

ROMÁN MENDOZA, ESPERANZA

GEORGE MASON UNIVERSITY, VIRGINIA, ESTADOS UNIDOS

EL ARTE DE FORMULAR PREGUNTAS PARA COMPRENDER LAS RESPUESTAS: CHATGPT COMO AGENTE CONVERSACIONAL EN EL APRENDIZAJE DE ESPAÑOL COMO SEGUNDA LENGUA

BIODATA

Esperanza Román Mendoza (eromanme@gmu.edu) es Catedrática de Español y Lingüística Española en George Mason University (Virginia, EE. UU.). Sus áreas de especialización son la educación a distancia, el *e-learning* y el estudio de las lenguas de herencia desde la perspectiva de la pedagogía crítica. Sus numerosos trabajos sobre tecnología educativa se enfocan en comprender las diferentes formas en las que el alumnado se percibe a sí mismo como aprendiente, y cómo el profesorado puede contribuir a que el estudiantado utilice la tecnología para afrontar retos de aprendizaje, superar la autocrítica, aprender de manera autónoma y convertirse en parte activa de la ciudadanía crítica. Más información en <https://mcl.gmu.edu/people/eromanme>.

RESUMEN

Este artículo examina el potencial de ChatGPT para el aprendizaje de español como segunda lengua en los niveles A1-B1 en cuanto a su capacidad tanto de descifrar el *input* del alumnado de estos niveles como de generar respuestas que puedan comprender los aprendientes. Para ello, y tras un breve repaso de las potencialidades y desafíos de la integración de las herramientas de generación de textos basadas en la inteligencia artificial, se describe el análisis realizado con dos corpus de interacciones con ChatGPT - compilados a partir de *prompts* del alumnado y de ejercicios propios de los cursos en estos niveles- y se presenta una serie de conclusiones preliminares a la implementación de ChatGPT en clase, resultantes del análisis de dichos corpus, y encaminadas a fomentar el autoaprendizaje mediante ChatGPT como agente conversacional.

PALABRAS CLAVE: EL/E, L2, LH, inteligencia artificial, ChatGPT

THE ART OF ASKING QUESTIONS TO UNDERSTAND THE ANSWERS: CHATGPT AS A CONVERSATIONAL AGENT IN LEARNING SPANISH AS A SECOND LANGUAGE

This article examines the potential of ChatGPT for learning Spanish as a second language at the A1-B1 levels in terms of its ability to both decipher input from learners at these levels and generate responses that learners can understand. To do this, and after a brief overview of the affordances and challenges of integrating AI-based text generative tools, it describes the analysis carried out with two corpora of interactions with ChatGPT —created from prompts from students and exercises from courses at these levels. The article also presents a series of preliminary conclusions regarding the implementation of ChatGPT in the classroom, resulting from the analysis of the two corpora, aimed at promoting self-learning through ChatGPT as a conversational agent.

KEYWORDS: SFL, L2, HL, artificial intelligence, ChatGPT

1. INTRODUCCIÓN

La utilización de los chatbots como herramienta de apoyo para la enseñanza / aprendizaje de segundas lenguas se lleva estudiando en las dos últimas décadas desde perspectivas muy diferentes (Belda-Medina y Calvo-Ferrer, 2022; Coniam, 2004; Coniam, 2008; Godwin-Jones, 2023; Huang, Hew y Fryer, 2022; Jia, 2003; Kim *et al.*, 2022; Vázquez-Cano, Mengual-Andrés y López-Meneses, 2021).

Sin embargo, la vertiginosa adopción de ChatGPT (*Chat Generative Pre-Trained Transformer*) por parte del público general tras su lanzamiento en noviembre de 2022 y la inmediata repuesta de las grandes compañías tecnológicas para posicionarse en la carrera de la generación de texto a través de inteligencia artificial han incrementado el interés por introducir esta tecnología en el sistema educativo, con reacciones que abarcan desde el entusiasmo incondicional hasta el rechazo más absoluto (por ejemplo, Roberts, Watumull y Chomsky, 2023). ChatGPT es un sistema avanzado de chatbot basado en un gran modelo de lenguaje (*Large Language Model* o LLM) capaz de generar texto en múltiples idiomas. A diferencia de los chatbots no basados en LLM, la producción textual de ChatGPT presenta un alto grado de corrección gramatical y léxica. Además, ChatGPT es capaz de contestar preguntas (*prompts*) de muy diversa índole y de organizar sus respuestas en diferentes formatos. Su potencial para aprender una segunda lengua y practicarla es innegable. Pero ¿hasta qué punto y en qué condiciones?

¹ Según Reber (2010), la capacidad del cerebro humano es equivalente a unos 2,5 millones de gigabytes.

1.1. INTELIGENCIA ARTIFICIAL. CONCEPTOS BÁSICOS

El término inteligencia artificial es acuñado por John McCarthy en 1956, aunque la comunidad científica ya se había planteado con anterioridad si las máquinas podrían llegar a pensar (Turing, 1950). Para responder esta cuestión, resulta fundamental establecer primero qué es realmente la inteligencia humana y si esta se basa exclusivamente en la capacidad de almacenamiento y procesamiento de datos que tiene el cerebro humano¹. No hay duda de que la inteligencia humana es multidimensional y aunque, en algunos contextos y ante tareas concretas, los sistemas de inteligencia artificial muestren comportamientos similares al de los humanos, el hecho es que no se ha podido desarrollar una máquina que emule completamente la inteligencia humana. Por lo tanto, para describir el grado de inteligencia de los sistemas computacionales creados hasta la fecha, resulta más apropiado utilizar el término “parcialmente inteligentes” (Feigenbaum, 2003).

