

Innovación pedagógica con inteligencia artificial generativa para fortalecer las habilidades productivas en inglés en estudiantes universitarios

Pedagogical Innovation with Generative Artificial Intelligence to Enhance Productive English Skills in University Students

<https://doi.org/10.5281/zenodo.17857927>

AUTORES: Alba Paulette León Morán¹

Kerly Jazmín Feijoo Rojas²

Maira María Rodríguez Torres³

Elma Verónica Ramírez Romero⁴

DIRECCIÓN PARA CORRESPONDENCIA: aleonm@utb.edu.ec

Fecha de recepción: 10 / 09 / 2025

Fecha de aceptación: 10 / 11 / 2025

RESUMEN

Este estudio analiza el impacto de una intervención pedagógica basada en herramientas de inteligencia artificial generativa (IA), como ChatGPT, Grammarly y QuillBot, en el desarrollo de habilidades productivas en inglés —*speaking* y *writing*— en estudiantes universitarios de nivel A2. Se empleó un diseño cuasi-experimental con grupos paralelos no aleatorizados (experimental y control) en una universidad pública ecuatoriana. La intervención, desarrollada durante ocho semanas, integró el uso pedagógico de herramientas de IA en tareas orales y escritas. Se aplicaron pruebas pretest y postest, rúbricas de evaluación según el MCER y un cuestionario tipo Likert para conocer la

¹ UNIVERSIDAD TÉCNICA DE BABAHOYO ORCID (0009-0004-4868-4058) aleonm@utb.edu.ec

² UNIVERSIDAD TÉCNICA DE BABAHOYO ORCID (0000-0002-3089-6739) kfeijoo@utb.edu.ec

³ UNIVERSIDAD TÉCNICA DE BABAHOYO ORCID (0000-0002-0590-5007) mrodriguezt@utb.edu.ec

⁴ UNIVERSIDAD TÉCNICA DE BABAHOYO ORCID (0000-0001-7799-9084) eramirezr@utb.edu.ec

percepción estudiantil. Los resultados revelaron mejoras estadísticamente significativas en el grupo experimental, tanto en *writing* como en *speaking*, en comparación con el grupo de control. Además, los subgrupos con mayor participación y uso frecuente de IA obtuvieron mejores desempeños. La percepción estudiantil fue mayoritariamente positiva, destacando beneficios como la retroalimentación inmediata, la mejora en la organización de ideas y el incremento de la autonomía. Sin embargo, también se identificaron limitaciones, como la necesidad de guía docente y el riesgo de dependencia automatizada. El estudio aporta evidencia empírica desde un contexto latinoamericano sobre el potencial de la IA generativa como herramienta para mejorar la enseñanza del inglés en educación superior, subrayando la necesidad de integrarla de forma ética, crítica y pedagógicamente estructurada.

Palabras clave: inteligencia artificial, enseñanza del inglés, habilidades productivas, educación superior, innovación pedagógica

ABSTRACT

This study analyzes the impact of a pedagogical intervention based on generative artificial intelligence (AI) tools—such as ChatGPT, Grammarly, and QuillBot—on the development of productive English skills (speaking and writing) in A2-level university students. A quasi-experimental design with non-randomized parallel groups (experimental and control) was used at a public university in Ecuador. The intervention, carried out over eight weeks, integrated the pedagogical use of AI tools in oral and written tasks. Pretests and posttests were administered, along with evaluation rubrics aligned with the CEFR and a Likert-type questionnaire to assess student perceptions. The results revealed statistically significant improvements in the experimental group in both writing and speaking compared to the control group. Additionally, subgroups with higher engagement and frequent AI use showed better performance. Student perceptions were mostly positive, highlighting benefits such as immediate feedback, improved organization of ideas, and increased autonomy. However, some limitations were also identified, including the need for teacher guidance and the risk of overreliance on automation. The study provides empirical evidence from a Latin American context on the potential of generative AI as a tool to enhance English

language teaching in higher education, emphasizing the need for its ethical, critical, and pedagogically structured integration.

Keywords: artificial intelligence, English teaching, productive skills, higher education, pedagogical innovation

INTRODUCCIÓN

En la era de la transformación digital, la educación se enfrenta a nuevos desafíos y oportunidades derivados de la incorporación de tecnologías emergentes. Una de las más influyentes en la actualidad es la inteligencia artificial (IA), cuyo impacto en el ámbito educativo ha sido objeto de estudio por su capacidad de personalizar el aprendizaje, mejorar la eficiencia en la retroalimentación y fomentar la autonomía del estudiante (Holmes et al., 2022; Luckin et al., 2016).

En este contexto, la IA generativa —particularmente a través de modelos lingüísticos como ChatGPT— ha abierto posibilidades inéditas en la producción de textos, la simulación de diálogos y la mejora de habilidades comunicativas, lo que representa un cambio de paradigma en las estrategias pedagógicas tradicionales (Zawacki-Richter et al., 2019).

Desde una perspectiva regional, las universidades latinoamericanas han avanzado en la digitalización de sus procesos formativos, especialmente a raíz de la pandemia por COVID-19, pero aún enfrentan limitaciones estructurales y metodológicas para integrar de forma efectiva las tecnologías emergentes en el aula (Selwyn, 2022). En el caso específico de la enseñanza del inglés como lengua extranjera (EFL), el desarrollo de competencias productivas como el speaking y el writing continúa siendo una de las mayores debilidades entre los estudiantes, especialmente en niveles preintermedios como el A2 (Council of Europe, 2020). Esta brecha se relaciona no solo con factores curriculares, sino también con enfoques pedagógicos desactualizados, limitaciones en la exposición al idioma y escasa retroalimentación personalizada (Rodríguez & Pérez, 2021). A pesar de que herramientas como ChatGPT, Grammarly y QuillBot están al alcance de los estudiantes universitarios, su

uso se ha mantenido mayoritariamente en espacios informales, sin un acompañamiento didáctico ni criterios claros de integración curricular.

