

AN APPROACH TO THE ADVANCE OF TOOLS BASED ON GENERATIVE
ARTIFICIAL INTELLIGENCE FOR LANGUAGE ASSESSMENT: A SWOT
ANALYSIS

*UMA ABORDAGEM PARA O AVANÇO DE FERRAMENTAS BASEADAS NA
INTELIGÊNCIA ARTIFICIAL GERATIVA PARA AVALIAÇÃO LINGÜÍSTICA DE
LÍNGUAS: UMA ANÁLISE SWOT*

*UNA APROXIMACIÓN AL AVANCE DE LAS HERRAMIENTAS BASADAS EN
INTELIGENCIA ARTIFICIAL GENERATIVA PARA LA EVALUACIÓN LINGÜÍSTICA
DE IDIOMAS: UN ANÁLISIS DAFO*



Jesús GARCÍA LABORDA¹
e-mail: jesus.garcialaborda@uah.es



Teresa MAGAL-ROYO²
e-mail: tmagal@degi.upv.es



Miguel FERNÁNDEZ ÁLVAREZ³
e-mail: m.fernandez@upm.es

How to reference this paper:

GARCÍA LABORDA, J.; MAGAL-ROYO, T.; FERNÁNDEZ
ÁLVAREZ, M. An approach to the advance of tools based on
generative artificial intelligence for language assessment: a SWOT
analysis. **Revista Ibero-Americana de Estudos em Educação**,
Araraquara, v. 19, n. 00, e024113, 2024. e-ISSN: 1982-5587. DOI:
<https://doi.org/10.21723/riaee.v19i00.19060>



| Submitted: 21/02/2024
| Revisions required: 30/03/2024
| Approved: 09/04/2024
| Published: 13/09/2024

Editor: Prof. Dr. José Luís Bizelli
Deputy Executive Editor: Prof. Dr. José Anderson Santos Cruz

¹ University of Alcalá (UAH), Madrid - Spain. University Professor, Franklin Institute.

² Polytechnic University of Valencia (UPV), Valencia - Spain. Associate Professor, Department of Graphic Engineering.

³ Polytechnic University of Madrid (UPM), Madrid - Spain. Contracted Doctor Professor, Department of Applied Linguistics to Science and Technology.

ABSTRACT: This article examines the integration of Generative Artificial Intelligence (AI) in language assessment. Employing a mixed approach, it reviews specialized literature and collects data through surveys and interviews with educational professionals, to understand the trends and challenges of these tools. The results indicate that while AI applications are efficient and adaptable to individual needs, they face significant challenges, including ethical issues and resistance to change by teachers and students. Finally, it concludes the necessity of strategic and careful planning to effectively incorporate AI in language education, improving teaching and assessment methodologies.

KEYWORDS: Generative artificial intelligence. Assessment. Competencies. Second language learning.

RESUMO: Este artigo analisa a integração da Inteligência Artificial (IA) Generativa na avaliação linguística de idiomas. Utilizando uma abordagem mista, revisa-se a literatura especializada e coletam-se dados por meio de pesquisas e entrevistas com profissionais da educação, para compreender as tendências e desafios dessas ferramentas. Os resultados indicam que, embora as aplicações de IA sejam eficientes e adaptáveis às necessidades individuais, enfrentam desafios significativos, incluindo questões éticas e resistência à mudança por parte de professores e estudantes. Finalmente, conclui-se a necessidade de um planejamento estratégico e cuidadoso para incorporar eficazmente a IA na educação linguística, melhorando as metodologias de ensino e avaliação.

PALAVRAS-CHAVE: Inteligência artificial generativa. Avaliação. Competências. Aprendizado de segunda língua.

RESUMEN: Este artículo analiza la integración de la Inteligencia Artificial (IA) Generativa en la evaluación lingüística de idiomas. Utilizando un enfoque mixto, se revisa literatura especializada y se recopilan datos a través de encuestas y entrevistas con profesionales educativos, para comprender las tendencias y desafíos de estas herramientas. Los resultados indican que, aunque las aplicaciones de IA son eficientes y adaptables a necesidades individuales, enfrentan retos significativos, incluyendo cuestiones éticas y resistencia al cambio por parte de docentes y estudiantes. Finalmente, se concluye la necesidad de una planificación estratégica y cuidadosa para incorporar eficazmente la IA en la educación lingüística, mejorando las metodologías de enseñanza y evaluación.

PALABRAS CLAVE: Inteligencia artificial generativa. Evaluación. Competencias. Aprendizaje de una segunda lengua.

Introduction

In recent years, language learning apps have evolved to adapt to the ever-changing needs of users, particularly in a technological environment increasingly dominated by mobile devices and tablets. These technological adaptations have progressed from their initial implementation in digital environments (Bonami; Piazzentini; Dala-Possa, 2020), through virtual classrooms and ad-hoc online assessment systems (Magal Royo; García Laborda, 2018), to the integration of Artificial Intelligence (AI) in content creation processes and computer-assisted testing. AI has demonstrated its value in language learning, primarily through applications based on Learning Analytics (Lahza; Khosravi; Demartini, 2023; Ouyang; Wu; Zheng; Zhang; Jiao, 2023), as well as in the generation of AI-generated digital content (Sonderegger, 2022).