De hecho, los primeros grandes avances en el campo de la inteligencia artificial se produjeron dentro del área de los llamados sistemas expertos, sistemas basados en conocimiento o sistemas basados en reglas. Estos sistemas se apoyan en la idea de que el conocimiento humano se puede codificar siguiendo reglas lógicas (Eliot, 2022). Aunque codificar todo el conocimiento humano en forma de reglas resulta inviable, sí ha sido posible hacerlo en campos específicos o para problemas concretos, en cuyo caso el resultado obtenido con la máquina puede incluso ser mejor que el realizado por un humano (Feigenbaum, 2003).

Sin embargo, el coste y la dificultad del desarrollo de sistemas expertos supuso al mismo tiempo un impedimento para la creación de aplicaciones que tuvieran un uso más general. En el caso concreto del procesamiento del lenguaje natural, los sistemas expertos avanzaron significativamente en la decodificación de las reglas gramaticales, pero en el plano de la semántica (Feigenbaum, 2003) y de la pragmática no consiguieron los resultados necesarios para producir, por ejemplo, traducciones de alta calidad. Ante estas circunstancias, se hizo necesario buscar otras formas que permitieran avanzar en el campo. Comienzan así nuevas líneas de investigación en inteligencia artificial basadas en redes neuronales artificiales (sistemas computacionales que imitan el funcionamiento del cerebro humano) y en el aprendizaje profundo (*deep learning*), es decir, algoritmos inspirados en la capacidad que tiene el cerebro humano para aprender de la experiencia (Ruiz, 2022; Russell y Norvig, 2020). Estos dos avances son fundamentales para la aparición de lo que se conoce como grandes modelos de lenguaje (a partir de ahora LLM) ya que “el aprendizaje profundo se basa en algoritmos que pueden manejar enormes conjuntos de datos que no estén estructurados ni etiquetados, [...] lo que lo convierte en un sistema ideal para el procesamiento del lenguaje natural” (Ruiz, 2022, par. 4). El tipo de arquitectura de redes neuronales que mayor impacto ha tenido hasta la fecha en el procesamiento del lenguaje natural son los llamados *transformers*, desarrollados por Google Brain en 2017 y utilizados en

² Los parámetros de un LLM basado en *transformers* son valores que el sistema aprende con los datos que se utilizan en su entrenamiento y que pueden actualizarse de forma autónoma con dicha información de forma autosupervisada. Se estima que cuanto mayor sea el número de parámetros, mayor será el grado de exactitud en las predicciones. Por ejemplo, GPT-3 se lanzó en 2020 con 175.000 millones de parámetros y GPT-4 tiene una cantidad seis veces mayor, es decir, cerca de un billón de parámetros.

³ No siempre es posible averiguar el volumen exacto de textos, ni su naturaleza, empleados para entrenar los diferentes LLM existentes. El informe publicado por

todas las versiones de GPT, el LLM en el que se basa ChatGPT. Los *transformers* evitan que la red neuronal tenga que ser entrenada con datos previamente etiquetados ya que el proceso de codificación y decodificación se realiza utilizando modelos matemáticos probabilísticos y millones de parámetros².

ChatGPT, por lo tanto, es un sistema que utiliza la inteligencia artificial para generar texto basándose en un LLM que ha sido entrenado con millones de *tokens* extraídos de ingentes cantidades de texto³, y que puede predecir qué palabra(s) u oraciones van a aparecer después de otra palabra concreta. Como toda herramienta basada en un LLM, es ideal para buscar, extraer, agrupar, clasificar, generar, resumir y reescribir información textual (Amer, 2022). Obviamente, ChatGPT no es capaz de pensar ni pasaría el test de Turing ni el de Feigenbaum, y, como muchos han señalado, se trata de “una boca con bozal y sin cerebro” (Goodlad y Baker, 2023), por lo que en lugar de inteligencia artificial parece más apropiado hablar de “astucia artificial”. Sin embargo, dada la vertiginosa carrera en la que las grandes empresas tecnológicas han entrado para posicionarse en el campo de las aplicaciones basadas en la inteligencia artificial generativa para el gran público⁴ y las grandes posibilidades que los sistemas de generación de texto como ChatGPT abren a la enseñanza de lenguas y al estudio de la lengua, resulta imprescindible que lingüistas y especialistas en la enseñanza de

Thompson (2020) incluye tanto los datos que han hecho públicos los desarrolladores de LLM como los que ha podido calcular él de forma indirecta. Las actualizaciones de este informe se encuentran en <https://lifearchitect.ai/whats-in-my-ai/>.

⁴ Por ejemplo, Microsoft con GPT de OpenAI, Google con BERT y LaMDa (el LLM que reemplazó a BERT en su *chatbot* Bard), Amazon con Teacher Model 2.0B y Meta con LLaMA. Apple también está mejorando su LLM para introducirlo en su asistente Siri y la empresa de Elon Musk, X, está trabajando en el desarrollo de X AI.

lenguas se familiaricen con las potencialidades de este chatbot avanzado y estén alertas ante la aparición de nuevas herramientas que puedan tener un impacto, positivo o negativo, en cualquier faceta de su ejercicio profesional y en la actividad del alumnado. Solo mediante la información, la formación y la experimentación se puede actuar adecuadamente ante la innovación digital.

