IA GENERATIVA Y RETROALIMENTACIÓN FORMATIVA EN LA ESCRITURA DE ALEMÁN L2 (A1): ESTUDIO CUASI-LONGITUDINAL[1]

GENERATIVE AI AND FORMATIVE FEEDBACK IN L2 GERMAN WRITING (A1): A QUASI-LONGITUDINAL STUDY

Almudena Mallo Dorado
Universidad del País Vasco / Euskal Herriko Unibertsitatea UPV/EHU

RESUMEN
Este estudio examina los efectos del uso de Inteligencia Artificial Generativa (IAGen) en la mejora de la producción escrita en alemán como lengua extranjera (nivel A1) en un entorno universitario. Mediante un diseño cuasi-experimental con grupos paralelos (experimental y control), se compararon dos intervenciones llevadas a cabo en la Universidad del País Vasco UPV/EHU durante el curso 2024–2025. En ambas, los estudiantes redactaron un texto sin apoyo externo; en el grupo experimental, revisaron el texto de un/a compañero/a con ayuda de *ChatGPT* antes de recibir retroalimentación docente, mientras que en el grupo de control el profesor proporcionó correcciones directas. Los resultados preliminares sugieren que el uso de IAGen favorece una actitud más consciente y reflexiva ante la lengua, aunque no permiten aún establecer conclusiones definitivas sobre su impacto. Ambas intervenciones forman parte de un análisis longitudinal que busca evaluar el potencial formativo y ético de la IAGen en contextos de enseñanza de lenguas extranjeras.

PALABRAS CLAVE: Inteligencia Artificial Generativa, enseñanza de alemán como lengua extranjera, producción escrita, pensamiento crítico, evaluación formativa

ABSTRACT
This study explores the impact of Generative Artificial Intelligence (GenAI) on the development of written production skills in German as a foreign language (A1 level) within a university context. A quasi-experimental design with parallel groups (experimental and control) was implemented at the University of the Basque Country UPV/EHU during the 2024–2025 academic year. In both groups, students composed a text without external assistance; the experimental group revised a peer's text using ChatGPT *before receiving teacher feedback, while the control group received direct corrections from the instructor. Preliminary findings suggest that GenAI-assisted revision fosters greater linguistic awareness and reflective*

[1] El presente trabajo ha sido posible gracias al apoyo del Gobierno Vasco (Language and Speech – LASLAB, IT1426-22; 2022–2025).

CC BY NC ND

engagement, although further data are needed to draw definitive conclusions about its effectiveness. Both interventions are part of a longitudinal study aimed at assessing the pedagogical and ethical implications of GenAI integration in foreign language writing instruction.

KEYWORDS: Generative Artificial Intelligence, teaching German as a foreign language, written production, critical thinking, formative assessment

1. INTRODUCCIÓN

La incorporación de la Inteligencia Artificial Generativa (IAGen) en contextos educativos de enseñanza y aprendizaje de lenguas no constituye el primer intento de trasladar a las máquinas procesos cognitivos humanos, sino que se inscribe en una trayectoria previa iniciada por la digitalización y consolidada por la traducción automática neuronal (NMT) a partir de 2016, con herramientas como *Google Translate* y *DeepL*, (Wang et al., 2022: 144). En este marco, la NMT se ha consolidado en los últimos años como una herramienta de uso habitual entre el alumnado para realizar tareas de escritura, a pesar de que se reconocen ampliamente sus limitaciones. Este uso suele circunscribirse por regla general a una mera reproducción parcial o íntegra de textos generados automáticamente (Jolley y Maimone, 2022: 39).

La integración sistemática, y con frecuencia acrítica, de la NMT ha dado lugar a un intenso debate académico en torno a la autoría, la evaluación y el desarrollo competencial, que se ha visto avivado por la reciente democratización de la IAGen, una tecnología capaz de asistir al estudiantado durante el proceso de escritura al resolver las dificultades inherentes a la producción textual – gramaticales, léxicas u ortográficas – y al ofrecer revisiones tanto de los borradores como del producto final. Como consecuencia, una parte considerable de los estudiantes de segundas lenguas recurre hoy a plataformas de IAGen – como *ChatGPT*, *Copilot* o *Gemini*– que generan, mediante respuestas inmediatas, textos aparentemente correctos.

Esta práctica ha provocado un notable disenso entre el profesorado de lenguas: mientras un sector propugna la prohibición del uso de IAGen, otro defiende una integración guiada y crítica, orientada tanto al desarrollo de competencias lingüísticas como de la alfabetización tecnológica (Muñoz-Basols et al., 2023: 173). Si bien la IAGen produce textos generalmente correctos en términos gramaticales, persisten errores semánticos frecuentemente vinculados a limitaciones en la representación de la diversidad cultural – como contextos específicos, traducciones complejas, matices culturales, expresiones idiomáticas o influencia del inglés (Román-Mendoza, 2023: 15). Tales limitaciones refuerzan la idea de que estas herramientas deben concebirse como complemento y no como sustituto de la retroalimentación humana (Navío-Inglés y Guzmán Mora, 2025: 252), lo que exige que el alumnado desarrolle habilidades críticas para valorar adecuadamente las sugerencias generadas.

La adopción de IAGen como metodología innovadora requiere, no obstante, de apoyo y regulación institucionales, de ahí que distintos agentes universitarios hayan instado a los responsables institucionales a establecer un marco normativo sólido que oriente su integración en la educación superior. En el contexto español, las universidades han comenzado a responder a esta demanda mediante la redacción de

CC BY NC ND

directrices institucionales[2] que parten generalmente del marco europeo vigente (Parlamento Europeo y Consejo de la Unión Europea, 2024) y pretenden garantizar un uso responsable de esta herramienta en los procesos de enseñanza. En esta línea, la UPV/EHU (2025) ha desarrollado un documento específico para regular la utilización de la IAGen como herramienta de apoyo a la producción escrita, en el que se exige la declaración explícita de su uso, se subraya que no debe sustituir la autoría ni el trabajo intelectual del estudiante y se enfatiza la necesidad de aplicar un pensamiento crítico para revisar y validar los resultados obtenidos. Asimismo, se prohíbe su empleo en contextos evaluativos restringidos y se advierte sobre protección de datos sensibles en plataformas externas.

En conjunto, estas pautas configuran un marco institucional general orientado a integrar la IAGen de manera ética, segura y pedagógicamente coherente en la escritura académica; sin embargo, dicho marco requiere aún adaptaciones específicas para los contextos de escritura en segundas lenguas, un ámbito en el que, además, la falta de estudios empíricos sobre su práctica real en el aula dificulta una adaptación e implementación verdaderamente eficaces de esta tecnología.

En respuesta a estas carencias, el estudio aquí presentado propone la integración ponderada de estas herramientas mediante una intervención basada en la asistencia de *ChatGPT* para la corrección de la competencia escrita en alemán de nivel A1. Con un diseño cuasi-experimental y un seguimiento longitudinal, en línea con la necesidad de este tipo de estudios señalada por Zhu y Wang (2025: 23), el estudio persigue contrastar la eficacia formativa de la IAGen frente a métodos más convencionales y establecer las bases de un modelo replicable de retroalimentación personalizada. El enfoque adoptado parte de la premisa de que *ChatGPT* puede ajustarse a las necesidades individuales del alumnado (Atlas, 2023: 86), analizando cómo reciben, valoran y justifican sus sugerencias, sin abordar por el momento la calificación formal de sus producciones.

A continuación, se formulan las preguntas de investigación que guían este estudio:

- P1: ¿En qué medida la retroalimentación proporcionada por *ChatGPT* mejora la calidad y coherencia de los textos en alemán de nivel A1 en comparación con otros métodos convencionales de corrección?
- P2: ¿Cómo incide el uso de *ChatGPT* en el nivel de reflexión metalingüística y en la autonomía del alumnado durante el proceso de revisión textual?
- P3: ¿Se mantienen a medio plazo los efectos formativos observados tras la intervención inicial con *ChatGPT*?