Esta situación genera una desconexión entre el potencial tecnológico disponible y su aplicación real en el proceso de enseñanza-aprendizaje, lo que impide la sistematización de experiencias y la evaluación de sus beneficios pedagógicos. En consecuencia, los docentes continúan implementando prácticas tradicionales que dependen de la corrección directa, el aprendizaje memorístico y la participación limitada del estudiante en la construcción del conocimiento (Kohnke & Jarvis, 2021). Esta resistencia o falta de conocimiento sobre el uso educativo de la IA genera una oportunidad crítica de intervención.

Además, se observa una carencia de estudios empíricos en contextos latinoamericanos que analicen rigurosamente el impacto de la IA generativa en la mejora de habilidades productivas en inglés, particularmente en entornos universitarios donde se espera que el dominio del idioma contribuya a la movilidad académica, el acceso a fuentes científicas y la inserción laboral global (Salinas & Cárdenas, 2022). La ausencia de políticas institucionales para regular, fomentar o incluso monitorear el uso de estas herramientas contribuye a la improvisación, el uso éticamente cuestionable y, en ocasiones, a una dependencia tecnológica sin fundamentos pedagógicos sólidos.

Ante este panorama, la presente investigación se justifica en su intención de llenar un vacío en la literatura académica y aportar evidencia cuantitativa sobre los efectos del uso de IA generativa en la enseñanza del inglés, específicamente en la producción escrita y oral. Al tratarse de habilidades activas que requieren no solo conocimiento lingüístico, sino también capacidad de argumentación, organización del discurso, fluidez y corrección, la intervención con IA ofrece una vía para optimizar la práctica constante, facilitar el aprendizaje autónomo y promover el desarrollo de la metacognición lingüística (Fitria, 2023; Kohnke & Jarvis, 2021).

Además, esta propuesta se alinea con los principios de la pedagogía innovadora, la cual promueve el uso consciente de la tecnología como una mediación que potencia las competencias del siglo XXI: pensamiento crítico, autonomía, competencia digital, y comunicación efectiva. Al evaluar empíricamente el impacto de herramientas específicas, el estudio no solo podrá sustentar decisiones pedagógicas basadas en evidencia, sino también proporcionar orientaciones concretas para docentes y diseñadores curriculares que busquen integrar la IA de manera ética, efectiva y contextualizada.

Por lo cual, el objetivo de este estudio es analizar el efecto de una intervención pedagógica basada en herramientas de inteligencia artificial generativa en el desarrollo de las habilidades productivas en inglés (speaking y writing) en estudiantes universitarios de nivel A2. Y entre los objetivos específicos esta, evaluar el rendimiento de los estudiantes en tareas de producción oral y escrita antes y después de la intervención con IA generativa, Comparar los resultados entre un grupo experimental que utiliza herramientas de IA y un grupo de control que emplea métodos tradicionales, e identificar fortalezas y limitaciones del uso de IA generativa en contextos reales de enseñanza del inglés.

Esta investigación se enmarca en un estudio cuasi-experimental desarrollado durante un período de seis a ocho semanas en una universidad pública de Ecuador. Participan dos grupos paralelos no aleatorizados de estudiantes universitarios inscritos en asignaturas de inglés de nivel A2 del Marco Común Europeo de Referencia para las Lenguas (MCER). El grupo experimental empleará herramientas de IA generativa (ChatGPT, Grammarly y QuillBot) para realizar tareas de producción oral y escrita, mientras que el grupo de control utilizará métodos tradicionales como el uso de diccionarios, corrección del docente y ejercicios estructurados.

El estudio se limita a evaluar únicamente las habilidades productivas del idioma inglés (writing y speaking), sin incluir la comprensión lectora ni auditiva. Asimismo, se centrará en los resultados inmediatos de la intervención, sin abordar impactos a largo plazo ni variables actitudinales profundas como la percepción, la ética del uso de IA o la motivación

estudiantil. Los resultados se circunscriben al contexto universitario ecuatoriano, por lo que su generalización a otros entornos requiere precaución.

El uso de inteligencia artificial en la educación ha ganado interés académico en los últimos años, con especial énfasis en la IA generativa como aliada en la enseñanza de lenguas extranjeras. Investigaciones recientes han explorado la efectividad de herramientas como Grammarly, ChatGPT y QuillBot en la mejora de habilidades lingüísticas, revelando avances significativos en la corrección de errores gramaticales, enriquecimiento del vocabulario y desarrollo del pensamiento crítico (Kohnke & Jarvis, 2021; Fitria, 2023).

Por ejemplo, Kohnke y Jarvis (2021) llevaron a cabo un estudio con estudiantes universitarios en Hong Kong, donde evaluaron el impacto de Grammarly en la calidad de textos escritos. Los resultados demostraron una mejora significativa en la precisión gramatical y la cohesión textual, especialmente en estudiantes con niveles intermedios. De manera similar, Choi et al. (2023) analizaron la interacción de estudiantes con ChatGPT como tutor conversacional, identificando un aumento en la confianza comunicativa y la fluidez oral, gracias a la retroalimentación instantánea y sin juicio.

En el caso de QuillBot, su uso como revisor de estilo ha sido analizado en estudios de escritura académica en contextos no nativos. Bui y Kong (2023) reportaron que los estudiantes universitarios que utilizaron esta herramienta durante la redacción de ensayos mejoraron su claridad y variación léxica, aunque también advirtieron la necesidad de acompañar su uso con orientación docente para evitar la dependencia automatizada.