1.2. ACTITUDES ANTE LA INTELIGENCIA ARTIFICIAL Y CHATGPT

Todo avance tecnológico, y más aquellos que se introducen rápidamente en el ámbito cotidiano, produce en sus (potenciales) usuarios una gran variedad de reacciones, que abarcan desde el entusiasmo completamente incondicional hasta el rechazo más absoluto. Estas posturas se pueden resumir utilizando el esquema propuesto por Kerschner y Ehlers (2016) reflejado en la Figura 1.



Figura 1. Actitudes ante la tecnología. Adaptado de Kerschner y Ehlers (2016).

Aunque la Figura 1 puede parecer indicar que las actitudes ante la tecnología son algo compartimentado e inamovible, no hay nada más

⁵ En <https://www.demandsage.com/chatgpt-statistics/> se encuentran estadísticas más completas sobre la penetración de ChatGPT comparadas con las de otros avances tecnológicos de la última década.

lejos de la realidad. De hecho, una persona puede mostrar una actitud diferente ante una determinada tecnología dependiendo del uso y del usuario/a que esté considerando. Por ejemplo, la misma persona puede aceptarla como herramienta para el ocio, pero censurarla para el estudio, o utilizarla como herramienta de investigación mientras que muestra escepticismo ante su implementación en la enseñanza. También la actitud cambia con el tiempo y se va perfilando de acuerdo con el conocimiento que se adquiera sobre la utilidad, real o imaginada, de la tecnología en cuestión.

El caso concreto de ChatGPT, que fue lanzado por OpenAI al público general en noviembre de 2022, es mucho más interesante todavía ya que se trata de una tecnología que en cinco días sobrepasó el millón de usuarios⁵ y en menos de seis meses ha llegado a la cifra de 1.800 millones de visitas en su sitio web. Dada su popularidad, unida a la falta de transparencia sobre el entrenamiento de los algoritmos en los que se basa y a la competencia despiadada que su lanzamiento ha activado entre las grandes empresas tecnológicas, resulta lógico que las actitudes ante su implementación en el aprendizaje de lenguas se sitúen a lo largo de un continuo fluido y que las razones que se alegan para aceptar o no su uso sean muy variadas, como se observa en el organigrama de la Figura 2 y se justifica en el siguiente apartado.

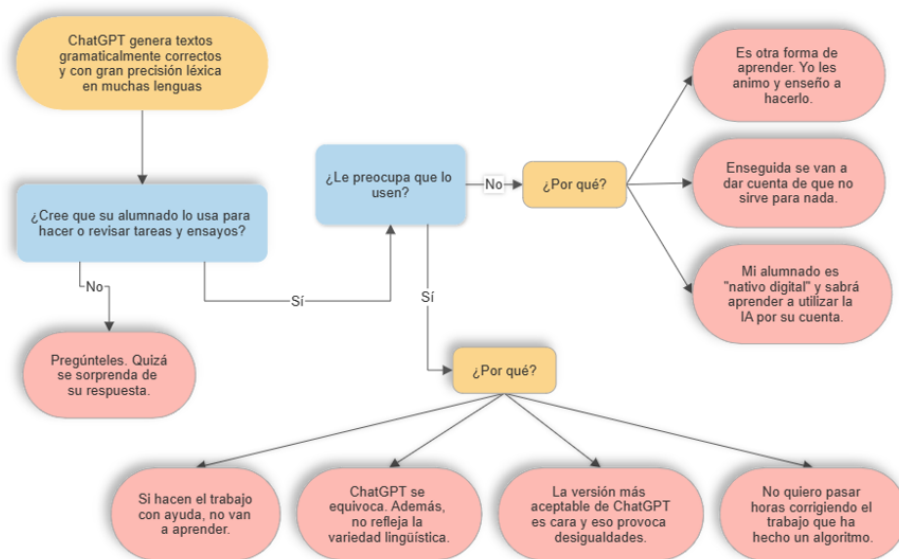


Figura 2. ChatGPT en el aprendizaje de L2. Qué piensa el profesorado. (Elaboración propia).

1.3. CHATGPT Y LA ENSEÑANZA DE SEGUNDAS LENGUAS

A pesar de la breve andadura de ChatGPT y de los pocos estudios que se han realizado a fecha de mayo de 2023 sobre su utilidad para la enseñanza / aprendizaje de segundas lenguas, resulta posible identificar ya algunos de sus retos más sobresalientes, tanto a nivel tecnológico como en cuanto al contenido y a las cuestiones éticas que rodean el entrenamiento de GPT (véase, por ejemplo, Bajohr, 2023; Bjork, 2023; Eliot, 2022; Rettberg, 2022). Entre ellos destacan, en cuanto al *output*, su falta de diversidad lingüística, estilística y cultural, las alucinaciones (esto es, la información inventada que presenta como si fuera real), la imposibilidad de verificar su corrección / veracidad mediante referencias externas al sistema, la autocensura, además del sexismo, el edadismo, la xenofobia, el normativismo y

otros sesgos en el *output* originados por la falta de neutralidad existente en los datos utilizados para el entrenamiento del sistema.

Con referencia a los aspectos tecnológicos hay que subrayar las interrupciones imprevistas e intermitentes en el servicio a determinadas horas sin motivo aparente o tras un número limitado de consultas, lo que impide programar su uso de forma regular, por ejemplo, para una clase. También se debe señalar la necesidad de abonarse si se quiere acceder a la versión más potente hasta la fecha, ChatGPT Plus, la poca transparencia en cuanto a cómo se generan las respuestas y cómo aprende el sistema de ellas y el ingente consumo de energía necesario para mantener la potencia de computación necesaria para entrenar y permitir las interacciones con el sistema. Otro gran desafío lo suponen los cambios vertiginosos en la herramienta y en su integración con otras aplicaciones.