La estructura del artículo es la siguiente: en primer lugar, se expone el marco teórico y la revisión del estado de la cuestión sobre la IAGen en el desarrollo de la competencia en la enseñanza de lenguas extranjeras; en segundo lugar, se describe la metodología, incluyendo el diseño y la implementación de la intervención; posteriormente, se presentan y analizan los resultados, tanto de la comparación entre los grupos de control y experimental como de la encuesta de satisfacción y la observación cualitativa; a continuación, se discuten los hallazgos a la luz de la

[2] Ruiz-Lázaro et al. (2025: 124) identifican 31 guías elaboradas por 24 universidades hasta el 27 de septiembre de 2024, cuyo objetivo es ofrecer orientaciones, recomendaciones y buenas prácticas para apoyar a la comunidad académica en la integración adecuada de tecnologías de IAGen en la educación superior.

CC BY NC ND

literatura existente; y, finalmente, se ofrecen las conclusiones junto con las limitaciones del estudio y las propuestas para futuras líneas de investigación.

2. MARCO TEÓRICO

En el ámbito de la enseñanza de lenguas extranjeras, el procesamiento del lenguaje natural ha dado lugar – como se ha indicado a modo introductorio – a dos grandes familias de tecnologías con potencial formativo: la IAGen, sustentada en Modelos Extensos de Lenguaje como *ChatGPT*, *Gemini* o *Copilot*, capaces no solo de sugerir estructuras gramaticales, vocabulario o patrones de uso, sino también de generar textos completos adaptados a distintos niveles, contextos y géneros comunicativos; y la NMT, que emplea redes profundas para facilitar la transferencia contextual de significados entre idiomas, como sucede en *Google Translate* o *DeepL* (Cassany y Casarin, 2025: 108).

Desde una perspectiva cognitiva aplicada al aprendizaje de lenguas, ambas tecnologías pueden brindar un "apoyo coactivo" al estudiante durante la composición de textos al favorecer la generación de ideas, el esbozo de la estructura argumental, la textualización y la revisión de forma y contenido (Hoffmann et al., 2025: 14; Jiang, 2023: 4; Steinhoff y Lehnen, 2025: 7; Syarifah y Fackruddin, 2024: 560; Wang, 2024: 828). No obstante, su integración efectiva en el aula exige que el alumnado desarrolle una competencia digital crítica que le permita valorar la calidad de las sugerencias generadas, identificar posibles sesgos y evitar una dependencia excesiva que comprometa tanto su autonomía lingüística como su capacidad de reflexión analítica (Hartmann, 2021: 692; Huang y Cassany, 2025: 24; Muñoz-Basols et al., 2023: 172; Shivani, 2024: 25).

Las revisiones recientes de la literatura (Imran y Almusharraf, 2023; Zhu y Wang, 2025) coinciden en señalar que, pese al potencial formativo de la IAGen en el campo de la enseñanza de lenguas, existen importantes lagunas en el conocimiento sobre su integración efectiva en el aula, en gran medida debido al carácter aún incipiente de esta tecnología. Estas revisiones reclaman estudios con seguimiento a largo plazo que permitan evaluar el impacto real de estas herramientas en el aprendizaje de lenguas y advierten, además, de que la investigación se ha centrado mayoritariamente en enfoques cuantitativos, en detrimento de aproximaciones cualitativas que exploren los procesos de uso y las dinámicas pedagógicas que los sustentan. A ello se suma que gran parte de estos trabajos se ha desarrollado en el ámbito del inglés como L2 (Adiguzel et al., 2023; Kohnke et al., 2023; Mohebi, 2024), lo que limita la aplicabilidad de sus conclusiones a otros contextos lingüísticos y deja menos explorados escenarios educativos igualmente relevantes.

Este panorama se agrava en el ámbito del alemán como L2, donde la evidencia empírica sigue siendo particularmente limitada. Tal como subraya Heinsch (2024), la investigación presenta tres déficits principales: (1) una brecha tecnológica entre el inglés y otras lenguas europeas; (2) un desajuste entre el desarrollo de herramientas de IAGen y su articulación en diseños pedagógicos coherentes; y (3) la escasez de estudios que analicen su uso real en contextos formativos de alemán L2, pese a su presencia creciente entre el alumnado. Aunque se han realizado aportaciones relevantes en alfabetización en IAGen, creatividad y corrección colaborativa (Ciężka, 2024; Kumari, 2024; Michel et al., 2025), siguen faltando

CC BY NC ND

investigaciones experimentales y longitudinales que vinculen de forma sistemática las prácticas de aula con la progresión en las distintas competencias lingüísticas (Muñoz-Basols et al., 2025: 304).

Esta falta de evidencia tiene implicaciones directas para la competencia digital crítica descrita en el párrafo anterior. Aunque se reconoce que dicha competencia resulta fundamental para promover la autonomía lingüística y la reflexión analítica del alumnado, la literatura aún no ofrece orientaciones claras sobre cómo desarrollarla e integrarla de manera efectiva en la práctica docente, precisamente debido a las lagunas existentes en torno a la implementación real de estas tecnologías en contextos educativos diversos.

En este contexto, las intervenciones educativas del presente artículo –basadas en el uso de *ChatGPT* como apoyo a la corrección de la expresión escrita en alemán de nivel A1– se conciben para atender estas carencias identificadas en la literatura. Mediante un enfoque cuasi-experimental con seguimiento longitudinal, el estudio examina la eficacia formativa de la IAGen en comparación con métodos convencionales y propone un modelo de retroalimentación personalizada que pueda servir como referencia para futuras implementaciones pedagógicas.

3. METODOLOGÍA

3.1. Diseño

Las intervenciones analizadas en este estudio se inscriben en el proyecto de innovación educativa 'Herramientas de Inteligencia Artificial en el aula universitaria de alemán (producción escrita y traducción)' (HPB/PIE i324-18), financiado por el Servicio de Asesoramiento Educativo (HELAZ/SAE) de la UPV/EHU para el periodo 2024–2025. Este proyecto se articula en torno a tres objetivos fundamentales: (1) integrar de forma sistemática el uso de herramientas de IAGen en la producción escrita en asignaturas de lengua alemana mediante equipos docentes institucionalmente respaldados; (2) capacitar tanto al profesorado como al alumnado en un manejo consciente, ético y eficaz de dichas tecnologías; y (3) evaluar su impacto en los procesos de corrección autónoma de textos en alemán. En conjunto, el proyecto pretende maximizar el potencial formativo de la IAGen –como la agilización de la retroalimentación y el fomento de la reflexión metalingüística – al tiempo que se minimizan sus riesgos, contribuyendo así a establecer un modelo formativo ético y reproducible para la enseñanza de la expresión escrita en alemán como lengua extranjera.

La Tabla 1 evidencia cómo este PIE retoma las intervenciones propedéuticas desarrolladas en 2023–2024 en las asignaturas 'Lengua CI' y 'Lengua CII: Alemán' del Grado en Traducción e Interpretación (TeI), incorporando sus principales hallazgos para diseñar y optimizar las actividades formales programadas para 2024–2025[3]. En esta segunda fase, se aplicó el mismo protocolo tanto al grupo experimental (TeI) como a un grupo de control compuesto por estudiantes de 'Segunda Lengua I' y

[3] A partir de los resultados obtenidos en esta fase y en el marco de un nuevo proyecto, se prevé una incorporación progresiva de herramientas de IA en todas las asignaturas impartidas por el equipo docente de lengua alemana de la Facultad de Letras de la UPV/EHU a lo largo de los próximos cursos académicos.

CC BY NC ND

'Segunda Lengua II: Alemán' de los Grados en Filología, Estudios Ingleses y Estudios Vascos (Filo-EEII-EEVV), cuyos niveles, contenidos y materiales eran equiparables. Este diseño comparativo posibilitó una evaluación sistemática de la eficacia de la metodología propuesta.

Tabla 1. Asignaturas implicadas en el HPB/PIE i324-18

Curso 2023–2024	Curso 2024–2025
Primer Cuatrimestre	
Lengua CI: Alemán (25474) (TeI) GRUPO EXPERIMENTAL	Lengua CI: Alemán (25474) (TeI) GRUPO EXPERIMENTAL Segunda Lengua I: Lengua Alemana (25382) (Filo-EEII-EEVV) GRUPO DE CONTROL
Segundo Cuatrimestre	
Lengua CII: Alemán (25477) (TeI) GRUPO EXPERIMENTAL	Lengua CII: Alemán (25477) (TeI) GRUPO EXPERIMENTAL Segunda Lengua II: Lengua Alemana (25383) (Filo-EEII-EEVV) GRUPO DE CONTROL

Las intervenciones se desarrollaron a lo largo del primer y segundo cuatrimestre del curso 2024-25, con una muestra total de 28 estudiantes [n=28] – 20 de TeI y 8 de Filo-EEII-EEVV. Se empleó un muestreo de conveniencia no probabilístico que, si bien limita la generalización de los hallazgos, resultaba el más adecuado para profundizar en el análisis de los procesos de aprendizaje implicados.