Sin embargo, a pesar de estos avances, la mayoría de los estudios se han desarrollado en contextos anglosajones o asiáticos, con poco enfoque en Latinoamérica, y con escasa sistematización sobre su aplicación conjunta en el aula como parte de un diseño pedagógico integral. Además, las investigaciones existentes tienden a centrarse en habilidades aisladas (como gramática o vocabulario) y no en habilidades productivas complejas como la escritura discursiva o la expresión oral argumentativa (Holmes et al., 2022).

Asimismo, existe una brecha en el diseño de estudios experimentales o cuasi-experimentales que permitan comparar de manera rigurosa el rendimiento de estudiantes que utilizan IA frente a aquellos que siguen métodos tradicionales. La literatura señala la necesidad de estudios que analicen el impacto de estas herramientas en contextos reales de enseñanza, con poblaciones diversas y bajo marcos normativos como el MCER (Zawacki-Richter et al., 2019).

En este sentido, el presente estudio se propone como una contribución original al estado del arte, al integrar herramientas de IA generativa en tareas de producción oral y escrita en un curso universitario de inglés nivel A2, aplicando un enfoque cuasi-experimental con evaluación pre y postest. La investigación busca aportar evidencia empírica desde un contexto latinoamericano, que permita orientar futuras decisiones metodológicas, curriculares y tecnológicas en la enseñanza del inglés como lengua extranjera.

1. La innovación pedagógica en la educación superior

La innovación pedagógica es entendida como la implementación de métodos, estrategias o herramientas que transforman y mejoran los procesos de enseñanza-aprendizaje, promoviendo el pensamiento crítico, la participación activa y la autonomía del estudiante (Area, 2009). En la educación superior, esta innovación implica superar modelos tradicionales centrados en el docente para adoptar enfoques más dinámicos, flexibles y personalizados. Las universidades, como espacios de formación integral, enfrentan el reto de integrar tecnologías emergentes con sentido pedagógico, no solo como elementos técnicos, sino como dispositivos transformadores del aprendizaje (Cabero & Llorente, 2015).

Diversas investigaciones destacan que la innovación debe estar alineada con el desarrollo de competencias del siglo XXI, como la creatividad, la colaboración y la competencia digital (Redecker & Punie, 2017). En este marco, las herramientas de inteligencia artificial —cuando son integradas intencionadamente en la práctica docente— representan una

oportunidad para reconfigurar las dinámicas de producción y retroalimentación en el aula universitaria.

2. La inteligencia artificial generativa en la enseñanza del inglés

La inteligencia artificial generativa (IAG) hace referencia a sistemas de aprendizaje automático capaces de producir contenido nuevo y coherente a partir de entradas humanas. En el campo del lenguaje, modelos como ChatGPT permiten generar textos, responder preguntas, sugerir correcciones y simular conversaciones en inglés, lo que ha ampliado las posibilidades de práctica para los estudiantes de lenguas extranjeras (Holmes et al., 2022). Herramientas como Grammarly y QuillBot se han convertido en asistentes útiles para mejorar la calidad gramatical, la cohesión y el estilo de textos académicos o personales (Kohnke & Jarvis, 2021).

Estas herramientas permiten una interacción personalizada, retroalimentación inmediata y un entorno de aprendizaje activo, lo cual es especialmente valioso en el desarrollo de habilidades productivas como el writing y el speaking. La literatura sugiere que su uso puede generar mayor confianza comunicativa, mejorar la autopercepción del estudiante y potenciar el aprendizaje autónomo (Fitria, 2023; Bui & Kong, 2023). Sin embargo, se señala también que su uso debe ser mediado pedagógicamente para evitar la dependencia, la superficialidad en el aprendizaje y el uso mecánico sin reflexión lingüística (Choi et al., 2023). La integración de la IAG requiere una guía docente que oriente el uso crítico, ético y significativo de estas tecnologías en el aula.

3. El desarrollo de habilidades productivas en inglés

Las habilidades productivas en el aprendizaje del inglés —writing y speaking— se consideran fundamentales para la competencia comunicativa, ya que implican la producción activa del lenguaje con propósito comunicativo real (Council of Europe, 2020). El desarrollo de estas destrezas requiere no solo conocimiento gramatical y léxico, sino también estrategias discursivas, fluidez, coherencia textual y capacidad de autorregulación en el uso del idioma (Brown & Lee, 2015).

En el caso del writing, se espera que el estudiante sea capaz de organizar ideas, redactar con claridad, corregir errores y utilizar un vocabulario apropiado al nivel. En cuanto al speaking, se valora la fluidez, la pronunciación, la coherencia de ideas y la capacidad para sostener una conversación o presentación (Richards, 2008). La evaluación de estas habilidades suele realizarse mediante rúbricas estandarizadas basadas en el Marco Común Europeo de Referencia para las Lenguas (MCER), que establecen descriptores por niveles (Council of Europe, 2020).

Diversos estudios coinciden en que los métodos tradicionales, centrados en ejercicios repetitivos o corrección unidireccional, resultan insuficientes para estimular el uso real del idioma. Por ello, se promueve el uso de tecnologías como mediadoras del aprendizaje productivo, permitiendo al estudiante practicar de forma autónoma, recibir retroalimentación y construir su conocimiento de manera activa (Blake, 2013).

4. Enfoques quasi-experimentales en la evaluación de innovaciones educativas

Los diseños quasi-experimentales permiten evaluar el impacto de una intervención pedagógica en contextos reales de aula, especialmente cuando no es posible asignar aleatoriamente a los participantes (Cohen et al., 2018). Este tipo de investigación es particularmente útil en entornos educativos donde existen condiciones preestablecidas, como los paralelos de clase, y donde se busca comparar el desempeño entre un grupo que recibe la innovación y otro que sigue con métodos tradicionales.