Asimismo, hay que mencionar los problemas relativos a la autoría del *output*, la cuestión de los derechos de autor de los textos utilizados para entrenar GPT, así como el fenómeno de antropomorfización de la herramienta, a la que muchas personas acuden como si se tratara de un oráculo o de una aplicación que puede pensar, dar consejos y proporcionarles compañía como si fuera un ser humano. Está claro también que durante el desarrollo de ChatGPT y de los planes para que el público general lo comenzara a usar no se prestó demasiada atención a lo que pudieran aportar colectivos como la comunidad educativa y las contribuciones de los campos de la ética y la filosofía.

Además, hay que recordar que, aunque para el entrenamiento de GPT se utilizaron textos de diversas lenguas, la proporción de textos y supervisión humana en inglés es desmesurada en comparación con otras lenguas, sobre todo aquellas con menor presencia en las fuentes que se usaron para su entrenamiento. Esto quiere decir que, si bien ChatGPT puede funcionar en más de 100 lenguas, la calidad

del *output* que genera es muy desigual ya que en los casos en los que no cuenta con suficiente entrenamiento en la lengua en cuestión recurre a la traducción, con lo cual los textos generados, aunque puedan ser gramaticalmente correctos, se basan en modelos textuales propios del inglés, como el famoso “ensayo de cinco párrafos” (Rettberg, 2022).

Si bien la variedad y relevancia de los retos mencionados en el párrafo anterior es preocupante, al mismo tiempo, su existencia es precisamente lo que abre un gran horizonte de posibilidades para nuevas líneas de investigación, como ya reflejan las experiencias realizadas y los estudios publicados hasta la fecha. En el campo de la enseñanza de lenguas y la educación cabe destacar los relativos al uso de ChatGPT para generar materiales, programas de asignaturas, actividades y exámenes. (Román Mendoza, 2023; Silió 2023), para ayudar al estudiantado como herramienta para el autoaprendizaje y la realización de actividades de refuerzo (Mills, 2023), para la formación del profesorado (Heimans *et al.*, 2023), para plantearse la enseñanza y la práctica de la escritura en una segunda lengua de forma que el uso de ChatGPT por parte del alumnado no sea considerado plagio u otro tipo de transgresión contra las normas de integridad académica (Warschauer *et al.*, 2023) y como herramienta de traducción (Calvo-Ferrer, comunicación personal, 8 de mayo de 2023). Dado que los chatbots basados en inteligencia artificial llevan estudiándose en el campo de la enseñanza asistida por ordenador desde principios del siglo XXI -véase una revisión de la literatura en Belda-Medina y Calvo-Ferrer (2022)- una de las primeras metas de investigación que resulta oportuno establecer consiste en determinar si ChatGPT presenta las mismas ventajas y retos que sus predecesores. Los apartados que siguen tienen como objetivo intentar dar respuesta a esta cuestión.

2. METODOLOGÍA

Los chatbots no son una herramienta desconocida en la enseñanza de lenguas. En su revisión de la literatura, Belda-Medina y Calvo-Ferrer (2022) señalan que su uso como agente conversacional contribuye a la reducción de la ansiedad en el alumnado y al aumento tanto de la cantidad de tiempo dedicada a utilizar la L2 como de la autoconfianza y motivación de los aprendientes, provocado todo ello, en parte, por el efecto “novedad” que supone la introducción de los chatbots. También señalan la poca naturalidad de la lengua generada por los chatbots y la aparición de frecuentes malentendidos y alucinaciones que resultan en una comunicación poco eficaz entre el alumnado y la herramienta.

Puesto que ChatGPT es un chatbot avanzado, cabe plantearse si su uso podría ayudar a superar las limitaciones planteadas por Belda-Medina y Calvo-Ferrer (2022). Dadas las dificultades para garantizar un acceso regular a la herramienta mencionadas en el apartado anterior, el presente estudio se planificó de forma que no fuera necesaria la participación directa del alumnado pero que, aun sin esta, permitiera analizar qué capacidad tiene ChatGPT para procesar y generar texto en la L2, en este caso español, que sea apropiado para el nivel en el que se quiera implementar. De esta forma, el objetivo del estudio fue evaluar de antemano los retos que plantearía la implementación de ChatGPT en el aula como agente conversacional y, por lo tanto, determinar estrategias que pudieran contrarrestar sus efectos negativos. A tal efecto, se partió de las siguientes preguntas:

1. ¿Entendería el alumnado las respuestas generadas por ChatGPT?
2. ¿Entendería ChatGPT las preguntas generadas por el alumnado?

Para proceder a su respuesta, la autora de este estudio elaboró dos corpus de preguntas junto a las contestaciones generadas por ChatGPT, las cuales fueron etiquetadas, categorizadas y analizadas en cuanto a su adecuación gramatical, ortográfica y pragmática, así como la de su contenido. Estas interacciones consistieron en preguntas individuales formuladas a partir de dos tipos de *input* procedente de dos cursos de español como L2. En el primer caso se usaron los enunciados o *prompts* de 32 actividades de respuesta abierta extraídas de un libro de texto de español de nivel inicial y, en el segundo, 124 dudas formuladas por el alumnado en el foro de preguntas de un curso de escritura en español a nivel intermedio-bajo impartido en 2014.