Desde el punto de vista pedagógico, se adoptó un enfoque competencial centrado en la expresión escrita, en el que la revisión colaborativa, asistida por IAGen y guiada por el profesorado, se concibe como un recurso clave para reforzar dicha competencia, promover la reflexión metalingüística y fomentar el aprendizaje autónomo.

Como herramienta de IAGen, se seleccionó *ChatGPT 3.5* debido a diversas ventajas pedagógicas: su acceso gratuito garantiza la equidad; permite la interacción tanto en alemán como en las lenguas maternas del alumnado (español y euskera), lo que facilita la comprensión y la toma de decisiones; y funciona mediante *prompts* escritos, que ofrecen al estudiante un control directo y personalizable el proceso de corrección. Además, este modelo constituye la base de buena parte de las aplicaciones actuales de IAGen textual (López Gil y Moreno Mosquera, 2025: 133).

En ambas intervenciones se solicitó a los dos grupos la redacción de un texto descriptivo, a partir del cual se diseñaron y aplicaron los distintos procedimientos de intervención conforme al marco metodológico del estudio. El grupo experimental empleó un protocolo de corrección por pares asistida por *ChatGPT*: tras intercambiar sus textos, los estudiantes evaluaron críticamente las sugerencias generadas por la IAGen, seleccionaron las opciones más adecuadas y justificaron sus decisiones. Posteriormente, la docente proporcionó una retroalimentación focalizada tanto en la corrección lingüística como en los procesos metalingüísticos y en la toma de decisiones implicada en la revisión del texto[4].

Por el contrario, el grupo de control siguió un procedimiento de corrección convencional: el profesorado identificó y señaló los errores en los borradores iniciales,

[4] Para una descripción exhaustiva de las tareas empleadas en la intervención, consúltese Mallo Dorado (2025).

CC BY NC ND

tras lo cual los estudiantes revisaron el texto de manera autónoma y entregaron una segunda versión para su valoración final. La comparación entre ambos grupos permitió examinar con detalle en qué medida el apoyo de la IAGen incide en la reflexión metalingüística, la autonomía y la calidad de la producción escrita en alemán A1.

Para abordar este análisis, se adoptó una metodología de enfoque mixto, combinando técnicas cuantitativas y cualitativas. Las primeras se centraron en el registro de usos específicos de estructuras lingüísticas, mientras que las segundas permitieron observar de manera detallada los procesos de revisión y retroalimentación. Esta integración metodológica buscó aprovechar la complementariedad entre ambos enfoques, aunque predominó el carácter cualitativo debido al tamaño reducido de la muestra (Creswell y Plano, 2018).

La recogida de datos se realizó mediante cinco instrumentos: pretest, postest, documento de corrección, observación directa y encuesta de valoración. Esta combinación permitió obtener información tanto sobre el desempeño y la ganancia de aprendizaje como sobre las percepciones del alumnado respecto al uso de la IAGen. El pretest y el postest fueron evaluados mediante una rúbrica de 55 puntos estructurada en once criterios: cumplimiento de la tarea; variedad y adecuación léxica; colocación del verbo; conjugación verbal; uso de artículos y pronombres; cohesión; uso de mayúsculas; ortografía y puntuación; extensión y fluidez; y comprensión global, cada uno valorado en una escala de 0 a 5. Por su parte, la encuesta – compuesta por ítems cerrados y preguntas abiertas – permitió explorar la utilidad percibida de *ChatGPT*, el grado de autonomía alcanzado, el desarrollo del pensamiento crítico, las dificultades encontradas y el impacto general en el proceso de aprendizaje.

El análisis de datos cuantitativos se llevó a cabo mediante estadística descriptiva, con especial atención al cálculo de la ganancia porcentual y a la comparación de las puntuaciones de la rúbrica entre el pretest y el postest. Paralelamente, las respuestas abiertas, los comentarios en los documentos de corrección y las notas de observación se sometieron a un análisis de contenido, lo que proporcionó una comprensión más profunda de las dinámicas de revisión y de las formas de retroalimentación desarrolladas por el alumnado.

La intervención requirió asimismo un rol docente activo como guía y facilitador en el uso de la tecnología. En lugar de limitarse a evaluar el producto final, la docente acompañó al alumnado en el análisis crítico de las propuestas generadas por la IAGen, fomentando así la autonomía, el pensamiento crítico y la conciencia lingüística. Este enfoque coincide con las recomendaciones de la literatura reciente (Muñoz-Basols et al., 2025), que subraya la importancia de una supervisión docente sólida para garantizar un uso pedagógicamente significativo de herramientas como *ChatGPT* en entornos de aprendizaje mediados por IAGen.

3.2. Implementación

La intervención se desarrolló en tres etapas – diseño, implementación y evaluación – que permitieron articular de forma progresiva los objetivos formativos del estudio. En la fase de diseño se planificó y estructuró la intervención educativa, definiendo los objetivos de aprendizaje, el tipo de tarea a realizar, la temática y los contenidos a trabajar, así como el formato de implementación y los criterios de

CC BY NC ND

evaluación. Asimismo, se elaboró un documento guía con instrucciones claras y detalladas para el alumnado y una rúbrica específica para la valoración de los textos producidos. Todos los participantes se incorporaron al estudio de manera voluntaria tras firmar un consentimiento informado, y el estudio se llevó a cabo conforme a las directrices éticas del *Comité de Ética para las Investigaciones con Seres Humanos de la UPV/EHU* (2023).

Tal como indica la Tabla 2, la fase de implementación se organizó en tres momentos para el grupo experimental:

1. Literacidad digital, orientada a familiarizar al alumnado con el uso crítico de *ChatGPT* y con el protocolo de interacción previsto.
2. Producción y corrección textual, en la que se redactó el texto descriptivo y se aplicó el procedimiento de corrección asistida por IAGen.
3. Valoración de la actividad, destinada a recoger percepciones, dificultades y aprendizajes derivados del proceso.

Tabla 2. Implementación: descripción de fases

		GRUPO EXPERIMENTAL	GRUPO DE CONTROL
1. Literacidad digital		5 sesiones (+ tareas no presenciales)	–
2. Producción-corrección textual	Pretest	Producción escrita presencial sin asistencia externa	
	Intervención	Revisión entre pares sin IA. Revisión asistida por ChatGPT. Evaluación crítica de las sugerencias de ChatGPT. Aceptación o rechazo de las propuestas generadas. Elaboración de la versión final, con justificación de cada decisión.	Corrección individual del docente.
	Postest	Producción escrita presencial sin asistencia externa	
3. Evaluación		Encuesta y reflexión grupal	-

La fase inicial se desarrolló a comienzos del primer cuatrimestre y se centró en fortalecer la competencia digital del alumnado, promoviendo un uso consciente, autónomo y crítico de herramientas de IAGen de redacción en alemán. Durante este periodo, los estudiantes exploraron el funcionamiento de *ChatGPT* y otras plataformas similares, identificaron sus fortalezas y limitaciones, practicaron la formulación de *prompts* eficaces y evaluaron de manera crítica las respuestas generadas, al tiempo que se familiarizaron con recursos especializados en corrección textual.

La segunda etapa, orientada específicamente a la producción escrita, consistió en dos intervenciones – una en cada cuatrimestre –, con sus respectivos pretest y postest. En la primera, el alumnado redactó en modalidad presencial[5] un texto descriptivo sobre la familia; en la segunda, una composición sobre la celebración de un cumpleaños. En todas las pruebas se prohibió el uso de dispositivos electrónicos

[5] A diferencia de Nguyen et al. (2024: 860), donde se permite el uso libre de IAGen para simular contextos reales de escritura con el objetivo de investigar hasta qué punto los participantes deciden utilizarlas durante la tarea, este estudio reproduce una situación de examen, restringiendo el acceso a herramientas externas para evaluar el impacto de la intervención en un entorno controlado.

CC BY NC ND

y de materiales externos (diccionarios, manuales, apuntes o gramáticas), con el fin de garantizar la validez de los datos obtenidos.