En el caso de la enseñanza del inglés, múltiples estudios han adoptado este enfoque para medir mejoras en producción escrita y oral, el desarrollo del vocabulario o la percepción de los estudiantes frente a metodologías activas (Yang et al., 2021). El uso de pruebas pretest y postest, junto con instrumentos estandarizados como rúbricas basadas en el MCER, permiten obtener resultados comparables y estadísticamente significativos. La aplicación de la prueba *t* de Student —tanto pareada como para muestras independientes— es común para establecer si las diferencias observadas en el rendimiento son atribuibles a la

intervención (Field, 2018). El presente estudio se enmarca dentro de esta tradición investigativa, al proponer una evaluación rigurosa del uso de IA generativa como innovación pedagógica para el desarrollo de habilidades productivas en inglés, mediante un diseño cuasi-experimental con evaluación antes y después de la intervención.

METODOLOGÍA

Esta investigación se desarrolló bajo un enfoque cuantitativo, con un diseño cuasi-experimental comparativo, y tuvo como objetivo analizar el efecto del uso de herramientas de inteligencia artificial generativa en el desarrollo de las habilidades productivas en inglés (*speaking* y *writing*) en estudiantes universitarios con nivel A2. El estudio se llevó a cabo durante un periodo de ocho semanas, en un entorno académico real, utilizando pruebas estandarizadas y recolección de datos con apoyo de escalas de percepción tipo Likert.

Diseño de la investigación

El estudio se enmarcó dentro de un diseño cuantitativo cuasi-experimental de tipo comparativo, con aplicación de pretest y postest, empleando dos grupos no aleatorizados: uno experimental y otro de control. Esta elección metodológica respondió a la necesidad de evaluar el efecto de una intervención pedagógica con herramientas de inteligencia artificial generativa en condiciones naturales de aula, donde la asignación aleatoria no era viable debido a la organización institucional preexistente (Cohen et al., 2018).

El grupo experimental recibió una intervención basada en el uso de herramientas de IA generativa (ChatGPT, Grammarly y QuillBot) durante el desarrollo de tareas de producción oral y escrita en inglés, mientras que el grupo de control trabajó los mismos contenidos utilizando metodologías tradicionales sin apoyo tecnológico. Ambos grupos fueron evaluados antes y después de la intervención mediante pruebas estandarizadas, lo que permitió analizar los cambios en su rendimiento y comparar los efectos del enfoque pedagógico aplicado.

Se adoptó un diseño de grupo intacto con intervención didáctica planificada y evaluación cuantitativa de resultados, adecuado para contextos educativos reales donde no se puede controlar totalmente la asignación de participantes. La estructura del diseño incluyó:

- Pretest aplicado a ambos grupos para establecer la línea base de las habilidades productivas en inglés (speaking y writing).
- Intervención de 8 semanas en el grupo experimental, integrando el uso pedagógico de IA generativa en las tareas lingüísticas.
- Postest al finalizar la intervención, con el fin de medir posibles diferencias en el desarrollo de habilidades productivas.
- Cuestionario de percepción tipo Likert aplicado únicamente al grupo experimental para complementar el análisis cuantitativo con datos subjetivos sobre la experiencia de uso de IA.

Participantes

La población del estudio estuvo compuesta por 150 estudiantes universitarios matriculados en asignaturas de inglés de segundo nivel, correspondientes al nivel A2 del Marco Común Europeo de Referencia para las Lenguas (MCER). Se trabajó con dos paralelos existentes: uno asignado como grupo experimental (75 estudiantes), y otro como grupo de control (75 estudiantes). Ambos grupos compartían características similares en cuanto a edad, formación académica, nivel lingüístico y modalidad de estudio.

Instrumentos y técnicas de recolección de datos

Se utilizaron tres instrumentos principales para recopilar los datos:

1. **Prueba escrita (writing):** se aplicó una tarea de redacción con un límite de 150–200 palabras, valorada mediante una rúbrica basada en el MCER que consideró cinco criterios: coherencia, cohesión, corrección gramatical, riqueza léxica y estructura textual.
2. **Actividad oral (speaking):** se implementó una simulación de entrevista o presentación oral individual, evaluada mediante una rúbrica que midió fluidez, pronunciación, organización del discurso y adecuación al nivel A2.

3. **Cuestionario de percepción estudiantil:** se aplicó al finalizar la intervención un cuestionario estructurado con ítems tipo Likert de 5 puntos (1 = totalmente en desacuerdo; 5 = totalmente de acuerdo), que midió la percepción del grupo experimental respecto a la utilidad, facilidad de uso, motivación y apoyo al aprendizaje que ofrecieron las herramientas de IA utilizadas (ChatGPT, Grammarly y QuillBot).

Intervención pedagógica

Se desarrolló a lo largo de ocho semanas en el grupo experimental. Los estudiantes recibieron inicialmente una capacitación básica sobre el uso responsable y académico de herramientas de inteligencia artificial generativa. Durante la intervención, se incorporaron estas herramientas como parte de las tareas de producción oral y escrita, bajo la guía del docente. ChatGPT fue utilizado para generar ideas, simular diálogos y recibir ejemplos lingüísticos; Grammarly para revisar redacciones y sugerir mejoras gramaticales; y QuillBot para explorar opciones de parafraseo y estilo.

En contraste, el grupo de control completó tareas equivalentes utilizando métodos convencionales: diccionarios impresos, corrección directa del docente y actividades estructuradas sin mediación tecnológica. Ambos grupos realizaron las mismas actividades en cuanto a objetivos lingüísticos, pero con distinto enfoque metodológico.