In this context, there has been a notable increase in the development of AI-generated applications (Escobar Hernández, 2021), which are capable of learning from user experience and the available digital environment, thus adapting to individual needs and demands (Blikstein, 2013; Magal-Royo; García Laborda, 2017).

This article focuses on the importance of analyzing linguistic and communicative competencies in the language learning process, highlighting the relevance of the Common European Framework of Reference for Languages (CEFR) in this context. Various approaches used to develop these competencies in digital environments are explored, covering areas such as listening and reading comprehension, oral and written expression, comprehension and production in conversations, as well as mediation competencies. Additionally, the crucial role of AI in assessing these competencies is emphasized, while addressing the advantages and disadvantages associated with its use.

The article proceeds to a SWOT analysis aimed at delving into the strengths, weaknesses, opportunities, and threats related to AI-based assessment tools in language learning. Advantages such as rapid assessment, adaptability to individual needs, and improved feedback to students are highlighted. However, ethical concerns, resistance to change from teachers and students, the need for teacher training, and the importance of data management and privacy of provided information are also discussed. The article concludes by emphasizing the need to address these issues strategically to ensure the successful and equitable implementation of AI-based assessment in education.

Analysis of Linguistic and Communicative Competencies in Language Learning

The Common European Framework of Reference for Languages (CEFR) establishes a robust framework for teaching and assessing linguistic competencies, including comprehension, expression, and mediation (Council of Europe, 2020). In the context of language learning in digital environments, adapting these competencies becomes a fundamental aspect, and various strategies have been developed to optimize the educational process (Warschauer; Healey, 1998).

In the realm of listening and reading comprehension, the importance of multimedia content as an essential tool is particularly notable (Mayer, 2001). The integration of authentic audio and video not only enhances listening comprehension but also enriches the student's experience (Gilmore, 2007). The inclusion of interactive texts that incorporate multimedia elements contributes to the development of reading comprehension by providing a more immersive and contextualized experience (Guthrie; Wigfield, 2000).

Linking the functionalities offered by digital platforms enhances oral and written expression. The addition of recording features allows students to practice and improve their speaking skills (Kormos; Dénes, 2004), while the use of online collaborative writing tools encourages interaction and joint practice of written expression, thus promoting collaborative learning (Storch, 2005).

Regarding comprehension and production in conversations, the use of videoconferencing tools and online forums emerges as an effective strategy (Hampel; Stickler, 2005). These tools facilitate real-time conversation practice, providing students with the opportunity to interact and communicate effectively in a digital environment, simulating real-life situations.

Mediation competencies, which involve the ability to translate and negotiate meanings between two parties, find practical application in online collaborative tasks (O'Dowd, 2006). Educational platforms can design activities that require students to work together to resolve linguistic or cultural issues, thereby promoting digital mediation as an integral component of language learning.

The use of digital resources is essential in this context, where specific learning applications offer games, interactive exercises, and playful activities to maintain interest and engagement (Cornillie; Thorne; Desmet, 2012). Adaptive platforms adjust to the individual progress of the student, provide personalized content according to their needs and learning pace, and optimize the learning experience (Xu; Peng, 2017).

Automatic feedback emerges as a crucial component for continuous improvement. Automated assessment systems provide instant feedback on pronunciation, grammar, and vocabulary, enabling students to correct errors immediately and adjust their learning approach more effectively (Neri *et al.*, 2002).

Furthermore, cultural interconnectedness is strengthened through the integration of online cultural resources. Videos, articles, and multimedia experiences explore the culture of the target language, providing an enriched cultural context for language learning and thereby contributing to a deeper and more contextualized understanding (Kramsch, 1993).

Ultimately, motivation and progress tracking are optimized through gamification and progress boards on digital platforms. Gamification elements, including rewards and virtual challenges, inspire students, while progress dashboards offer a clear visualization of their development and achievements, encouraging persistence and commitment in the learning process (Lee; Martelo, 2011). These combined strategies foster an effective and engaging learning environment in the digital context of language acquisition.

Methodology

This study employs a mixed-methods approach to investigate the development of Generative Artificial Intelligence-based tools for linguistic assessment in language learning, using a SWOT analysis to understand the trends, capabilities, and challenges of these tools. A comprehensive review of academic literature and current technological developments was conducted. The SWOT analysis allowed for the evaluation of strengths, such as efficiency and personalization, and weaknesses, including ethical challenges and resistance to change.

In addition, surveys and interviews were conducted to gather perspectives from professionals in the educational sector. Surveys, designed with structured and semi-structured questions, were distributed to a diverse group of educators and educational technology developers, focusing on collecting feedback on the utility, effectiveness, and perceived challenges of AI tools. Interviews, on the other hand, were conducted in a more open format, allowing respondents to express their experiences and viewpoints in depth. This mixed-methods approach to data collection provided a comprehensive understanding of attitudes toward AI in education.

The information obtained from surveys and interviews was qualitatively analyzed to identify common trends and expert opinions. The results from the SWOT analysis, influenced

by these data, reveal promising areas for the future integration of AI in language teaching and highlight critical areas that require attention.