Tras completar el pretest en cada intervención, el grupo experimental llevó a cabo la corrección del texto elaborado por un/a compañero/a en dos fases: en primer lugar, mediante una revisión independiente, y posteriormente con el apoyo de *ChatGPT*. A partir de ello realizaron un análisis comparativo crítico de ambas versiones, aceptando o rechazando las sugerencias proporcionadas por la IAGen y justificando cada decisión en una versión final anotada. Además, este grupo recibió retroalimentación individualizada tanto sobre sus propias redacciones como sobre las revisiones efectuadas a sus pares, complementadas con explicaciones en el aula sobre los errores más recurrentes. Por el contrario, el grupo de control siguió un procedimiento más convencional de retroalimentación: indirecta no codificada, aplicada de manera no enfocada (Ferris, 2011: 44). Así, el profesorado señaló los errores en los borradores iniciales y los estudiantes elaboraron una segunda versión de manera autónoma, que fue sometida posteriormente a valoración final.

En la fase final, el grupo experimental completó un cuestionario sobre su experiencia con *ChatGPT*, que incluía preguntas relativas al uso previo de la herramienta, su utilidad percibida en la producción escrita, la valoración de la actividad, su impacto en el aprendizaje y su disposición a seguir empleándola en el futuro.

La evaluación de la intervención fue continua y formativa, tomándose en consideración cuatro dimensiones: participación en las sesiones de literacidad digital; calidad de los textos producidos; ejecución de las tareas de revisión y reflexión; y cumplimentación de los cuestionarios de valoración. Esta actividad formó parte del bloque de "Tareas no presenciales" dentro del sistema de evaluación continua (60% de la nota final) y representó el 10% de la calificación de la asignatura.

El análisis de resultados se articuló en torno a cuatro aspectos: (1) la ganancia de aprendizaje, estimada mediante la comparación porcentual pretest–postest; (2) la evolución media en los once criterios de la rúbrica de expresión escrita; (3) las percepciones sobre utilidad, autonomía e impacto de *ChatGPT* recogidas en la encuesta; y (4) una interpretación integrada de todos los datos obtenidos. Conviene subrayar que las sesiones de literacidad digital constituyeron una fase formativa imprescindible, en línea con Heinsch (2024: 166), ya que prepararon al alumno para afrontar con solvencia las tareas posteriores de revisión y toma de decisiones lingüísticas.

4. RESULTADOS

En esta sección se presentan los resultados derivados del contraste entre las puntuaciones de pretest y postest, organizados en dos bloques complementarios. En primer lugar, se analiza la variación porcentual individual tras cada una de las dos intervenciones ("Meine Familie" y "Mein Geburtstag"). En segundo lugar, se examina la evolución media de los once criterios de la rúbrica de expresión escrita, lo que permite evaluar no solo el impacto global de la intervención, sino también las áreas específicas de mejora o retroceso en cada grupo. Finalmente, se sintetizan los datos obtenidos mediante la encuesta de valoración, con el fin de integrar la percepción del alumnado en la interpretación de los resultados cuantitativos.

CC BY NC ND

4.1. Variación porcentual del rendimiento

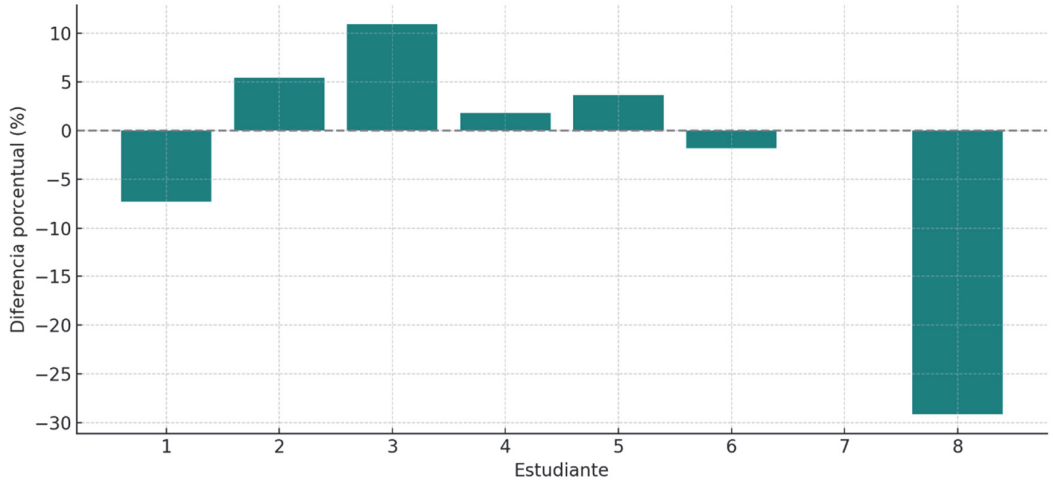
En los Gráficos 1–4 se muestra, para cada participante, la variación porcentual entre las puntuaciones de pretest y postest. En el eje X se enumeran los estudiantes (1, 2, … N) y en el eje Y se expresa el porcentaje de cambio, calculado como sigue:

$$\%\Delta = \frac{\text{Postest} - \text{Pr e test}}{\text{Pr e test}} \times 100$$

*4.1.1. Intervención I (*Meine Familie*)*

En el grupo de control (Gráfico 1) las variaciones porcentuales son, en general, moderadas y relativamente homogéneas. Solo dos estudiantes experimentan un ligero retroceso en su puntuación (alumno 1: –7%; alumno 6: –15%), mientras que uno mantiene exactamente el mismo rendimiento (0%). El resto muestra mejoras discretas: los estudiantes 2, 4 y 5 incrementan sus resultados entre un 4% y un 6%, mientras que el estudiante 3 alcanza la mejora más elevada dentro del grupo, cerca de un 11%.

Gráfico 1. IE 1 – Diferencia porcentual entre pretest y postest en grupo de control

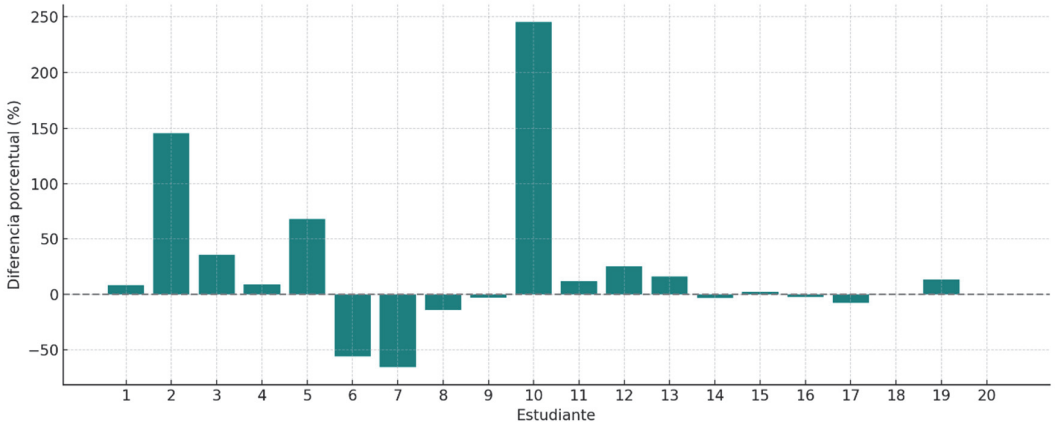


Frente a este perfil relativamente estable, el grupo experimental (Gráfico 2) presenta una variabilidad mucho mayor en los resultados. La mayoría de los estudiantes se sitúa entre variaciones del –10% al +30%, pero el conjunto muestra una dispersión amplia, con casos extremos de mejora muy significativa, como el alumno 10 (+245%) y el alumno 2 (+145%). También se observan progresos destacables en otros participantes, como los alumnos 5 (+65%) y 11 (+30%).

En contraste, varios estudiantes experimentan descensos pronunciados (alumnos 7 y 8, entre –30% y –50%), lo que evidencia un impacto desigual de la intervención y sugiere la necesidad de considerar factores individuales y procesuales en la interpretación de los resultados.