Tabla 1.

Cronograma de la intervención pedagógica con herramientas de inteligencia artificial generativa

Semana	Actividad Principal
Semana 1	Inducción y capacitación sobre el uso pedagógico de ChatGPT, Grammarly y QuillBot.
Semana 2	Producción escrita 1: redacción de un párrafo descriptivo con apoyo de Grammarly.
Semana 3	Producción oral 1: práctica de presentación personal utilizando ChatGPT como simulador.

Semana 4	Producción escrita 2: redacción de un correo formal con corrección asistida por QuillBot.
Semana 5	Producción oral 2: simulación de entrevista académica con apoyo de ChatGPT.
Semana 6	Producción escrita 3: redacción de un texto narrativo y revisión automatizada.
Semana 7	Producción oral 3: exposición breve sobre tema académico con retroalimentación entre pares.
Semana 8	Evaluación final: aplicación de postest escrito y oral + encuesta de percepción (Likert).

Nota. Las actividades fueron desarrolladas semanalmente en el grupo experimental como parte de una intervención de ocho semanas orientada a fortalecer las habilidades productivas en inglés a través del uso pedagógico de herramientas de IA generativa.

Análisis de datos

Tabla 2.

Resumen del análisis estadístico aplicado en el estudio

Prueba Estadística	Propósito	Variables Analizadas	Software Utilizado
Prueba t para muestras relacionadas	Comparar el rendimiento del pretest y postest dentro de cada grupo (experimental y control)	Puntajes pretest y postest en writing y speaking	SPSS
Prueba t para muestras independientes	Comparar el rendimiento del postest entre grupo experimental y grupo de control	Puntajes postest en writing y speaking	SPSS
ANOVA de una vía	Analizar diferencias dentro del grupo experimental según variables como participación, frecuencia de uso	Puntajes postest vs. subgrupos por participación/frecuencia/género	SPSS

Analisis descriptivo con escala Likert	y género	Evaluar la percepción estudiantil sobre el uso de herramientas de IA generativa	Respuestas en escala Likert (1-5) sobre utilidad, facilidad, motivación y retroalimentación	SPSS
---	----------	---	---	------

Nota: Todas las pruebas estadísticas fueron procesadas mediante el software SPSS, estableciendo un nivel de significancia de $p < 0.05$ para la interpretación de resultados.

Se presenta un resumen de las pruebas estadísticas utilizadas en el análisis de los datos obtenidos durante la intervención pedagógica. Se aplicaron tanto pruebas inferenciales como descriptivas, en función de los objetivos específicos de la investigación. La prueba t para muestras relacionadas se utilizó para comparar el rendimiento del pretest y postest dentro de cada grupo (experimental y control), mientras que la prueba t para muestras independientes permitió contrastar el desempeño final entre ambos grupos.

Adicionalmente, se empleó un ANOVA de una vía para identificar diferencias significativas entre subgrupos del grupo experimental según variables como nivel de participación, frecuencia de uso de herramientas de IA y género. Finalmente, se realizó un análisis descriptivo de los resultados del cuestionario tipo Likert, permitiendo evaluar la percepción estudiantil sobre el uso de IA generativa en el aprendizaje de inglés.

Tabla 3.

Comparación de medias pretest y postest en habilidades productivas en inglés del grupo experimental

Grupo	Media	Desviación Estándar	N.Estudantes
Experimental - Pretest Writing	6.2	0.8	75
Experimental - Postest Writing	8.1	0.7	75
Experimental - Pretest Speaking	6.0	0.9	75
Experimental - Postest Speaking	7.9	0.6	75

Nota: Los resultados corresponden a una muestra de 75 estudiantes universitarios. La mejora observada fue confirmada estadísticamente mediante la prueba t para muestras relacionadas con un nivel de significancia de $p < 0.05$.

Se muestran los puntajes obtenidos por el grupo experimental en las pruebas de producción escrita (*writing*) y producción oral (*speaking*), aplicadas antes y después de la intervención pedagógica con inteligencia artificial generativa. Se observa un incremento significativo en ambas habilidades tras la intervención. En *writing*, la media pasó de 6.2 (DE = 0.8) en el pretest a 8.1 (DE = 0.7) en el postest, mientras que en *speaking* se incrementó de 6.0 (DE = 0.9) a 7.9 (DE = 0.6). Estos resultados reflejan una mejora general en el rendimiento del grupo experimental, sugiriendo que el uso de herramientas como ChatGPT, Grammarly y QuillBot tuvo un efecto positivo en el desarrollo de las habilidades productivas del idioma inglés.

Los valores más altos en el postest, junto con una reducción en la desviación estándar, indican además una mayor homogeneidad en el rendimiento de los estudiantes, lo que podría estar relacionado con el apoyo personalizado y la retroalimentación continua que ofrecieron las herramientas de IA durante la intervención. Para garantizar la confiabilidad y validez de los instrumentos utilizados, se aplicó una prueba piloto con un grupo similar de estudiantes, lo cual permitió ajustar ítems ambiguos y mejorar la claridad de las instrucciones. En el caso del cuestionario de percepción, se calculó el coeficiente alfa de Cronbach, obteniéndose un valor de 0.87, lo que indica una alta consistencia interna.

Tabla 4.

Resultados comparativos del pretest y postest del grupo experimental

Grupo / Subgrupo	Media	Desviación Estándar	N (Estudiantes)
Control - Postest Writing	6.5	0.7	75
Control - Postest Speaking	6.4	0.8	75

Nota. La diferencia entre pretest y postest fue analizada con la prueba t para muestras relacionadas, estableciendo un nivel de significancia de $p < 0.05$.