The findings from the SWOT analysis were used to enhance the discussion and conclusions of the study, indicating promising areas for future AI integration in language teaching and highlighting critical areas requiring special attention. This analysis serves as a foundation for a more comprehensive study on AI-based assessment tools, with a particular focus on the challenges of using AI for language assessment (Zawacki-Richter *et al.*, 2019; Helms; Nixon, 2010).

Analysis of Generative AI Tools in Language Learning

The analysis of generative AI tools designed for assessing linguistic and communicative competencies in second language learning is crucial in the current educational context. The need for effective methods to evaluate students' language and communication skills has driven the development of various generative AI applications (Chapelle; Voss, 2016). This analysis focuses on understanding and categorizing these applications to identify their advantages and challenges.

Generative AI applications for assessing second language proficiency fall into three main categories:

- General AI development applications: Designed to adapt flexibly to user needs, these applications lack a specific adaptation profile and do not have rigorous control over assessment or authorship (Griffiths; Tagg, 2018).
- Generative AI applications with adapted profiles: These tools adjust to the specific needs of the user by generating personalized content. They offer formal control over aspects such as spelling, structure, writing style, and type of dialogue, making them more suitable for accurate assessment of language and communication skills (Levy; Stockwell, 2006).
- Generative AI development applications with adapted profiles and formal control: Similar to the previous category, these applications provide adapted content with more rigorous control over formal aspects, ensuring spelling accuracy, structural coherence, appropriate writing style, and the required type of dialogue (Warschauer, 2010).

This study explores the advantages and challenges of each approach in assessing linguistic and communicative competencies. Additionally, it considers the perspectives and opinions of teachers and education professionals to understand the impact of these applications on the teaching and learning process of languages as a second language (Heift; Schulze, 2007).

Table 1 presents a detailed description of various AI-based tools used in language learning. These applications vary in tasks and functionalities, ranging from information extraction from PDF files to the generation of personalized educational content. Each tool is classified according to its general orientation, the linguistic and communicative competencies it promotes, and the type of assessment it provides, reflecting significant advancements in the integration of AI into the language learning process and its potential to enhance students' linguistic and communicative skills (Blake *et al.*, 2020).

Table 1 - AI-based tools for language learning

Application	Task or Resource Description	Application Guidance	Language and Communication Skills	Assessment Type
PDFgear Chatbot	Integrated into the new ChatGPT API is an application that extracts information from large PDF files such as books, textbooks, essays, legal contracts, research papers, etc., and generates answers and references. (PDFgear, 2023)	General AI development application that recognizes texts in PDF to generate summaries or new information.	It can contribute indirectly by improving reading comprehension and writing skills. Its main focus is extracting and processing information from PDF documents.	It allows you to assess reading comprehension and written expression skills in multiple languages efficiently.
Quizlet/ Q-Chat	Q-Chat has AI technology built on OpenAI's ChatGPT API. Quizlet is a web application that trains learners through flashcards, games, and quizzes (Quizlet, 2023).	Generative AI development application tailored to user needs in the field of communication through online chats with a machine.	Its language and communication skills include improving comprehension, written expression, and real-time responsiveness in an interactive and adaptive learning environment.	It allows you to assess language skills such as vocabulary, written comprehension and real-time responsiveness in different languages.
Duolingo	Duolingo has integrated with the GPT4 API. Duolingo is a web platform focused on free language learning. Duolingo follows the 1-to-1 computational model, that is, the machine constantly teaches the student and adapts to him/her, a method that is improved day by day with the data collected by the learning of each user who has already	Generative AI development application tailored to user needs through content generation.	It includes developing reading, writing, listening, and speaking skills in multiple languages. Users can improve their pronunciation, grammar, and vocabulary through interactive and practical lessons, enabling them to communicate effectively in real-life situations in the language they are learning.	It is used for language assessment, offering interactive lessons and exercises that assess language skills, including reading, writing, listening, and speaking in multiple languages. Users can measure their progress through tests and practice exercises, allowing their level of proficiency in the