CC BY NC ND

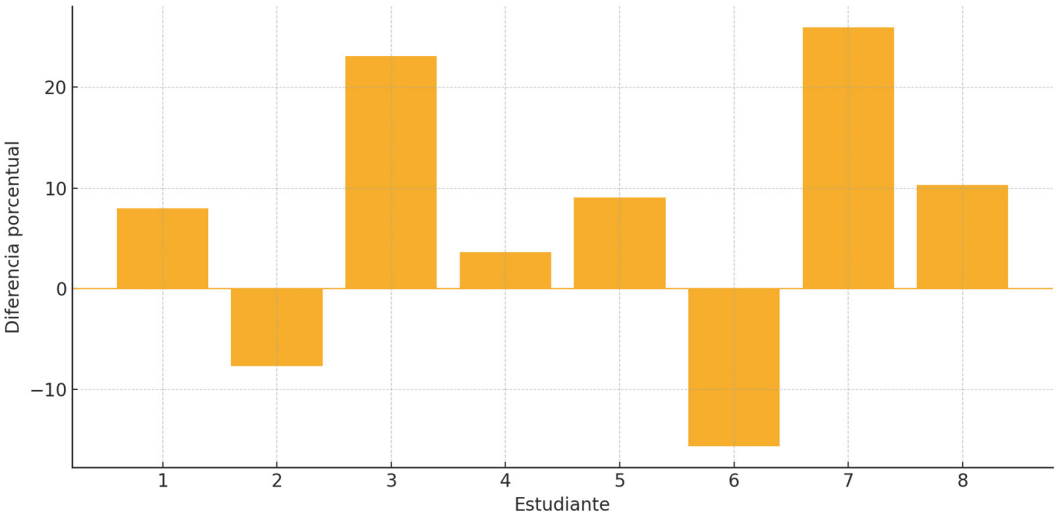
Gráfico 2. IE 1 – Diferencia porcentual entre pretest y postest en GE



*4.1.2. Intervención II (***Mein Geburtstag*)**

En el grupo de control (Gráfico 3) las variaciones porcentuales tras la segunda intervención mantienen un perfil moderado y homogéneo. Únicamente dos estudiantes presentan descensos – alumno 2 (–8%) y alumno 6 (–15%) –, mientras que los restantes muestran mejoras discretas. En concreto, los alumnos 1, 4, 5 y 8 registran incrementos comprendidos entre +3% y +10%. Por su parte, los estudiantes 3 y 7 alcanzan las ganancias más elevadas del grupo, con aproximadamente +23% y +25%, respectivamente.

Gráfico 3. IE 2 – Diferencia porcentual entre pretest y postest en grupo de control

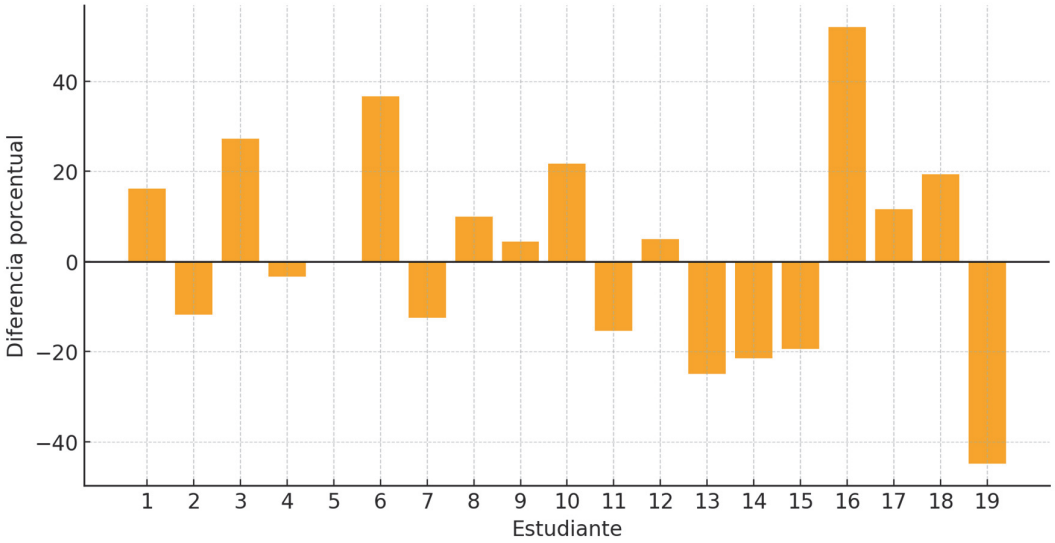


En el grupo experimental (Gráfico 4) se reproduce la elevada dispersión observada en la primera intervención. La mayoría de los estudiantes presenta variaciones comprendidas entre –20% y +20%; no obstante, vuelven a registrarse casos extremos. Así, los alumnos 16 (+48%), 6 (+37%) y 3 (+28%) muestran incrementos

CC BY NC ND

especialmente destacados, mientras que los descensos más acusados corresponden a los alumnos 19 (–45%), 13 (–25%) y 15 (–20%).

Gráfico 4. IE 2 – Diferencia porcentual entre pretest y postest en grupo experimental



4.1.3. Conclusiones de la variación porcentual

La comparación entre ambas intervenciones pone de manifiesto un patrón consistente: el grupo de control presenta variaciones moderadas y relativamente homogéneas –con mejoras generalmente situadas entre +3% y +11% y descensos puntuales de hasta –15%–, mientras que el grupo experimental exhibe una dispersión claramente mayor, con incrementos muy elevados (+28% a +245% en la Intervención I; hasta +53% en la Intervención II) y retrocesos igualmente acusados (entre –20% y –50%).

En las dos intervenciones se observa, además, que tanto el grupo de control como el grupo experimental muestran mejoras en la mayoría del alumnado, junto con algunos casos de descenso. En conjunto, ambos enfoques parecen favorecer un aumento del rendimiento medio; no obstante, la variabilidad individual es sustancialmente superior en el grupo experimental, lo que sugiere que el impacto de la intervención asistida por IAGen resulta más heterogéneo y depende en mayor medida de las características y estrategias de cada estudiante.

4.2. Resultados por criterio de rúbrica

En los Gráficos 5–8 se presentan los diagramas de barras comparativas que ilustran la puntuación media obtenida en cada uno de los once criterios de la rúbrica de expresión escrita; las barras amarillas representan los resultados del pretest y las naranjas los del postest. En el eje X se numeran los criterios – (1) Cumplimiento de la tarea; (2) Variedad léxica; (3) Adecuación léxica; (4) Uso del verbo en 2.ª posición; (5) Conjugación verbal; (6) Artículos y pronombres; (7) Cohesión y conectores; (8) Uso de mayúsculas; (9) Ortografía y puntuación; (10) Extensión y fluidez; (11)

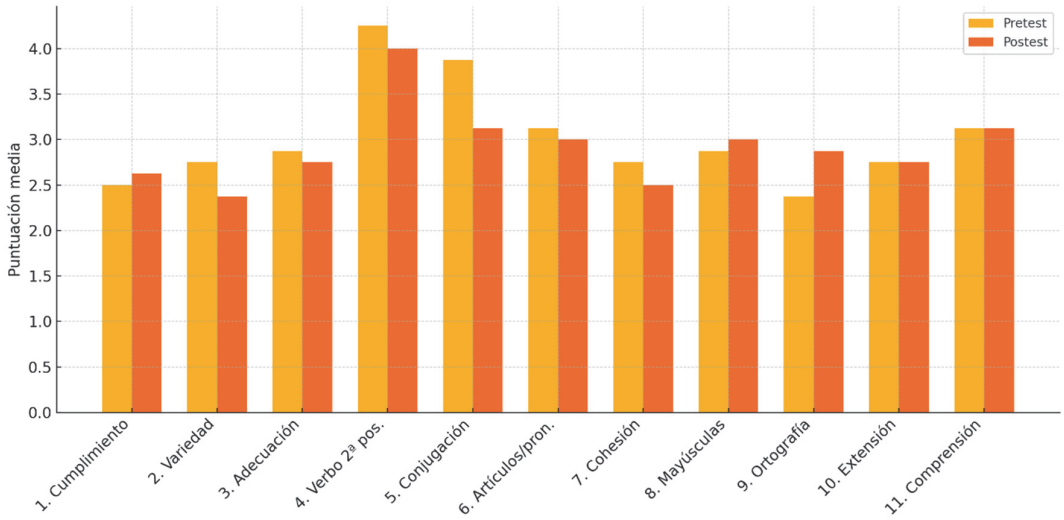
CC BY NC ND

Comprensión global –, mientras que el eje Y muestra la puntuación media obtenida en una escala de 0 a 5.