Se muestra los puntajes promedio obtenidos por los estudiantes del grupo experimental en las pruebas de producción escrita (*writing*) y oral (*speaking*) antes y después de la intervención pedagógica basada en el uso de herramientas de inteligencia artificial generativa. Se observa un incremento significativo en ambos indicadores tras la

implementación de la estrategia innovadora, evidenciando mejoras en la competencia productiva del idioma inglés en nivel A2.

Tabla 5.

Resultados del postest según nivel de participación y frecuencia de uso de IA generativa en el grupo experimental

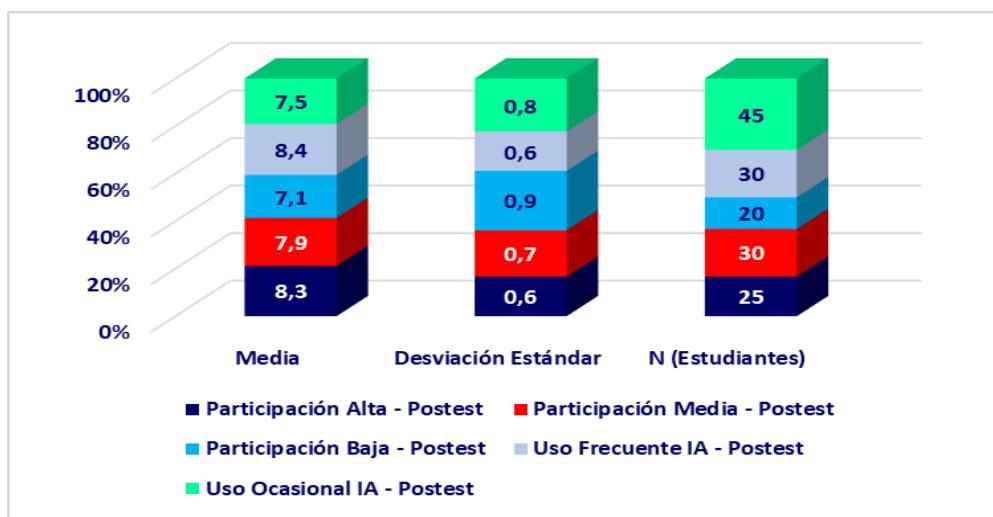
Subgrupo	Media	Desviación Estándar	N (Estudiantes)
Participación Alta - Postest	8.3	0.6	25
Participación Media - Postest	7.9	0.7	30
Participación Baja - Postest	7.1	0.9	20
Uso Frecuente IA - Postest	8.4	0.6	30
Uso Ocasional IA - Postest	7.5	0.8	45

Nota. La comparación entre subgrupos fue analizada mediante un ANOVA de una vía, mostrando diferencias estadísticamente significativas ($p < 0.05$).

La tabla presenta los puntajes promedio obtenidos por subgrupos del grupo experimental en la prueba postest de habilidades productivas en inglés, diferenciados según dos variables: el nivel de participación en las actividades durante la intervención y la frecuencia de uso de herramientas de inteligencia artificial generativa.

Figura 1.

Desempeño postest en habilidades productivas según nivel de participación y frecuencia de uso de herramientas de IA



Nota. Los valores indican que los estudiantes con alta participación y uso frecuente obtuvieron mejores resultados en el postest.

Se evidencia que los estudiantes con mayor participación activa y uso frecuente de IA obtuvieron mejores resultados (medias de 8.3 y 8.4 respectivamente), en comparación con aquellos que participaron menos o usaron ocasionalmente las herramientas. Este comportamiento sugiere que la implicación activa y la práctica constante con IA tienen un efecto positivo en el desarrollo de las competencias productivas del idioma.

RESULTADOS

Con base en el objetivo general de analizar el efecto de una intervención pedagógica basada en herramientas de inteligencia artificial generativa en el desarrollo de las habilidades productivas en inglés (*speaking* y *writing*), se realizó una comparación de los puntajes obtenidos por los estudiantes universitarios de nivel A2 antes y después de dicha intervención, así como entre un grupo experimental y un grupo de control.

Los resultados muestran una mejora significativa en el rendimiento del grupo experimental en ambas habilidades productivas tras la intervención. En la habilidad de *writing*, la media pasó de 6.2 (DE = 0.8) en el pretest a 8.1 (DE = 0.7) en el postest. De forma similar, en *speaking*, la media aumentó de 6.0 (DE = 0.9) a 7.9 (DE = 0.6). La prueba t para muestras relacionadas confirmó que estas diferencias fueron estadísticamente significativas con un valor de $p < 0.05$, lo que respalda la efectividad de la intervención con herramientas de IA generativa como apoyo para mejorar la producción escrita y oral de los estudiantes.

La comparación entre los puntajes postest del grupo experimental y del grupo de control evidenció también diferencias significativas. Mientras que el grupo experimental alcanzó una media general de 8.0 en producción escrita y oral, el grupo de control obtuvo una media inferior de 6.45, con una desviación estándar ligeramente mayor. La prueba t para muestras independientes indicó diferencias estadísticamente significativas entre ambos grupos ($p < 0.05$), lo que sugiere que la incorporación de herramientas como ChatGPT, Grammarly y QuillBot generó un impacto positivo en el proceso de aprendizaje del inglés.

en el grupo experimental, en comparación con las metodologías tradicionales empleadas en el grupo de control.

Tabla 6.

Impacto del uso de herramientas de inteligencia artificial generativa (ChatGPT, Grammarly y QuillBot) en el desarrollo de habilidades productivas en inglés.