	taken the course (Duolingo, 2023)			language they are learning to be assessed.
JasperAI /Jasper Chat	It is an AI-based natural language processing (NLP) tool that focuses on improving the quality of writing and written communication. It offers grammar corrections, style suggestions, and improvements to written content to help users create clearer, more consistent, and effective texts in various contexts. This application is useful for improving skills of writing and written communication in several languages.	Generative AI development application tailored to user needs in reference to text correction based on style patterns.	It focuses on improving writing and writing skills. The app helps users create high-quality written content by correcting grammatical errors and offering suggestions to improve clarity and coherence in written communication, which contributes to developing effective writing and written communication skills in various contexts.	The tool provides grammar corrections and suggestions for written expression, contributing to the development of effective written communication skills in multiple languages.
WriteMage	WriteMage has integrated the GPT4 API. WriteMage is an AI application designed specifically for macOS and iOS devices.	Generative AI development application tailored to user needs in reference to text correction based on style patterns.	It focuses on improving the quality of writing. The app offers grammatical suggestions and corrections, making it easier to create clear and effective written content. Users can improve their writing and written communication skills in various contexts, contributing to more accurate and consistent communication.	It contributes to the development of language skills by providing accurate and personalized feedback to improve the quality of written communication in different languages.
Ivy.ai	It is an AI platform designed for communication in the higher education field. It offers chatbot and virtual assistant solutions that help students and academic staff get answers and access relevant information efficiently. Ivy.ai facilitates written communication in different languages and contributes to quick access to academic resources, improving the user experience in educational environments.	Generative AI development application tailored to user needs in the field of oral communication through online lectures and multilingual written communication.	It focuses on natural interaction with students through chatbots and virtual assistants. It provides answers to frequently asked questions, guides students through academic resources, and provides relevant information, improving communication and accessibility of information in educational settings and contributing to the development of students' communication and comprehension skills.	It facilitates written communication and the understanding of academic information, contributing to the development of language and communication skills in an educational environment.
Kahoot	It is an online learning platform that offers interactive games and quizzes for learning and assessment in educational environments. Users can create and participate in playful activities that	Gamification-oriented learning application that generates playful assessment content.	It includes the promotion of reading comprehension and oral expression skills. Users can participate in educational games that require reading questions and answers, as well as oral communication in groups, which contributes	Users can participate in playful activities that assess their level of proficiency in the language they are learning in an effective and fun way.

	<p>assess knowledge and skills across a variety of topics and subjects. Kahoot is widely used in the education field to engage students in a fun way and not effective assessment and learning.</p>		<p>to the development of communication and comprehension skills in a playful and collaborative educational environment.</p>	
<i>Khan Academy</i>	<p>Khan Academy has integrated the GPT4 API. Jasper AI is an AI writing tool and has a natural language processing (NLP) system, integrates with software tools such as WordPress and HubSpot.</p>	<p>Generative AI development application tailored to user needs by generating content with an adapted profile where there is formal control (spelling, structure, writing style, type of dialogue)</p>	<p>It focuses on supporting reading comprehension and written communication. The platform offers interactive explanations and exercises that help students understand complex concepts and improve their written expression skills. In addition, Khan Academy promotes online communication between students and teachers, which contributes to the development of communication and comprehension skills in an educational environment.</p>	<p>Students can measure their progress and proficiency in the language they are learning through practical exercises and assessments.</p>
<i>Texta</i>	<p>It is an application that allows you to review, improve, and develop writing ideas and styles depending on the user's needs.</p>	<p>Generative AI development application tailored to user needs, generating content with an adapted profile where there is formal control (spelling, structure, writing style, type of dialogue)</p>	<p>It includes reviewing, improving, and suggesting written content. Users can improve their grammar, writing style, spelling, and coherence, contributing to the development of more effective and accurate written communication skills in various contexts. Texta makes it easy to create high-quality content and improve users' writing skills.</p>	<p>Texta is used for language assessment, reviewing and improving writing in different languages, providing suggestions for written expression, and correcting grammatical errors.</p>
<i>Grammarly</i>	<p>Grammarly has reviewed, improved, and suggested texts based on text orientation, lexicon, style, and/or vocabulary.</p>	<p>Generative AI development application tailored to user needs, generating content with an adapted profile where there is formal control (spelling, structure, writing style, type of dialogue)</p>	<p>Improves writing and language skills. It provides grammar corrections, style, and vocabulary suggestions. Improves accuracy and consistency in written communication, contributing to the development of more effective and professional written communication skills in various contexts.</p>	<p>The application contributes to the development of language skills by providing accurate and personalized feedback for effective and professional written communication in various contexts.</p>

Source: Authors, 2024.

Almost all generative AI tools, such as Chat GPT (in its various versions) and SlideGo, are designed to develop content focused on data collection and presentation tasks. For instance,

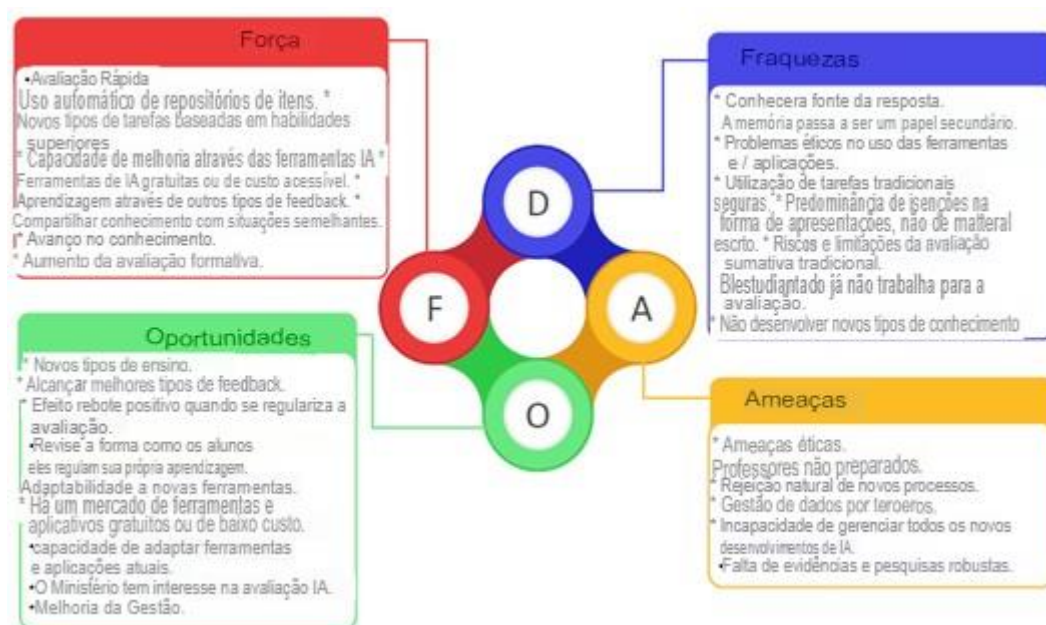
Chat GPT specializes in adapting response styles (Vaswani *et al.*, 2017), while SlideGo facilitates the creation of presentations that can be exported to PowerPoint. This advancement presents significant challenges in the educational field, particularly for language teachers, due to the difficulty in distinguishing between natural language production by students and that generated by these tools (O'Shea *et al.*, 2020).