2.1. CURSO DE NIVEL INICIAL

El primer corpus de preguntas-respuestas se elaboró utilizando los enunciados de todos los ejercicios de respuesta abierta de cuatro⁶ de los ocho capítulos asignados en un curso intensivo de español de nivel inicial impartido de forma totalmente asincrónica en primavera de 2023 en una universidad pública de la costa este de Estados Unidos⁷. Las interacciones con ChatGPT -realizadas indistintamente con sus versiones gratuita y de pago- las llevó a cabo la profesora (y autora de este artículo), al igual que el posterior análisis. De las 32 actividades asignadas, 26 estaban basadas en enunciados textuales, 2 eran de comprensión auditiva, 2 de interpretación de imágenes y dos basadas en vídeos. Todas, excepto una, exigían respuestas escritas.

⁶ El capítulo preliminar, el 2, el 4 y el 6.

En cuanto al nivel, el curso equivale a un A1 y parte de un A2 del *Marco común europeo de referencia para las lenguas*, aunque hay que señalar que este curso atrae a hablantes de español de herencia con un nivel más alto que desean repasar lo que ya saben o no encuentran una alternativa más adecuada para encajar un curso de español en sus horarios. La diferencia de nivel entre los estudiantes se hace patente desde las primeras semanas, como queda reflejado en los ejemplos de la Tabla 1.

Tabla 1. Ejemplo de respuestas de estudiantes en el curso de español inicial.

PROMPT	ESTUDIANTE 1	ESTUDIANTE 2
¿Cómo es tu mejor amigo?	Mi mejor amigo es responsable.	Mi mejor amigo es activo.
Make a list of your favorite athletes and artists, then write their nationality. Choose at least three different nationalities. Modelo: <i>Shakira es colombiana.</i>	Bad Bunny is Puertorriqueño Camila Cabello es Cubana Sofía Vergara es Colombiana Leonel Messi es Argentino Ursula Corbero es Española	LeBron James es norteamericano. Sergio Ramos es español. Luis Hernández es mexicano.
Write a list of the qualities that you think a good friend must have. Use complete sentences.	Un buen amigo es divertido, simpático y comprensión.	Un buen amigo es trabajador. Un buen amigo es divertido Un buen amigo es listo Un buen amigo es lindos

⁷ El libro de texto utilizado en este curso es *Mosaicos: Spanish as a World Language* de Elizabeth Guzmán, Paloma Lapuerta y Judith Liskin-Gasparro, de la editorial Pearson.

Para las interacciones con ChatGPT se utilizó siempre que se pudo el *prompt* literal del libro (13 de ellos aparecían en español y 19, en inglés). Hay que recordar que ChatGPT contesta en español, aunque los *prompts* estén en inglés siempre que la sesión de trabajo haya parecido indicar que el usuario prefiere recibir las respuestas en español, como se ve en los ejemplos de las Tablas 2 y siguientes. En algunas preguntas de carácter personal, ChatGPT no generó ninguna respuesta o lo hacía negándose a contestar alegando que al ser una inteligencia artificial no podía responder preguntas de índole personal. En estos casos se procedió a cambiar el enunciado con objeto de obtener una respuesta (véase la Tabla 2 y la Tabla 3). También hubo que modificar los *prompts* para las preguntas que se apoyaban en gráficos. Para las preguntas basadas en la interpretación de vídeos la mejor manera de generar una respuesta por parte de ChatGPT consistió en copiar el guion y plantear a continuación las preguntas sobre dicho guion. También se probó una técnica combinada usando ChatGPT como traductor ya que, en estos niveles, ChatGPT traduce casi perfectamente, por lo menos en el par de lenguas (español e inglés) utilizado en este estudio. Para las preguntas de comprensión auditiva se transcribió el texto.

Tras llevar a cabo las interacciones y el análisis de las respuestas a los 32 *prompts*, se puede afirmar que ChatGPT:

- Acepta indistintamente *prompts* en inglés o español, y bilingües, tanto si se trata de los enunciados literales del libro como si son traducciones de estos o enunciados parafraseados.

- No genera respuestas con graves errores ni gramaticales ni léxicos.
- Normalmente, genera textos con estructuras y léxico más avanzados que los correspondientes a un A1-A2.
- No contesta bien todas las preguntas en cuanto al contenido, incluso tras parafrasear el *prompt*.
- Introduce diferentes tipos de advertencias en el caso de preguntas con contenido personal, pero sin sistematicidad (es decir, a veces contesta preguntas personales y en otras ocasiones alega que su condición de inteligencia artificial le impide responder).

En las siguientes tablas se recogen algunos ejemplos que ilustran las afirmaciones anteriores.

Tabla 2. Ejemplos de respuestas generadas por ChatGPT.⁸

OBSERVACIÓN	RESPUESTA DE CHATGPT
ACEPTACIÓN DE PROMPTS BILINGÜES	
¿Qué fecha es hoy?	Hoy es 17 de abril de 2023.
Write a descriptive word in Spanish for these celebrities: Miley Cyrus, Lebron James, Jennifer Lopez, Antonio Banderas	Miley Cyrus: polémica LeBron James: atlético Jennifer Lopez: talentosa Antonio Banderas: carismático

⁸ En las siguientes tablas aparecen señalados en LETRAS MAYÚSCULAS los cambios introducidos para que ChatGPT generara respuestas o adaptara la respuesta a un nivel menos avanzado.