*4.2.1. Intervención I (*Meine Familie*)*

En el grupo de control (Gráfico 5), las diferencias absolutas entre pretest y postest varían entre 0,00 y 0,75 puntos. El cambio más destacado se registra en C5, que desciende 0,75 puntos. Por el contrario, C9 muestra la mejora más relevante, con un aumento de 0,50. Otros criterios presentan ajustes menores: C2 desciende 0,38.; C4 y C7 disminuyen 0,25; mientras que C1, C3, C6 y C8 experimentan variaciones marginales, inferiores a ±0,10. Finalmente, C10 y C11 permanecen sin cambios.

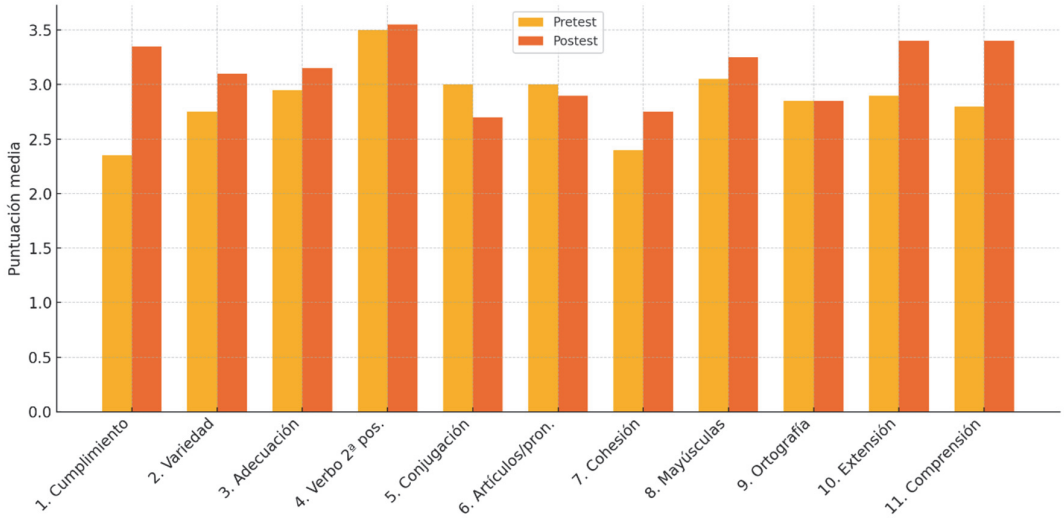
Gráfico 5. IE I – Comparación de rúbricas por criterio entre pretest y postest en grupo de control



En el grupo experimental (Gráfico 6), las diferencias absolutas entre pretest y postest muestran una variabilidad más amplia, con oscilaciones entre 0,00 y 1,00 puntos. El avance más destacado corresponde al criterio C1 (+1,00), seguido de mejoras sustanciales en C11 (+0,60) y C10 (+0,50). También se observan incrementos en C2 y C7 (+0,35 cada uno), así como en C8 y C3 (+0,20). En cambio, C4 apenas varía (+0,05), mientras que C5 (–0,30) y C6 (–0,10) experimentan descensos. El criterio C9 no registra cambios.

CC BY NC ND

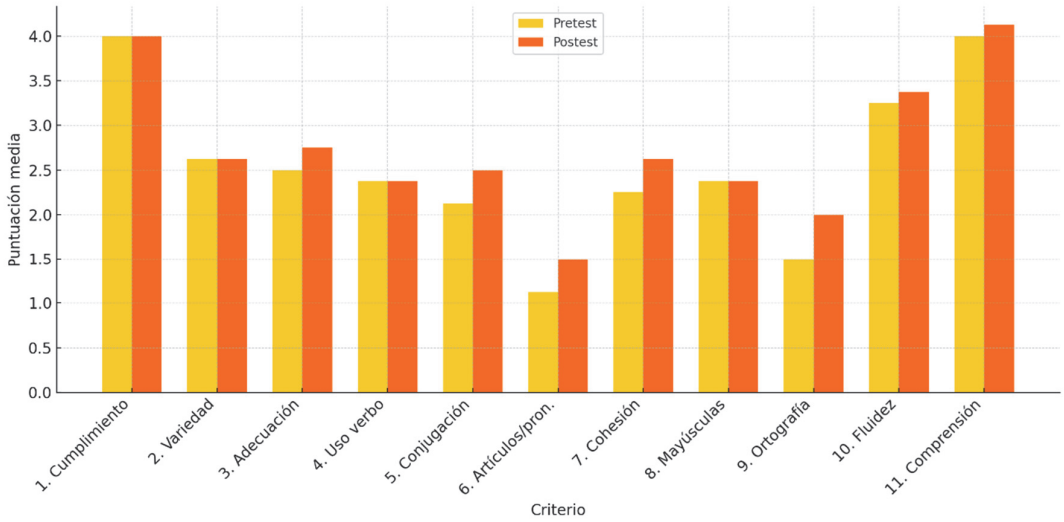
Gráfico 6. IE I – Comparación de rúbricas por criterio entre pretest y postest en grupo experimental



*4.2.2. Intervención II (*Mein Geburtstag*)*

En el grupo de control, las variaciones absolutas entre pretest y postest son discretas y uniformes: los mayores avances se dan en C5, C6 y C7 (+0,38), seguidos de C3 (+0,25) y de C10 y C11 (+0,12), mientras que C1, C2, C4, C8 y C9 no registran cambios.

Gráfico 7. IE II – Comparación de rúbricas por criterio entre pretest y postest en grupo de control

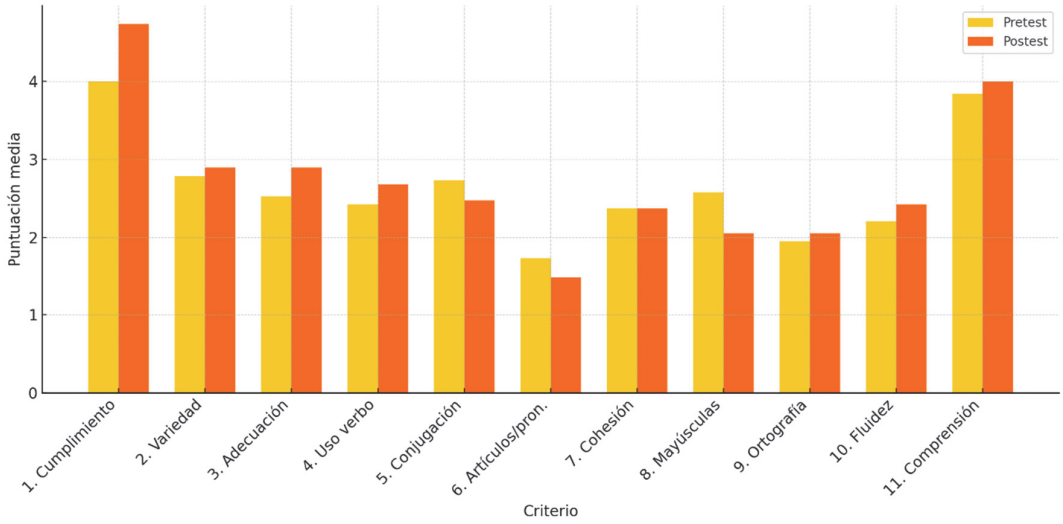


En el grupo experimental (Gráfico 8), las diferencias entre pretest y postest oscilan entre –0,53 y +0,50 puntos. Las mejoras más acusadas se observan en C1 (+0,50), C3 (+0,40) y C4 (+0,30), mientras que los descensos más marcados

CC BY NC ND

corresponden a C8 (–0,53), C6 (–0,25) y C5 (–0,20). En los criterios restantes (C2, C7, C9, C10 y C11), los ajustes son reducidos, con variaciones positivas que se sitúan entre +0,10 y +0,25.

Gráfico 8. IE II – Comparación de rúbricas por criterio entre pretest y postest en grupo experimental



4.2.3. Conclusiones de la variación absoluta de las puntuaciones por criterio

En conjunto, los resultados indican que la intervención aplicada en el grupo experimental tiende a favorecer el desarrollo de determinadas competencias discursivas, especialmente aquellas vinculadas con la organización textual y la coherencia global (C1, C3 y C4). No obstante, este avance convive con descensos o progresos limitados en criterios de naturaleza más formal o mecánica (C5, C6 y C8), lo que sugiere que la integración de la IAGen no garantiza por sí sola una mejora equilibrada de todos los componentes de la expresión escrita.