Results of the Analysis of the Current Situation of Generative AI Applications for Assessing Language Competencies

Following the SWOT analysis, it can be noted that the most significant criticisms of generative AI come from the education sector, especially in contexts where data collection and direct tasks predominate (Johnson *et al.*, 2020). In contrast, competency-based assessment, case studies, and formative assessment are less common in the Spanish university system.

Preliminary results, shown in Figure 1, indicate that the point of **Debilidades** are relevant and stronger compared to the challenges. The latter are primarily associated with the novelty of the technology and the need for further research (Bryant *et al.*, 2014).

Figure 1 - SWOT Analysis of Language Assessment through Generative AI



Source: Authors' elaboration, 2023.

It is essential to understand the strengths and weaknesses presented by this innovative technology. The strengths of AI-based assessment include speed and efficiency in generating results, the ability to design tasks focused on higher-order skills, and the capacity to provide

personalized and real-time feedback. Additionally, these tools are becoming increasingly accessible, promoting formative assessment and contributing to the advancement of knowledge.

However, it is also important to address the weaknesses, such as the opacity in recognizing the sources of AI-generated responses, ethical concerns related to privacy and the use of original information, and algorithmic biases that could improperly influence the final content. All of these issues carry the risk of diminishing cognitive learning processes, such as memory, assimilation, attention, and critical thinking, thereby potentially limiting the progress of pedagogical innovation in human intelligence.

AI-based assessment represents a significant advancement but requires a thoughtful and measured approach to address these challenges and maximize its benefits in the educational environment, as shown below:

Strengths:

- **Automated and Rapid Assessment:** One of the main strengths of AI-based assessment is its ability to perform evaluations quickly and efficiently. The automation of the process allows for instant responses, benefiting both teachers and students by providing immediate feedback.
- **Automatic Use of Item Repositories:** AI facilitates the management and automatic selection of assessment items from repositories, streamlining the creation of tests and ensuring a greater diversity of questions.
- **New Types of Tasks and Processes Based on Higher-Order Skills:** AI enables the design of assessment tasks that go beyond rote memorization, focusing on the development of higher-order skills such as critical thinking and problem-solving.
- **Ability to Enhance Continuous and Adaptive Learning through AI Tools:** The capacity to continuously adapt and improve is an exceptional strength. AI tools can learn from students' response patterns and adjust assessments to provide more accurate and personalized feedback.
- **Free or Accessible AI Tools:** The availability of free or accessible AI tools facilitates their integration into educational environments, democratizing access to advanced technologies.
- **Learning through Other Types of Feedback:** AI allows feedback to be incorporated beyond correct or incorrect answers, offering detailed insights into the student's thought process, thereby facilitating deeper learning.

- **Knowledge Sharing with Similar Situations:** AI enables the comparison of results and learning patterns across different contexts, allowing effective knowledge and strategies to be shared between similar educational situations.
- **Advancement of Knowledge Focused on Educational Assessment and Metrics:** AI-based assessment contributes to the advancement of knowledge by providing valuable data on student performance, enabling educators and researchers to gain different perspectives on the same concept or process insights that facilitate the continuous improvement of existing assessment methods.
- **Increase in Processes and Applications Aimed at Formative Assessment:** AI promotes the development of formative assessments, which focus on student progress and development over time, rather than solely on summative assessment at the end of the process.

Weaknesses:

- **Understanding the Digital Source of Responses:** The opacity in the source of AI-generated responses can present challenges, as understanding the decision-making process is crucial for fair and transparent assessment.
- **Memory as a Cognitive Process in Learning Becomes Secondary:** AI-based assessment may diminish the importance of memory, as it focuses on the ability to apply and comprehend information in real-time, which might undervalue certain types of knowledge.
- **Properly Understanding and Interpreting Ethical Issues in the Use of Tools and Applications:** The use of AI in assessment raises ethical concerns, such as student data privacy and the potential for algorithmic biases, which require careful consideration.
- **Seeking Synergies with the Use of Traditional Safe Tasks:** A preference for traditional safe tasks may limit pedagogical innovation, as educators might opt for conventional approaches due to a lack of trust in AI tools.
- **Dominance of Exams in the Form of Presentations Rather Than Written Material:** AI-based assessment might favor presentation-based exams over written tests, which could negatively impact certain learning styles.
- **Risks and Limitations of Traditional Summative Assessment:** Although AI-based assessment offers significant advancements, the inherent limitations of traditional summative assessment still persist, especially in measuring complex and contextual skills.

