in experimentation using artificial intelligence and large language models: A comparative study with human raters. *Computers and Education: Artificial Intelligence*, 5, e100177. <https://doi.org/10.1016/j.caeai.2023.100177>.
- Bhutoria, A. (2022). Personalized education and Artificial Intelligence in the United States, China, and India: A systematic review using a Human-In-The-Loop *Lenguaje* 53(1), e20314387
<https://doi.org/10.25100/lenguaje.v53i1S.14387>

model. *Computers and Education: Artificial Intelligence*, 3, e100068.

<https://doi.org/10.1016/j.caeai.2022.100068>

- Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, 21(4), 1-41. <https://doi.org/10.1186/s41239-023-00436-z>
- Bower, J., & Christensen, C. (1995). Disruptive technologies: catching the wave. *Harvard Business Review*, 73(1), 43–53. <https://www.infona.pl/resource/bwmeta1.element.elsevier-315e6fb9-6d1b-39e4-9036-cb79369f2c2c>
- Bozkurt, A., Xiao, J., Farrow, R., Bai, J. Y. H., Nerantzi, C., Moore, S., Dron J., Stracke, C. M., Singh, L., Crompton, H., Koutropoulos, A., Terentev, E., Pazurek, A., Nichols, M., Sidorkin, A. M., Costello, E., Watson, S., Mulligan, D., Honeychurch, S., ..., Asino, T. I. (2024). The Manifesto for Teaching and Learning in a Time of Generative AI: A Critical Collective Stance to Better Navigate the Future. *Open Praxis*, 16(4), 487–513. <https://doi.org/10.55982/openpraxis.16.4.777>
- Carless, D. (2023, june 8). *Assessment and feedback redesign for the generative AI era* [Conference presentation]. CUHK Generative AI conference, Pok Fu Lam, Hong Kong.
- Castañeda-Trujillo, J. E., (2021). Pedagogical strategies used by English teacher educators to overcome the challenges posed by emergency remote teaching during the COVID-19 pandemic. *Íkala, Revista De Lenguaje Y Cultura*, 26(3), 697–713. <https://doi.org/10.17533/udea.ikala/v26n3a12>
- Chiu, T., Xia, Q., Zhou, X., Chai, S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, e100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Cinganotto, L. (2023). Learning technologies for ELT during the pandemic in Italy: Teachers' attitudes. In S. Kourieos, & D. Evripidou (Eds.), *Language teaching and learning during the COVID-19 pandemic: A Shift to a new era* (pp. 37-57). Cambridge Scholars Publishing.
- Claro, M. F., Díaz, C. D., & Román, K. J. (2024). *Incorporating artificial intelligence and gamification in EFL education: Strategies for enhancing teaching and learning outcomes of EFL teachers* [Master's thesis, Universidad Santo Tomás]. USTA Repository. <http://hdl.handle.net/11634/55866>
- Corbin, T., Dawson, P., & Liu, D. (2025). Talk is cheap: why structural assessment changes are needed for a time of GenAI. *Assessment & Evaluation in Higher Education*, 1–11. <https://doi.org/10.1080/02602938.2025.2503964>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design* (5th ed.). SAGE Publications.