Herramienta de IA	Habilidad principal desarrollada	Aplicaciones observadas	Percepción estudiantil	Resultado académico
ChatGPT	<i>Speaking</i>	Simulación de conversaciones, práctica oral, interacción inmediata	78% consideró útil la interacción con ChatGPT para practicar conversaciones	Mejora en <i>speaking</i> de 6.0 (DE=0.9) a 7.9 (DE=0.6)
Grammarly	<i>Writing</i>	Corrección automática, sugerencias gramaticales, mejora de estilo y redacción	Alta percepción de mejora en claridad, coherencia y corrección de textos	Mejora en <i>writing</i> de 6.2 (DE=0.8) a 8.1 (DE=0.7)
QuillBot	<i>Writing</i>	Reformulación de oraciones, enriquecimiento de vocabulario, claridad de ideas	Se reportó mayor organización de ideas y apoyo para reestructurar textos complejos	Complementó el rendimiento en tareas de redacción

Nota: Los resultados muestran mejoras significativas en las habilidades de speaking y writing del grupo experimental tras la implementación de estas herramientas.

Los resultados del análisis por subgrupos dentro del grupo experimental revelaron diferencias internas asociadas al grado de participación y frecuencia de uso de las herramientas de IA. Los estudiantes con alta participación obtuvieron una media de 8.3 (DE = 0.6), frente a una media de 7.1 (DE = 0.9) en el grupo de baja participación. Asimismo, aquellos con uso frecuente de herramientas de IA alcanzaron una media de 8.4 (DE = 0.6), mientras que los usuarios ocasionales obtuvieron una media de 7.5 (DE = 0.8). El análisis ANOVA de una vía confirmó que estas diferencias fueron estadísticamente significativas ($p < 0.05$), reflejando que el grado de implicación activa y el uso continuo de IA están directamente asociados con mejores resultados en habilidades productivas.

Finalmente, los resultados del cuestionario de percepción tipo Likert aplicado al grupo experimental mostraron una valoración mayoritariamente positiva de la intervención. Un 82% de los estudiantes estuvieron de acuerdo o totalmente de acuerdo en que el uso de IA facilitó la mejora de su escritura en inglés, mientras que un 78% consideró útil la interacción con ChatGPT para practicar conversaciones. Además, se reportó una alta percepción de autonomía, retroalimentación inmediata y claridad en la organización de ideas, especialmente en tareas de redacción.

Para confirmar la validez estadística de las diferencias observadas entre los resultados pretest y postest, se aplicó una prueba t de Student para muestras relacionadas en ambos grupos. Los resultados mostraron diferencias estadísticamente significativas ($p < 0.05$) en las calificaciones del grupo experimental, tanto en la habilidad de speaking como en la de writing, lo que refuerza la efectividad de la intervención con herramientas de inteligencia artificial generativa.

Sin embargo, algunos estudiantes indicaron como limitación la necesidad de guía docente para interpretar adecuadamente las respuestas generadas por las herramientas, especialmente en estructuras gramaticales más complejas. En conjunto, los resultados sugieren que la intervención pedagógica con IA generativa no solo mejoró el desempeño académico en habilidades productivas, sino que también promovió la autonomía y la autoconfianza lingüística, siempre que existiera un uso responsable, frecuente y mediado por el docente.

DISCUSIÓN

Los resultados obtenidos en esta investigación confirman que la integración pedagógica de herramientas de inteligencia artificial generativa —como ChatGPT, Grammarly y QuillBot— puede tener un impacto significativo en el desarrollo de las habilidades productivas en inglés (*speaking* y *writing*) en estudiantes universitarios de nivel A2. Este hallazgo es consistente con investigaciones como las de Bui y Kong (2023), quienes

demostraron que el uso de QuillBot durante tareas de redacción mejoró la variedad léxica y la claridad expresiva en estudiantes universitarios no nativos.

El incremento observado en los puntajes del postest en el grupo experimental también respalda lo planteado por Kohnke y Jarvis (2021), quienes concluyeron que el uso de Grammarly permite una mejora sustancial en la precisión gramatical y la coherencia textual, especialmente cuando se integra como parte de un diseño pedagógico guiado. A su vez, Choi et al. (2023) evidenciaron que la interacción con ChatGPT incrementa la fluidez y la confianza comunicativa de los estudiantes en tareas orales, gracias a la retroalimentación inmediata y sin juicio.

En términos de eficacia comparativa, los resultados de esta investigación se alinean con lo señalado por Zawacki-Richter et al. (2019), quienes destacan que el uso de IA en educación superior es más efectivo cuando se basa en estrategias pedagógicas activas y se articula con objetivos lingüísticos claros. En este caso, los estudiantes que participaron activamente en la intervención y utilizaron con frecuencia las herramientas de IA mostraron avances significativamente superiores, confirmando lo planteado por Area (2009) sobre la importancia del compromiso activo del estudiante en procesos de innovación pedagógica.

Desde la perspectiva del aprendizaje autónomo, los hallazgos también confirman los principios de la pedagogía andragógica propuestos por Knowles (1984), especialmente en cuanto a la autorregulación, la toma de decisiones y el uso de recursos tecnológicos para resolver problemas reales de comunicación. La IA generativa no solo actuó como un corrector o asistente textual, sino como un mediador cognitivo que permitió a los estudiantes explorar, modificar y refinar sus producciones lingüísticas con autonomía.

En cuanto a la percepción estudiantil, los datos obtenidos mediante la escala de Likert coinciden con los hallazgos de Fitria (2023), quien señala que el uso de herramientas de IA promueve una mayor motivación en el aprendizaje del inglés cuando los estudiantes perciben un beneficio directo en su desempeño. No obstante, también se ratifican

advertencias expresadas por autores como Holmes et al. (2022) y Selwyn (2022), quienes alertan sobre el uso automatizado y acrítico de estas tecnologías, lo cual puede derivar en dependencia o pérdida de habilidades de autocorrección si no se cuenta con una mediación docente intencionada.