- **Students No Longer Work for Assessment:** Automating assessment could reduce students' intrinsic motivation, as they may feel they are working less for evaluations by relying on automatically generated responses.
- **Lack of Detailing New Types of Knowledge:** AI-based assessment might focus on the application of existing knowledge but may not be capable of evaluating students' ability to generate new types of knowledge and creative thinking.

Analysis of Results

The analysis conducted in this study reveals a multifaceted picture. The identified strengths, such as efficiency and customization capabilities, indicate a notable advancement in these tools' adaptability to various learning styles and paces. This level of adaptability is crucial for meeting the diverse needs of students in an increasingly diverse educational context. Additionally, the ability of these applications to enhance complex language skills and promote autonomous and critical learning represents a significant leap in language pedagogy.

However, alongside these advancements, significant challenges arise. One of the most prominent is the opacity in the algorithmic processes of AI, which raises concerns about the reliability and fairness of assessments. This lack of transparency can lead to inadvertent algorithmic biases that negatively impact the objectivity of learning and evaluation. Furthermore, the ethical management of personal data and student privacy is an increasing concern in an environment where AI plays an increasingly prominent role.

Research and interviews with educators and other education professionals revealed cautious interest and optimism about the use of AI in language teaching. While they acknowledge its potential to transform language teaching and learning, they also expressed concerns about ensuring ethical and transparent practices in its implementation. Additionally, the importance of teacher training in the use of these technologies was emphasized, highlighting the need to prepare educators not only in technical aspects but also in innovative pedagogical methodologies that effectively integrate AI into the curriculum.

This analysis also underscores the importance of addressing disparities in access to technology. To ensure that AI is an inclusive and equitable tool in language teaching, it is essential to consider the availability of technological resources in different educational contexts. Similarly, the need for ongoing collaboration between AI developers, educators, and linguistics experts is evident to ensure that AI applications are both technically sound and pedagogically effective.

While generative AI applications in language assessment offer promises of innovation and improvement in education, their practical implementation requires careful consideration of technical, pedagogical, and ethical aspects. Interdisciplinary collaboration and ongoing training are emerging as critical elements to maximize the benefits of AI while ensuring its use in educational settings is equitable, fair, and enriching for all students.

Final considerations

The integration of AI into educational assessment presents significant opportunities to revolutionize teaching and learning. By facilitating the personalization of academic content, AI enables adaptive pedagogical approaches tailored to individual students' needs. This more student-centered approach promises substantial improvements in learning effectiveness.

One of the primary advantages of AI-based assessment is the opportunity to obtain more effective feedback in learning. AI provides immediate and detailed feedback that goes beyond merely correcting answers, enhancing understanding and skill development. This enriched feedback not only improves the quality of education but can also increase student motivation and engagement.

A positive rebound effect is anticipated as AI-based assessment becomes normalized in educational systems. The widespread acceptance of these tools may boost confidence in their effectiveness, encouraging their adoption and optimizing educational benefits.

Another significant opportunity is the review of how students regulate their learning. Personalized feedback facilitates a greater understanding of individual progress, encouraging more effective self-regulation. Adaptability to new tools is essential in an ever-changing educational environment, and AI supports the rapid adaptation to new methodologies and technologies.

However, there are threats associated with AI-based assessment. Ethical concerns, such as data privacy and algorithmic biases, require careful attention to ensure the integrity and fairness of the process. Teacher preparation is crucial, as effective implementation of AI-based assessment requires specific technical and pedagogical skills. Training programs are essential for a smooth transition into this new era of educational evaluation.

Some educators and students' natural resistance to new processes can be an obstacle. This resistance should be addressed through effective communication and demonstrations of tangible benefits. Additionally, managing third-party data and keeping pace with AI advancements are challenges that require ongoing vigilance. It is vital for educational

institutions to establish clear policies and safeguards to protect data integrity and ensure the ethical use of technology.

Limitations and Prospects

A major limitation of AI-based assessment is the lack of robust evidence and research to support its long-term effectiveness and benefits. Additionally, concerns arise regarding the reliance on technology and students' tendency to focus more on the technology than learning content. Further research is needed to fully understand the impact of AI across various educational contexts and demographic groups.

Looking ahead, it is crucial to continue exploring and addressing these challenges to ensure that AI-based assessment is both technologically advanced and pedagogically sound and equitable. Successful implementation of AI-based evaluations requires careful planning and management to overcome threats and ensure a stimulating and fair educational environment. This includes developing strategies for effectively integrating AI into curricula, ensuring ongoing teacher training, and fostering collaborations among educators, technologists, and policymakers.