- Edmett, A., Ichaporia, N., Crompton, H., & Crichton, R. (2023). *Artificial intelligence and English language teaching: Preparing for the future*. British Council.
<https://doi.org/10.57884/78ea-3c69>
- España, J. A. (2023). Kahoot, Quizizz, and Quizalize in the English class and their impact on motivation. *HOW*, 30(1), 65–84. <https://doi.org/10.19183/how.30.1.641>
- Galaczi, E. (2023). *English language education in the era of generative AI: our perspective*. Cambridge Press.
- Gómez-Orjuela, Y. R. (2021). Digital skills for communication and content creation: Can B-learning greatly influence them? *HOW*, 28(1), 45–68.
<https://doi.org/10.19183/how.28.1.568>
- Guio, A., Muñoz, V., & Tamayo, E. (2021). *The Colombian case: adopting collaborative governance as a path for implementing ethical artificial intelligence*. Universidad de San Andrés. <http://hdl.handle.net/10908/18743>
- Hidayah, N. S., Hasyim, F. Z., & Yunita, V. (2023). Empowering language assessment: The pioneering role of artificial intelligence. *Dewantara: Jurnal Pendidikan Sosial Humaniora*, 2(3), 262-269. <https://doi.org/10.30640/dewantara.v2i3.1578>
- Hoyos-Pipicano, Y. A., & Jaime-Osorio, M. F. (in press). Technology normalisation in rural education: Public school teachers' narratives from Colombia. *Educación y Educadores*.
- Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, 26(1), 112-131.
[https://doi.org/10.30191/ETS.202301_26\(1\).0009](https://doi.org/10.30191/ETS.202301_26(1).0009)
- Inbar-Lourie, O., & Donitsa-Schmidt, S. (2009). Exploring classroom assessment practices: the case of teachers of English as a foreign language. *Assessment in Education: Principles, Policy & Practice*, 16(2), 185-204.
<https://doi.org/10.1080/09695940903075958>
- Inoue, A. B. (2023) *Labor-Based Grading Contracts. Building Equity and Inclusion in the Writing Classroom* (2nd ed.). The WAC Clearinghouse.
- Jamshed, M., Alam, I., Sultan, S. A., & Banu, S. (2024). Using artificial intelligence for English language learning: Saudi EFL learners' opinions, attitudes and challenges. *Journal of Education and E-Learning Research*, 11(1), 135–141.
<https://doi.org/10.20448/jeelr.v11i1.5397>
- Juárez-Díaz, C., & Perales, M. D. (2019). Experiencias en el aprendizaje de inglés en la educación superior. *Lenguaje*, 47(2), 358-378.
<https://doi.org/10.25100/lenguaje.v47i2.6734>
- Juárez-Díaz, C., & Perales, M. (2021). Language teachers' emergency remote teaching experiences during the COVID-19 confinement. *Profile: Issues in Teachers'*

Professional Development, 23(2), 121–135.

<https://doi.org/10.15446/profile.v23n2.90195>

Kurni, M., Mohammed, M. S., & Srinivasa, K. G. (2023). *A beginner's guide to introduce artificial intelligence in teaching and learning*. Springer eBooks.

<https://doi.org/10.1007/978-3-031-32653-0>

Kusuma, I. P. I. (2022). EFL preservice teachers' technology integration in managing and teaching speaking skills during emergency remote teaching. *Profile: Issues in Teachers' Professional Development*, 24(2), 149–165.

<https://doi.org/10.15446/profile.v24n2.97497>

Li, M., Enkhtur, A., Cheng, F., & Yamamoto, B. A. (2024). Ethical implications of ChatGPT in higher education: A scoping review. *Journal of Interdisciplinary Studies in Education*, 13(1), 55-68. <https://doi.org/10.48550/arXiv.2311.14378>

Makeleni, S., Mutongoza, B. H., & Linake, M. A. (2023). Language education and artificial intelligence: An exploration of challenges confronting academics in global South universities. *Journal of Culture and Values in Education*, 6(2), 158-171.

<https://doi.org/10.46303/jcve.2023.14>

McLaren, P. (2014). *Life in schools. An introduction to critical pedagogy in the foundations of education* (6th ed.). Paradigm Publishers.

26/29 Merriam, S., & Tisdell, E. (2016). *Qualitative research: a guide to implementation* (4th ed.). Jossey-Bass.

Ministerio de Ciencias, Tecnología e Innovación. (2024). *Hoja de ruta para el desarrollo y aplicación de la inteligencia artificial en Colombia*. Gobierno Nacional de Colombia.

Owan, V., Abang, K., Idika, D., Etta, O., & Bassey, (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *EURASIA Journal of Mathematics, Science and Technology Education*, 19(8), em2307.

<https://doi.org/10.29333/ejmste/13428>

Panizzon, D. (2019). Assessment, learning and teaching – a symbiotic relationship. In Grady, V. & Dawson, V. (Eds.), *The art of teaching science* (3rd ed., pp. 140-158). Allen & Unwin.

Patton, M.Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). SagePublications.

Pelletier, K., Robert, J., Muscanell, N., McCormack, M., Reeves, J., Arbino, N., & Grajek, S. (2023). *EDUCAUSE Horizon Report. Teaching and Learning Edition*. EDUCAUSE Publications. <https://www.educause.edu/horizon-report-teaching-and-learning-2023>

Poehner, M. E., Zhang, J., & Lu, X. (2015). Computerized Dynamic Assessment (C-DA): Diagnosing L2 development according to learner responsiveness to mediation. *Language Testing*, 32(3), 337–357

Rahman, M. M. (2023). Sample size determination for survey research and non-probability sampling techniques: A review and set of recommendations. *Journal*