Al documentar de forma estructurada los efectos de su uso en un entorno universitario real, este estudio responde al llamado de Salinas y Cárdenas (2022) sobre la necesidad de fortalecer la investigación aplicada en innovación educativa en la región, especialmente en lo relativo a enseñanza de lenguas extranjeras con apoyo tecnológico.

CONCLUSIONES

La presente investigación permitió comprobar que la integración de herramientas de inteligencia artificial generativa en el proceso de enseñanza-aprendizaje del inglés como lengua extranjera tiene un impacto positivo y significativo en el desarrollo de las habilidades productivas —*speaking* y *writing*— en estudiantes universitarios de nivel A2. La intervención pedagógica basada en el uso de ChatGPT, Grammarly y QuillBot no solo fortaleció el rendimiento académico en ambas competencias, sino que también favoreció la autonomía, la retroalimentación inmediata y la mejora de la calidad lingüística de las producciones estudiantiles.

Los resultados demostraron mejoras sustanciales en los puntajes *posttest* del grupo experimental en comparación con el grupo de control, validando la eficacia de estas tecnologías como mediadoras del aprendizaje cuando se emplean dentro de un marco metodológico planificado y guiado por el docente. Además, el análisis por subgrupos reveló que la participación activa y el uso frecuente de las herramientas potenciaron aún más los logros alcanzados, lo que confirma la importancia del involucramiento del estudiante como factor determinante en el éxito de este tipo de innovaciones.

Asimismo, la percepción positiva del estudiantado sobre el uso de IA generativa refuerza su potencial como recurso pedagógico, siempre que su implementación sea ética,

contextualizada y complementaria a la enseñanza humana. No obstante, se identificaron limitaciones que deben considerarse en futuras aplicaciones, como la necesidad de capacitación docente, la regulación institucional del uso de IA en el aula, y el riesgo de dependencia si no se promueve una reflexión crítica sobre las respuestas generadas por estas herramientas.

Finalmente, se destaca que el uso de ChatGPT no se limita únicamente a la mejora de la escritura académica, sino que también constituye una herramienta valiosa para la práctica de la expresión oral (*speaking*), al permitir la interacción simulada en tiempo real tipo chat, lo cual favorece el desarrollo comunicativo autónomo en entornos virtuales.

REFERENCIAS BIBLIOGRÁFICAS

- Anderson, J. (2017). *Reflective teaching in second language classrooms*. Cambridge University Press.
- Beckett, G. H., & Miller, P. C. (Eds.). (2006). *Project-based second and foreign language education: Past, present, and future*. Information Age Publishing.
- Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. International Society for Technology in Education.
- Brown, H. D. (2007). *Principles of language learning and teaching* (5th ed.). Pearson Education.
- Celce-Murcia, M., Brinton, D. M., & Snow, M. A. (2014). *Teaching English as a second or foreign language* (4th ed.). National Geographic Learning.
- Chen, W., & Kent, S. (2020). The effects of digital storytelling on EFL students' speaking fluency. *Journal of Educational Technology Development and Exchange*, 13(1), 65–80. <https://doi.org/10.18785/jetde.1301.04>

Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge University Press.

EF Education First. (2022). *EF English Proficiency Index: A ranking of 111 countries and regions by English skills*. <https://www.ef.com/wwen/epi/>

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Godwin-Jones, R. (2018). Using mobile technology to develop language skills and cultural understanding. *Language Learning & Technology*, 22(3), 1–17.
<https://www.lltjournal.org>

Goh, C. C. M., & Burns, A. (2012). *Teaching speaking: A holistic approach*. Cambridge University Press.

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: a review of technology types and their effectiveness. *Computer Assisted Language Learning*, 27(1), 70–105.
<https://doi.org/10.1080/09588221.2012.700315>

Harmer, J. (2007). *The practice of English language teaching* (4th ed.). Pearson Longman.

Hockly, N., & Dudeney, G. (2018). *Current and future digital trends in ELT*. *RELATE Journal*, 35(2), 22–31.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132. <https://doi.org/10.1111/j.1540-4781.1986.tb05256.x>

Isbell, R. (2002). Telling and retelling stories: Learning language and literacy. *Young Children*, 57(2), 26–30.

Ladousse, G. P. (1987). *Role play*. Oxford University Press.

Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer-assisted language learning*. Routledge.

Ministerio de Educación del Ecuador. (2016). *Lineamientos para la implementación del currículo de inglés como lengua extranjera*. <https://educacion.gob.ec>

Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. Routledge.

Nunan, D. (2015). *Teaching English to speakers of other languages: An introduction*. Routledge.

Oxford, R. (1997). Cooperative learning, collaborative learning, and interaction: Three communicative strands in the language classroom. *The Modern Language Journal*, 81(4), 443–456. <https://doi.org/10.1111/j.1540-4781.1997.tb05510.x>

Richards, J. C. (2006). *Communicative language teaching today*. Cambridge University Press.

Stoller, F. L. (2006). Establishing a theoretical foundation for project-based learning in second and foreign language contexts. In G. H. Beckett & P. C. Miller (Eds.), *Project-based second and foreign language education: Past, present, and future* (pp. 19–40). Information Age Publishing.

Tajeddin, Z., & Aghababazadeh, F. (2018). Technology-enhanced task-based instruction: A critical analysis. *Teaching English with Technology*, 18(1), 22–40.

Thornbury, S. (2005). *How to teach speaking*. Pearson Education.

Werbach, K., & Hunter, D. (2012). *For the win: How game thinking can revolutionize your business*. Wharton Digital Press.

Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford University Press.