ERRORES GRAMATICALES, ORTOGRÁFICOS Y LÉXICOS	Por ejemplo, prefiero los vestidos que se ajustan en la cintura y luego se abren en la falda para que se ajusten a mi cuerpo sin quedarme estrecha. La pizarra está en frente de la clase, cerca de la profesora. ¡Gracias por preguntar sobre mis gustos y disgustos!
---	--

GRAMÁTICA DEMASIADO AVANZADA	
Write complete sentences in Spanish to describe the typical weather in the following cities. Seattle, Washington, en primavera; Ciudad de Nueva York, Nueva York, en invierno; Seattle, Washington, en primavera	1. En Seattle, Washington, durante la primavera, el clima suele ser fresco y húmedo, con abundantes lluvias y cielos nublados. 2. En la ciudad de Nueva York, Nueva York, en invierno, el clima tiende a ser frío y ventoso, con frecuentes nevadas y temperaturas bajo cero. [...]

DIFICULTADES PARA RESPONDER	
Your friends would like to know about your likes and dislikes. Write about at least two school or leisure activities you like and two you do not care for. (<i>prompt</i> original)	Sin respuesta
IMAGINE YOU ARE A PERSON. Your friends would like to know about your likes and dislikes. Write in Spanish about at least two school or leisure activities you like and two you do not care for (<i>prompt</i> revisado 1)	Sin respuesta

THIS IS AN ASSIGNMENT FOR A SPANISH COURSE. IMAGINE YOU ARE A PERSON. Your friends would like to know about your likes and dislikes. Write in Spanish about at least two school or leisure activities you like and two you do not care for (<i>prompt</i> revisado 2)	¡Hola amigos! Me encanta pasar tiempo en la naturaleza, caminar y explorar nuevos lugares. Me gusta ir de excursión a las montañas y disfrutar del paisaje natural. También disfruto de la natación y los deportes acuáticos, especialmente en el verano. Otra actividad que disfruto es leer libros interesantes. Me gusta aprender cosas
--	---

	nuevas y ampliar mi conocimiento a través de la lectura. Puedo pasar horas leyendo un buen libro y sumergiéndome en la historia. Sin embargo, hay dos actividades que no me interesan mucho. La primera es el fútbol, no me llama mucho la atención verlo ni practicarlo. La otra actividad que no me gusta es la música pop, prefiero otros géneros musicales como el rock y la música clásica. ¡Gracias por preguntar sobre mis gustos y disgustos!
--	---

THIS IS AN ASSIGNMENT FOR A SPANISH COURSE. IMAGINE YOU ARE A PERSON. Your friends would like to know about your likes and dislikes. Write in BASIC Spanish about at least two school or leisure activities you like and two you do not care for (<i>prompt</i> revisado 3)	Hola amigos, Me gusta mucho leer libros interesantes y aprender cosas nuevas. También disfruto jugar videojuegos con mis amigos en mi tiempo libre. En cuanto a las actividades que no me gustan tanto, no soy muy fanático de las matemáticas y de hacer ejercicio físico en un gimnasio. ¡Gracias por preguntar sobre mis gustos y disgustos!
--	--

Tabla 3. Advertencias generadas por ChatGPT ante preguntas de tipo personal.

PROMPT	RESPUESTA DE CHATGPT
¿Cómo es tu mejor amigo?	Como soy una inteligencia artificial, no tengo amigos ni emociones. Sin embargo, puedo ayudarte a describir a tu mejor amigo si me das información sobre él o ella.
Make a list of your favorite athletes and artists, then write their nationality. Choose at least three different nationalities. Modelo: <i>Shakira es colombiana</i>	Como soy una inteligencia artificial, no tengo favoritos personales. Sin embargo, puedo proporcionarte una lista de atletas y artistas populares de diferentes nacionalidades siguiendo el modelo que proporcionaste: Rafael Nadal es español. Beyoncé es estadounidense.

	Cristiano Ronaldo es portugués. Céline Dion es canadiense.
Write a list of the qualities that you think a good friend must have. Use complete sentences.	Aunque soy una inteligencia artificial y no tengo emociones ni amigos, puedo proporcionar una lista de cualidades que, en general, se consideran importantes en una buena amistad. Aquí tienes algunas: [...]
¿Tienes una familia grande? ¿Cómo se llaman tus padres? ¿Cuántos hermanos tienes? ¿Cómo se llaman tus primos?	Como soy una inteligencia artificial, no tengo familia en el sentido humano. No tengo padres, hermanos ni primos. Estoy aquí para responder tus preguntas y ayudarte en lo que necesites. Si tienes alguna otra pregunta, no dudes en hacerla.

Uno de los retos mayores consistió en intentar que ChatGPT generara respuestas más acordes con el nivel del curso, como se refleja en los ejemplos de la Tabla 4. Para ello, se añadió “in basic Spanish” o expresiones similares en inglés o español. En otros casos, se incluyó un límite de palabras para la respuesta o una especificación sobre el tipo de verbos que podía usar en ella. Por último, también se solicitó que regenerara la respuesta para comprobar si el resultado era más sencillo.

Tabla 4. Reformulación de *prompts* para generar respuestas más acordes al nivel inicial.