La ausencia de incrementos consistentes en aspectos como la ortografía, la conjugación verbal o el uso de mayúsculas pone de relieve la necesidad de complementar el enfoque con apoyos específicos dirigidos a estas áreas, especialmente en niveles iniciales. Al contrastar ambos grupos, puede apreciarse que la metodología experimental genera mejoras más visibles en la dimensión discursiva del texto, mientras que el grupo de control mantiene un perfil de evolución más homogéneo, aunque menos pronunciado. En conjunto, los datos sugieren que la intervención con IAGen ofrece un potencial formativo relevante, pero requiere un diseño más equilibrado para atender simultáneamente a los distintos niveles de la competencia escrita.

4.3. Encuesta de valoración

Los resultados de la encuesta revelan una percepción mayoritariamente positiva por parte del alumnado respecto al uso de herramientas de IAGen durante la intervención. El 80% de los participantes indicó haber utilizado previamente

aplicaciones como *ChatGPT*, *DeepL* o *DeepLWrite*, lo que explica el bajo nivel de reticencia inicial, situado mayoritariamente entre los niveles 2 y 3 de una escala de 1 a 5.

Entre las ventajas más mencionadas destacan la rapidez, la claridad de las respuestas y su utilidad como apoyo en la revisión de textos, elementos que refuerzan el potencial de estas herramientas en la producción escrita. No obstante, el alumnado también identificó limitaciones relevantes, como la fiabilidad desigual de la información, la invención de datos y el riesgo de fomentar una dependencia excesiva que pueda disminuir el pensamiento crítico.

Respecto a las preferencias, el grupo mostró una inclinación clara hacia *ChatGPT*, valorado por su versatilidad y aplicabilidad a tareas diversas. Aunque la percepción del impacto directo en el aprendizaje fue moderada (niveles 2–3), la intención de uso futuro resultó elevada, ya que la mayoría expresó que probablemente seguiría recurriendo a estas herramientas ("sí" o "tal vez").

En términos generales, la experiencia fue considerada positiva, útil e innovadora, lo que respalda la pertinencia de integrar la IAGen de manera crítica y reflexiva en contextos de enseñanza de lenguas extranjeras.

4.4. Valoración global

La comparación porcentual de ambos bloques revela un patrón consistente: el grupo de control registra variaciones moderadas y relativamente homogéneas, entre +3% y +11% (con descensos puntuales de hasta –15%), mientras que el grupo experimental exhibe una dispersión más amplia, con incrementos de +28% a +245% en IE I y de hasta +53% en IE II, junto a descensos de –20% a –50%. A pesar de que ambos enfoques elevan la puntuación media en la mayoría de los criterios, la variabilidad individual en el grupo experimental es notablemente mayor que en el grupo de control.

En términos absolutos – es decir, considerando la diferencia directa en puntos dentro de la rúbrica (0–5) sin convertirlos en porcentajes–, el grupo de control presenta fluctuaciones muy reducidas, con un rango máximo de ±0,75 puntos, concentradas principalmente en los criterios de conjugación, uso de artículos y cohesión. Por su parte, el grupo experimental evidencia una variabilidad más marcada, con diferencias que oscilan desde descensos de –0,53 hasta mejoras de +1,00. Esto indica que, aunque ambos grupos mejoran o empeoran en distintos criterios, el de control mantiene un progreso muy estable en todas las áreas, mientras que el experimental registra cambios más intensos en algunos criterios específicos.

Las tendencias cualitativas complementan esta lectura cuantitativa. En el grupo de control, las mejoras parecen responder más a la memorización de correcciones previas que a una comprensión transferible de las estructuras: el alumnado tiende a reproducir fragmentos corregidos del pretest sin reestructurar el texto, lo cual denota un aprendizaje limitado y puntual. Por el contrario, en el grupo experimental el uso de *ChatGPT* fomentó una escritura más libre y expresiva, con ampliación del texto, espontaneidad y narrativa desarrollada al centrar la atención en el contenido antes que en lo formal.

En conjunto, los datos indican que la intervención didáctica potencia de manera sostenida las competencias discursivas globales (estructura, adecuación y coherencia) en ambos grupos. No obstante, el grupo experimental presenta descensos

CC BY NC ND

en precisión ortográfica y morfosintáctica, lo que refuerza la necesidad de complementar las actividades asistidas por IAGen con ejercicios específicamente orientados a fortalecer los aspectos normativos y mecánicos de la escritura, a fin de garantizar un desarrollo equilibrado de la competencia escrita.

5. DISCUSIÓN

En respuesta a las preguntas de investigación planteadas al comienzo, los resultados permiten afirmar lo siguiente:

- P1: La retroalimentación proporcionada por *ChatGPT* mejora la calidad y coherencia de los textos en alemán A1 en mayor medida que la corrección convencional, ya que el grupo experimental muestra avances más destacados en los criterios discursivos principales.
- P2: Tanto las observaciones cualitativas como la encuesta evidencian un incremento notable de reflexión metalingüística en el grupo experimental. El alumnado empleó *ChatGPT* para planificar y estructurar sus textos de manera más autónoma y consciente, focalizando su atención en el desarrollo del contenido y en la articulación de ideas, en lugar de limitarse a corregir errores aislados.
- P3: El seguimiento a medio plazo confirma que los progresos vinculados a competencias discursivas tienden a mantenerse. No obstante, las habilidades normativo-mecánicas muestran un retroceso cuando no se refuerzan de forma específica, lo que pone de manifiesto la necesidad de incorporar actividades focalizadas en estos aspectos para consolidar la competencia escrita de manera equilibrada.

Nuestros hallazgos confirman el "apoyo coactivo" de la IAGen en la redacción (Jiang, 2023), en la medida en que los estudiantes del grupo experimental lograron planificar y estructurar sus textos con mayor coherencia. Este resultado ilustra la ampliación del modelo de interacción de Moore (1989) propuesta por Muñoz-Basols et al. (2025), según la cual la IAGen adopta un rol formativo previo a la intervención docente, facilitando procesos de generación, organización y revisión textual.

Paralelamente, el estancamiento observado en el grupo de control reproduce el patrón de corrección puntual descrito por Navío-Inglés y Guzmán Mora (2025), una práctica que no garantiza la interiorización duradera de las estructuras lingüísticas. La tendencia del alumnado a reproducir fragmentos previamente corregidos, sin reformular o ampliar el texto, coincide con las advertencias de Zhu y Wang (2025) y de Shivani (2024) acerca del riesgo de una memorización mecánica sin reflexión metalingüística, característica de enfoques centrados exclusivamente en la corrección explícita del error.

Por su parte, los descensos en ortografía y conjugación registrados en el grupo experimental respaldan las advertencias de Hartmann (2021) y de Huang y Cassany (2025) sobre la fiabilidad limitada de la IAGen y el riesgo de generar dependencia si no se acompaña de una alfabetización digital crítica. Estos resultados subrayan la conveniencia de combinar la retroalimentación automatizada con tutoría metalingüística específica, de modo que la potencia generativa de la IAGen no eclipse el desarrollo de la precisión normativa.

CC BY NC ND

Finalmente, la valoración positiva del alumnado, pese a los sesgos detectados, enlaza con las reflexiones de Farazouli et al. (2024: 369) sobre los retos éticos y de autoría en la era de la IAGen, y refuerza la necesidad de marcos normativos claros que regulen su uso en entornos educativos (Heinsch, 2024). Estos elementos, considerados conjuntamente, evidencian el potencial pedagógico de la IAGen, pero también la necesidad de una integración crítica, guiada y éticamente informada.

6. CONCLUSIONES

Los resultados de este estudio muestran la eficacia de *ChatGPT* como corrector-guía para mejorar la competencia discursiva y pragmática en la escritura de alemán A1, fomentar la autonomía del alumnado y promover la reflexión metalingüística. A diferencia de los enfoques meramente memorísticos, el uso sistemático de IAGen permitió a los estudiantes centrar sus esfuerzos en la organización del contenido y la coherencia del texto, generando producciones más ricas y elaboradas.