REFERENCES

BLAKE, R. J.; GUILLÉN G.; THORNE S. L. **Brave New Digital Classroom: Technology and Foreign Language Learning**. Georgetown University Press, 2020. DOI: 10.2307/j.ctv1nc6rkf.

BLIKSTEIN, P. Multimodal learning analytics. *In: INTERNATIONAL CONFERENCE ON LEARNING ANALYTICS AND KNOWLEDGE*, 3., 2013. **Proceedings [...]**. [S. l.: s. n.], 2013. DOI:10.1145/2460296.2460316.

BONAMI, B.; PIAZENTINI, L.; DALA-POSSA, A. Education, big data and artificial intelligence: Mixed methods in digital platforms. **COMUNICAR: Media Education Research Journal**, [S.I], v. 28, n. 65, p. 43-51, 2020. Available at: <https://www.proquest.com/scholarly-journals/education-big-data-artificial-intelligence-mixed/docview/2488221174/se-2>. Accessed in: 14 Nov. 2023.

BRYANT, P.; COOMBS, A.; PAZIO, M.; WALKER, S. **Disruption, destruction, construction or transformation?** The challenges of implementing a university wide strategic approach to connecting in an open world. 2014. Available at: <http://eprints.lse.ac.uk/56577> Accessed in: 14 Nov. 2023.

CHAPELLE, C. A.; VOSS, E. 20 years of technology and language assessment in Language Learning & Technology. **Language Learning & Technology**, [S.I.], v. 20, n. 2, p. 116-128, 2016. Available at: <http://llt.msu.edu/issues/june2016/chapellevoss.pdf>. Accessed in: 14 Nov. 2023.

CONSEJO DE EUROPA. **Marco Común Europeo de Referencia para las Lenguas: aprendizaje, enseñanza, evaluación**. 2020. Available at: <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>. Accessed in: 14 Nov. 2023.

CORNILLIE, F.; THORNE, S. L.; DESMET, P. ReCALL special issue: Digital games for language learning: challenges and opportunities: Editorial Digital games for language learning: from hype to insight? **ReCALL**, [S.I.], v. 24, n. 3 p.243-256, 2012 DOI: 10.1017/S0958344012000134.

DUOLINGO. 2023. Available at: <https://es.duolingo.com>. Accessed in: 14 Nov. 2023.

ESCOBAR HERNÁNDEZ, J. C. La Inteligencia Artificial y la Enseñanza de lenguas. Una aproximación al tema. **Revista del Centro de Enseñanza para Extranjeros**, [S.I.], v. 21 n. 25, p. 29-44, 2021. DOI:10.22201/cepe.14059134e.2021.21.25.3.

GILMORE, A. Authentic materials and authenticity in foreign language learning. **Language Teaching**, [S.I.], v. 40, n. 2, p. 97-118, 2007. DOI:10.1017/S0261444807004144.

GRIFFITHS, C.; TAGG, C. The effects of a digital intervention on ESL learners' general English proficiency. **Language Learning & Technology**, [S.I.], v. 22, n. 1, p. 1-20, 2018.

GUTHRIE, J. T.; WIGFIELD, A. Engagement and motivation in reading. In: KAMIL, M. L.; MOSENTHAL, P. B.; PEARSON, P. D.; BARR, R. (ed.). **Handbook of reading research**. Mahwah: Erlbaum, 2000.

HAMPEL, R.; STICKLER, U. New skills for new classrooms: Training tutors to teach languages online. **Computer Assisted Language Learning**, [S.I.], v. 18, n. 4, p. 311-326, 2005. DOI:10.1080/09588220500335455.

HEIFT, T.; SCHULZE, M. **Errors and intelligence in computer-assisted language learning: Parsers and pedagogues**. New York: Routledge, 2007.

HELMS, M. M.; NIXON, J. Exploring SWOT analysis – where are we now? A review of academic research from the last decade. **Journal of Strategy and Management**, [S.I.], v. 3 n. 3 p. 215-251, 2010. DOI:10.1108/17554251011064837.

JOHNSON, M.; GRIFFITHS, M.; WANG, Y. Artificial intelligence in education: Challenges and opportunities for sustainable development. **UNESCO Working Papers on Education Policy**, [S.I.], n. 7, 2020. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000366994.locale=es>. Accessed in: 14 Nov. 2023.

KORMOS, J.; DÉNES, M. Exploring measures and perceptions of fluency in the speech of second language learners. **System**, [S.I.], v. 32, n. 2, p. 145-164, 2004. DOI: 10.1016/j.system.2004.01.001

KRAMSCH, C. **Context and culture in language teaching**. Oxford: Oxford University Press, 1993.

LAHZA, H.; KHOSRAVI, H.; DEMARTINI, G. Analytics of learning tactics and strategies in an online learner sourcing environment. **Journal of Computer Assisted Learning**, [S.I.], v. 39, n. 1, p. 94-112, 2023. DOI:10.1111/jcal.12729.

LEE, J. J.; HAMMER, J. Gamification in education: What, how, why bother? **Academic Exchange Quarterly**, [S.I.], v. 15, n. 2, 2011.