PROMPT	RESPUESTA DE CHATGPT
<p>1a. Think about your university life and describe the following people, places, or things (Mis clases, la cafetería, mi mejor amiga). Be sure to modify the adjectives as necessary so that they agree with the things or people being described. Modelo: <i>el/la profesor/a de español</i> <i>La profesora de español es inteligente y simpática.</i></p>	<p>Mis clases: Mis clases son interesantes y desafiantes, lo que me ayuda a aprender y crecer. La cafetería: La cafetería es amplia y acogedora, con una gran variedad de opciones deliciosas para comer. Mi mejor amiga:</p>

	Mi mejor amiga es divertida y leal, siempre está dispuesta a escuchar y apoyarme en momentos difíciles.
--	---

<p>1b. Think about your university life and describe the following people, places, or things (Mis clases, la cafetería, mi mejor amiga). Be sure to modify the adjectives as necessary so that they agree with the things or people being described. USE BASIC SPANISH. Modelo: <i>el/la profesor/a de español</i> <i>La profesora de español es inteligente y simpática.</i></p>	<p>Mis clases: Mis clases son buenas y difíciles, aprendo mucho. La cafetería: La cafetería es grande y tiene comida variada. Mi mejor amiga: Mi mejor amiga es divertida y fiel, me apoya siempre.</p>
--	---

<p>2a. Write complete sentences in Spanish to describe the typical weather in the following cities. Seattle, Washington, en primavera; Ciudad de Nueva York, Nueva York, en invierno; Seattle, Washington, en primavera.</p>	<ol style="list-style-type: none"> 1. En Seattle, Washington, durante la primavera, el clima suele ser fresco y húmedo, con abundantes lluvias y cielos nublados. 2. En la ciudad de Nueva York, Nueva York, en invierno, el clima tiende a ser frío y ventoso, con frecuentes nevadas y temperaturas bajo cero. 3. En Seattle, Washington, en primavera, las temperaturas suelen ser templadas, y los días son mayormente lluviosos y grises, dando lugar a una exuberante vegetación.
---	--

<p>2b. Write complete sentences OF NO MORE THAN 10 WORDS in Spanish to describe the typical weather in the following cities. Seattle, Washington, en primavera; Ciudad de Nueva York, Nueva York, en invierno; Seattle, Washington, en primavera</p>	<ol style="list-style-type: none"> 1. En Seattle, Washington, durante la primavera, el clima suele ser fresco y húmedo, con abundantes lluvias y cielos nublados. 2. En la ciudad de Nueva York, Nueva York, en invierno, el clima tiende a ser frío y ventoso, con frecuentes nevadas y temperaturas bajo cero. 3. En Seattle, Washington, en primavera, las temperaturas suelen ser templadas, y los días son mayormente lluviosos y grises, dando lugar a una exuberante vegetación.
---	--

<p>3a. How would you describe the following people in Spanish (Miley Cyrus, mi amiga, mi padre)? Write a sentence for each, expressing your opinion. Remember to give the correct form of the adjective according to the gender of the subject, when necessary.</p>	<p>Miley Cyrus es una cantante talentosa y polémica. Mi amiga es inteligente, divertida y siempre solidaria. Mi padre es sabio, trabajador y tiene un gran corazón.</p>
--	---

<p>3b. How would you describe the following people in Spanish (Miley Cyrus, mi amiga, mi padre)? Write a sentence for each, expressing your opinion. Remember to give the correct form of the adjective according to the gender of the subject, when necessary. DO NOT USE VERBS OTHER THAN SER.</p>	<p>Miley Cyrus es talentosa, polémica y carismática. Mi amiga es inteligente, divertida y solidaria. Mi padre es sabio, trabajador y bondadoso.</p>
---	---

2.2. CURSO DE EXPRESIÓN ESCRITA (NIVEL INTERMEDIO)

El segundo corpus de preguntas-respuestas⁹ se elaboró utilizando las dudas formuladas por el alumnado matriculado en un curso de expresión escrita en español de nivel intermedio-bajo impartido de forma totalmente asincrónica en primavera de 2018 en la misma universidad pública de la costa este de Estados Unidos mencionada en el apartado anterior. Las interacciones con ChatGPT las llevó a cabo la profesora (y autora de este artículo), al igual que el posterior análisis. El contenido de las 124 dudas utilizadas se distribuía de la siguiente manera:

- 42 Dudas sobre vocabulario.
- 31 Dudas sobre acentuación, puntuación, ortotipografía y pronunciación.

⁹ Se puede solicitar una copia de este corpus enviando un correo electrónico a la autora de este artículo.

- 25 Cuestiones específicas sobre los contenidos de la clase (es decir, el lenguaje académico y el lenguaje comercial).
- 19 Dudas sobre gramática (sintaxis y morfología).
- 6 Preguntas sobre aspectos técnicos relativos al curso o a la plataforma.
- 1 Pregunta incomprensible.

El nivel de los participantes en esta clase correspondía a un B1-B2, aunque, como en el caso anterior, también había un porcentaje de hablantes de español de herencia con un nivel mucho más alto. Las características del español escrito por el alumnado de esta clase se pueden observar en los *prompts* de los incluidos en las tablas en este apartado, empezando con la Tabla 5, que contiene dos ejemplos extraídos del foro del curso que sirvió de base para compilar el segundo corpus con las respuestas de la profesora.

Tabla 5. Ejemplo de dudas de estudiantes en el curso de español escrito.