No obstante, los hallazgos también subrayan la necesidad de una mediación pedagógica sólida que evite la mecanización del proceso de corrección. Resulta esencial complementar el uso de IAGen con actividades que fomenten el pensamiento crítico – como la justificación de las sugerencias aceptadas o rechazadas – y con ejercicios específicos destinados a consolidar la competencia gramatical y ortográfica, especialmente en aquellos aspectos donde se detectaron retrocesos.

Entre las limitaciones del estudio destaca el reducido tamaño de la muestra, lo que exige cautela a la hora de extrapolar los resultados. Con todo, los datos obtenidos aportan evidencias relevantes en un campo aún emergente y abren vías para futuras investigaciones. Entre las líneas de trabajo propuestas se incluyen la ampliación de la muestra, la diversificación de las tareas de escritura, la incorporación de niveles lingüísticos superiores y la comparación de distintos modos de interacción con herramientas de IAGen. Asimismo, se prevé mantener el seguimiento longitudinal en los próximos cursos y extender la intervención a otras asignaturas, intensificando la colaboración docente para garantizar una transferencia efectiva de esta innovación pedagógica.

REFERENCIAS

Adiguzel, T., Kaya, M. H. y Cansu, F. K. 2023. Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology* 15(3): ep429.

Atlas, S. 2023. *ChatGPT for higher education and professional development: A guide to conversational AI*. College of Business Faculty Publications. https://digitalcommons.uri.edu/cba_facpubs/548 (acceso: 11/06/2024)

Cassany, D. y Casarin, M. 2025. Éxito, fracaso y retos letrados con la IA. *Traslaciones. Revista Latinoamericana de Lectura y Escritura* 12(23): 105–127.

Ciężka, A. 2024. Generative KI-Tools: Die Zukunft des kreativen Lernens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 29(1): 375–405.

CC BY NC ND

Comité de Ética para las Investigaciones con Seres Humanos – Universidad del País Vasco/Euskal Herriko Unibertsitatea. 2023. *Procedimiento normalizado de trabajo (PNT).* https://www.ehu.eus/es/web/ceid/comite-de-etica-para-las-investigaciones-con-seres-humanos-ceish/presentacion-composicion-pnt (acceso: 11/06/2024)

Creswell, J. y Plano, V. L. 2018. *Designing and conducting mixed methods research*. Sage.

Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K. y McGrath, C. 2024. Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education* 49(3): 363–375.

Ferris, D. R. 2011. *Treatment of error in second language student writing*. University of Michigan Press.

Hartmann, D. 2021. Künstliche Intelligenz im DaF-Unterricht? Disruptive Technologien als Herausforderung und Chance. *Info DaF* 48(6): 683–696.

Heinsch, B. 2024. Lenguas extranjeras e inteligencia artificial. Adquisición de competencias comunicativas en alemán como lengua extranjera en la educación superior: Un caso práctico. En A. Barrientos-Báez, D. Caldevilla Domínguez y J. Sierra Sánchez eds. *Inteligencia artificial ¿amiga o enemiga?* Peter Lang, 157–170.

Hoffmann, I., Engelmayr-Hofmann, K. y Gür-Şeker, D. 2025. *KI für AI-nsteiger – Künstliche Intelligenz im Kontext DaF/DaZ. Ein Einführungspapier*. Versión 2.0, febrero 2025.

Huang, Sh. y Cassany, D. 2025. Spanish language learning in the AI era: AI as a scaffolding tool. *Journal of China Computed-Assisted Language Learning*. https://doi.org/10.1515/jccall-2024-0026

Imran, M. y Almusharraf, N. 2023. Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology* 15(4): e464.

Jiang, Z., Xu, Z., Pan, Z., He, J. y Xie, K. 2023. Exploring the role of artificial intelligence in facilitating assessment of writing performance in second language learning. *Languages* 8(4): 1–16.

Jolley, J. R. y Maimone, L. 2022. Thirty years of machine translation in language teaching and learning: A review of the literature. *L2 Journal* 14(1): 26–44.

Kohnke, L., Moorhouse, B. L. y Zou, D. 2023. ChatGPT for language teaching and learning. *RELC Journal* 54(2): 537–550.

Kumari, P. V. 2024. The impact of AI models on students' writing competence and writing motivation in the context of learning German as a foreign language in India. *International Journal For Multidisciplinary Research* 6(2): 1–11.

López Gil, K. y Moreno Mosquera, E. 2025. Retroalimentación formativa en la escritura de tesis en posgrado: Comparación entre ChatGPT y revisores pares en un círculo de escritura. *Revista Virtual Universidad Católica del Norte* 74: 123–160.

Mallo Dorado, A. 2025. ChatGPT como herramienta de corrección y reflexión para la producción textual en lengua alemana en el ámbito universitario. En M. L. Carrió Pastor y C. Ureña Tormo eds. *Análisis del discurso y enseñanza de lenguas en la era de la inteligencia artificial*. Octaedro, 207–238.

Michel, M., Bazhutkina, I., Abel, N. y Strobl, C. 2025. Collaborative writing based on generative AI models: Revision and deliberation processes in German as a foreign language. *Journal of Second Language Writing* 67: art. 101185.

Mohebi, L. 2024. Empowering learners with ChatGPT: Insights from a systematic literature exploration. *Discov Educ* 3, 36. https://doi.org/10.1007/s44217-024-00120-y (acceso: 8/04/2025)

Muñoz-Basols, J., Fuertes Gutiérrez, M. y Cerezo, L. 2025. *Technology-mediated language teaching: From social justice to artificial intelligence*. Multilingual Matters.

Muñoz-Basols, J., Neville, C., Lafford, B.A. y Godev, C. 2023. Potentialities of applied translation for language learning in the era of artificial intelligence. *Hispania* 106(2): 171–194.

CC BY NC ND

Navío-Inglés, M. y Guzmán Mora, J. 2025. Docentes y ChatGPT. Una comparativa en la evaluación de la expresión escrita de los maestros en formación. En M. L. Carrió Pastor y C. Ureña Tormo eds. *Análisis del discurso y enseñanza de lenguas en la era de la inteligencia artificial*. Octaedro, 239–256.

Nguyen, A., Hong, Y., Dang, B. y Huang, X. 2024. Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education* 49(5): 847–864.

Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial. *Diario Oficial de la Unión Europea.* https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX:32024R1689

Román-Mendoza, E. 2023. El arte de formular preguntas para comprender las respuestas: ChatGPT como agente conversacional en el aprendizaje de español como segunda lengua. *marcoELE. Revista de Didáctica Español Lengua Extranjera* 36: 1–18.

Ruiz-Lázaro, J., Redondo-Duarte, S., Jiménez-García, E., Martínez-Requejo, S. y Galán-Íñigo, A. 2025. Análisis de las guías de uso de inteligencia artificial en educación superior: comparación entre las universidades españolas. *Bordón. Revista de Pedagogía* 77(1): 121–153.

Shivani, S. 2024. Teaching and learning German as a foreign language in the age of Artificial Intelligence: Possibilities and challenges of integrating AI technologies. *International Journal of Language, Literature and Culture* 4(6): 20–31.

Steinhoff, T. y Lehnen, K. 2025. Schreiben mit Künstlicher Intelligenz: Das GPT-Modell (Ghost, Partner, Tutor). *Leseräume. Zeitschrift für Literalität in Schule und Forschung* 12(11): 1–14.

Syarifah, E. F. y Fakhruddin, A. 2024. Exploring students' experience using AI to assist their writing. *Journal of English Language Learning* 8(1): 558–564.

UPV/EHU. 2025. Principios y recomendaciones para el uso de la IA generativa en la UPV/EHU. *Universidad del País Vasco*. https://www.ehu.eus/documents/d/bilboko-hezkuntza-fakultatea/principio-y-recomendaciones-uso-ia_ehu (acceso: 17/09/2025)

Wang, Ch. 2024. Exploring students' Generative AI-Assisted writing processes: Perceptions and experiences from native and nonnative. *Tech Know Learn* 30: 1825–1846.

Wang, H., Wu, H., He, Z., Huang, L. y Church, K. W. 2022. Progress in machine translation. *Engineering* 18: 143–153.

Zhu, M. y Wang, C. 2025. A systematic review of research on AI in language education: Current status and future implications. *Language Learning & Technology* 29(1): 1–29.

CC BY NC ND