LEVY, M.; STOCKWELL, G. **Call dimensions: Options and issues in computer-assisted language learning**. Mahwah: Lawrence Erlbaum, 2006.

MAGAL-ROYO, T.; GARCÍA LABORDA, J. Una aproximación del efecto en el aprendizaje de una lengua extranjera debida a la obtención de datos a través de exámenes en línea de idiomas. **RED. Revista de Educación a Distancia**, [S.I.], v. 53, p. 1-14, 2017. Available at: <https://revistas.um.es/red/article/view/289881>. Accessed in: 14 Nov. 2023.

MAGAL-ROYO, T.; GARCÍA LABORDA, J. Standardization of design interfaces applied to language test on-line through ubiquitous devices. **International Journal of Interactive Mobile Technologies (iJIM)**, [S.I.], v. 12, n. 4, p. 21-31, 2018. DOI:10.3991/ijim.v12i4.9197.

MAYER, R. E. **Multimedia learning**. Cambridge: Cambridge University Press, 2001. DOI:10.1017/CBO9781139164603.

NERI, A.; CUCCHIARINI, C.; STRIK, H.; BOVES, L. The pedagogy-technology interface in Computer Assisted Pronunciation Training. **Computer Assisted Language Learning**, [S.I.], v. 15, n. 5, p. 441-467, 2002. DOI:10.1076/call.15.5.441.13473

O'DOWD, R. **Telecollaboration and the development of intercultural communicative competence**. Munich: Langenscheidt-Longman, 2006. ISBN: 9783526508359.

O'SHEA, K.; BANDAR, Z.; CROCKETT, K. A survey of chatbot systems through a Loebner prize competition. **Artificial Intelligence Review**, [S.I.], v. 53, n. 8, p. 5613-5649, 2020. DOI:10.1007/s10462-020-09870-7.

OUYANG, F.; WU, M.; ZHENG, L.; ZHANG, L.; JIAO, P. Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. **International Journal of Educational Technology in Higher Education**, [S.I.], v. 20, n. 4, 2023. DOI:10.1186/s41239-022-00372-4.

PDFGEAR, COM. **Intelectualizar y chatear archivos PDF con Asistente de IA**. Disponible en: <https://www.pdfgear.com>. Acceso el: 14 Nov. 2023.

QUIZLET.COM. **Quizlet/Q-Chat**. Available at: <https://quizlet.com>. Accessed in: 14 Nov. 2023.

SONDEREGGER, S. How generative language models can enhance interactive learning with social robots. *In: INTERNATIONAL CONFERENCE ON COGNITION AND EXPLORATORY LEARNING IN DIGITAL AGE*, 19., 2022. **Proceedings** [...]. [S. l.: s. n.], 2022. ISBN: 978-989-8704-43-6.

STORCH, N. Collaborative writing: Product, process, and students' reflections. **Journal of Second Language Writing**, [S.I.], v. 14, n. 3, p. 153-173, 2005.

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, L.; POLOSUKHIN, I. Attention is all you need. **Advances in Neural Information Processing Systems**, [S.I.], v. 30, 2017, p. 5998-6008. ISBN: 9781510860964. Available at: <https://arxiv.org/abs/1706.03762>. Accessed in: 14 Nov. 2023.

WARSCHAUER, M.; HEALEY, D. Computers and language learning: An overview. **Language Teaching**, [S.I.], v. 31, n. 2, p. 57-71, 1998.

WARSCHAUER, M. Invited commentary: New tools for teaching writing. **Language Learning & Technology**, [S.I.], v. 14, n. 1, p. 3-8, 2010. Available at: <http://lt.msu.edu/vol14num1/commentary.pdf>. Accessed in: 14 Nov. 2023.

WRITEMAGE. Available at: <https://writemage.com/>. Accessed in: 14 Nov. 2023.

XU, J.; PENG, H. Investigating mobile-assisted oral feedback in teaching Chinese as a second language. **Computer Assisted Language Learning**, [S.I.], v. 30, n. 3-4, p. 173-182, 2017.

ZAWACKI-RICHTER, O.; MARÍN, V. I.; BOND, M.; GOUVERNEUR, F. Systematic review of research on artificial intelligence applications in higher education – where are the educators? **International Journal of Educational Technology in Higher Education**, [S.I.], v. 16, n. 1, p. 1-27, 2019. DOI:10.1186/s41239-019-0171-0.

Acknowledgements: Not applicable.

Funding: Dr. Jesús García Laborda would like to express his gratitude to the Vice-Rectorate of the Universidad de Alcalá for providing funding for this research through the EPU-INV-UAH/2022/004 Research Program.

Conflicts of interest: Not applicable.

Ethical approval: Not applicable.

Availability of data and material: Not applicable.

Author contributions: Jesús García Laborda is the framework designer and theoretical supervisor; Teresa Magal worked on the implementation section of chatbot usage; and Miguel Fernandez was responsible for the implementation of taxonomy. All three authors wrote their respective sections and added the appropriate references.

Processing and editing: Editora Ibero-Americana de Educação.
Proofreading, formatting, normalization and translation.