- of *Entrepreneurship, Business and Economics*, 11(1), 42-62.
<http://www.scientificia.com/index.php/JEBE/article/view/201>
- Ricart, A. (2024). ChatGPT as a tool to improve written expression in English as a foreign language. *Íkala, Revista de Lenguaje y Cultura*, 29(2), 1-16.
<https://doi.org/10.17533/udea.ikala.354584>
- Rhokila, R., Akhmadaliev, D., & Matlubaxon, A. (2023). Using computers and online technologies in teaching English. *Scientific Focus*, 1(8), 996-1000.
<http://nauchniyimpuls.ru/index.php/sf/article/view/13524>
- Romero, M., Reyes, J., & Kostakos, P. (2024). Generative artificial intelligence in higher education. In A. Urmeneta, & M. Romero (Eds.), *Creative Applications of Artificial Intelligence in Education* (pp. 129-143). Palgrave Studies in Creativity and Culture.
https://doi.org/10.1007/978-3-031-55272-4_10
- Salas-Pilco, S.Z., & Yang, Y. (2022). Artificial intelligence applications in Latin American higher education: A systematic review. *International Journal of Educational Technology in Higher Education*, 19(21), 1-20. <https://doi.org/10.1186/s41239-022-00326-w>
- Serafimovska, S. (2023). *The effects of blended learning and smart technologies on ELT*. Lambert Academic Publishing. <https://eprints.ugd.edu.mk/32095/>
- Singha, S., Singha, R., & Jasmine, E. (2024). Enhancing language teaching materials through artificial intelligence. In F. Pan. (Ed), *AI in Language Teaching, Learning, and Assessment* (pp. 22–42). IGI Global. <https://doi.org/10.4018/979-8-3693-0872-1.ch002>
- Stake, R. E. (2006). *Multiple case study analysis*. The Guilford Press.
- Strasser, T. (2023). ELT in the digital age: We have come a long way. *Works in English and American Studies*, 48(1), 121-146. <https://doi.org/10.24053/AAA-2023-0006>
- Sued, G. E. (2023). A Cultural vision of algorithms: Agency, practices, and resistance in Latin America. In M. Cebral-Loureda, E. G. Rincón-Flores, & G. Sanchez-Ante, (Eds.), *What can I do: Strengths and limitations of artificial intelligence* (pp. 53-68). Chapman and Hall/CRC <https://doi.org/10.1201/b23345>
- Trace, J. (2021). Classroom assessment and assessment as learning. In H. Mohebbi, & C. Coombe (Eds.), *Research questions in language education and applied linguistics: A reference guide* (pp. 317-322). Springer. <https://doi.org/10.1007/978-3-030-79143-8>
- Villegas-Torres, P. (2022). The use of social networks in L2 instruction: Discussing its opportunities and obstacles. *HOW*, 29(1), 173–193.
<https://doi.org/10.19183/how.29.1.622>

APPENDIX

Sample question Focus Group

Romper el hielo

Por favor mencione su nombre, cargo y el tiempo de experiencia en el área de la enseñanza de idiomas.

I. Uso del GenAI en los procesos de enseñanza, aprendizaje y evaluación

1. ¿Cuáles son sus pensamientos sobre el papel de IAGen en los procesos de enseñanza, aprendizaje y evaluación de lenguas y el desarrollo de las diferentes habilidades receptivas y productivas?

II. Estrategias y Adaptación (diseño y principios de evaluación)

- 28/29
1. ¿Cómo creen que se puede aprovechar la IAGen para crear evaluaciones del lenguaje más efectivas o eficientes?
 - a. ¿Cómo creen que se puede hacer uso de la IAGen para generar procesos evaluativos más justos y que respondan a las diferentes necesidades lingüísticas de los estudiantes?
 2. ¿Qué estrategias de diseño de evaluación podrían ayudar a mitigar el potencial de que los estudiantes abusen de las herramientas GenAI?

III. Prácticas de Integración al Programa

1. ¿Cómo prepara usted a futuros docentes para los desafíos y oportunidades que presenta GenAI en la evaluación del lenguaje?

IV. Consideraciones Futuras

1. ¿Cuáles son las mayores preocupaciones o desafíos que tienen con respecto a la integración de IAGen en la evaluación del lenguaje?
 - a. ¿Qué apoyo o recursos adicionales serían útiles para que los educadores de idiomas naveguen por el panorama cambiante de IAGen y la evaluación?

Preguntas Adicionales:

- Si tiene, por favor comparta un ejemplo de una tarea de evaluación específica donde GenAI podría ser beneficioso o problemático. (opcional)