DUDAS DE ESTUDIANTES	RESPUESTAS DE LA PROFESORA
<p>Allí, aquí y allá son adjetivos demostrativos y llevan tilde, correcto? Porqué dice el texto que los demostrativos nunca llevan tilde?</p>	<p>No, en realidad son adverbios de lugar que hacen referencia a un lugar más o menos cercano. Los demostrativos son este, ese, aquel y sus formas para femenino y plural. Allí, aquí y allá llevan tilde, sí.</p>
<p>Entiendo que hay reglas de la acentuación depende de si una palabra es una palabra aguda o una palabra llana, pero tengo problemas para distinguir la diferencia cuando yo oigo la palabra. Yo puedo memorizar cuando</p>	<p>La mayoría de las palabras en español son llanas, pero tienes razón: es difícil saber dónde van las tildes si no se conoce cómo se pronuncia una palabra. Puedes ir a https://forvo.com/languages/es/ para escuchar la pronunciación y a</p>

una palabra necesita una tilde pero ¿hay una manera a ayudarme oír donde la acentuación está en una palabra?	http://tip.iatext.ulpgc.es/silabas/default.aspx para ver la separación en sílabas. Además, mediante la lectura y la escritura irás mejorando. Es una cuestión de práctica.
--	--

Tras llevar a cabo las interacciones y el análisis de las respuestas a las 124 dudas, se puede afirmar que ChatGPT:

- Entiende todo tipo de enunciados escritos por el alumnado, incluso aquellos que contienen errores. También intenta generar respuestas para preguntas incomprensibles.
- Genera graves errores de contenido y frecuentes alucinaciones. Solo el 39,9% de las respuestas fueron aceptables (véase la Tabla 6).
- Produce textos que frecuentemente presentan la misma estructura (cuatro o cinco párrafos) plagada de fórmulas repetitivas.
- Genera textos con muchos calcos del inglés en cuanto a la estructura de la respuesta y la ortotipografía.
- A veces proporciona recomendaciones sobre sitios web, reales o inventados sin dar los enlaces o incluyendo enlaces inexistentes.
- Normalmente, genera textos con estructuras y léxico comprensibles por aprendientes de español B2, pero demasiado largos.

Para ilustrar las afirmaciones anteriores se han elaborado las siguientes tablas.

Tabla 6. Porcentaje de respuestas de ChatGPT aceptables.

TIPO DE PREGUNTA	NÚMERO DE RESPUESTAS ACEPTABLES	PORCENTAJE
Vocabulario (N=42)	11	26,2%
Acentuación, puntuación, pronunciación y ortotipografía (N=31)	9	29,0%
Estilo -específicas sobre temas del curso (N=25)	17	68,0%
Gramática (N=19)	2	10,5%
Aspectos técnicos (N=6)	3	50,0%
Incomprensible (N=1)	0	0,0%
Total=124	42	39,9%

Tabla 7. Respuestas generadas por ChatGPT.

OBSERVACIÓN Y EJEMPLO DE DUDA	RESPUESTA DE CHATGPT
<p>RESPUESTAS A PREGUNTAS IMPOSIBLES</p> <p>Una pregunta si tiene alguna sugerencia sobre cómo escribir el octographia que está separado o juntos o con un acento o no cuando se trata de verbal? Desde que cuando lo escribes sabes que necesitas escribir, pero cuando lo oyes puede ser difícil diferenciarlo.</p>	<p>No estoy seguro de entender exactamente a qué te refieres con "octographia". Si te refieres a la ortografía en general, hay algunas reglas generales que pueden ayudarte a distinguir la escritura correcta de las palabras, incluso cuando las escuchas.</p> <p>En cuanto a la separación o unión de palabras, en español, las palabras compuestas pueden escribirse juntas o separadas, dependiendo de la convención y el uso. Por ejemplo, "sobretudo" y "sobre todo" son dos formas aceptables de escribir la misma expresión, pero cada una puede tener un uso y un significado ligeramente diferente.</p> <p>En cuanto al acento, las reglas de acentuación en español se basan en la posición de la sílaba tónica o acentuada en una palabra. Existen reglas específicas para determinar la sílaba tónica en</p>

	<p>diferentes tipos de palabras, como agudas, graves o esdrújulas.</p> <p>Si tienes dificultades para identificar la sílaba tónica en una palabra, puede ser útil practicar con ejercicios de acentuación y estudiar las reglas específicas para cada tipo de palabra.</p> <p>En general, la práctica y la familiarización con las reglas de ortografía y acentuación en español pueden ayudarte a mejorar tu escritura y a identificar los errores de manera más efectiva.</p>
--	---

<p>FÓRMULAS EXTRAÑAS O AJENAS AL ESPAÑOL</p> <p>Hay otra forma adecuada para decir buen provecho?</p>	<p>Sí, hay otras formas adecuadas para decir "buen provecho" en español, dependiendo del país o la región en la que te encuentres. Algunas de estas formas son:</p> <p>Que aproveche Disfruta tu comida ¡Buen apetito! ¡Que tengas un buen almuerzo / cena! Buen provecho en tu comida</p> <p>Todas estas formas son comunes y apropiadas para desear a alguien que disfrute su comida.</p>
--	---

<p>FÓRMULAS EXTRAÑAS O AJENAS AL ESPAÑOL</p> <p>¿Básicamente la lectura dice que no utilizan el "Oxford Comma" del inglés en español? La lectura da un ejemplo sin el "Oxford Comma"; " In the park there were trees, bushes, flowers and a lovely statue". ¿Entonces, el "Oxford Comma" del inglés es un error si lo utilizamos en español después del penúltimo elemento?</p>	<p>En español, el uso del "Oxford comma" (también conocido como "coma vocativa" o "coma enumerativa") es opcional y no se considera un error si se utiliza o no después del penúltimo elemento en una lista.</p> <p>Aunque en inglés se utiliza comúnmente el "Oxford comma" para evitar confusiones en las listas, en español la estructura gramatical y la posición de los elementos de una lista son diferentes, lo que hace que el uso del "Oxford comma" sea menos común.</p> <p>En general, el uso del "Oxford comma" en español depende del estilo y preferencias del escritor o del medio en el que se publique el texto. Si decides utilizarlo,</p>
--	--

	<p>asegúrate de ser consistente en su uso en todo el texto para evitar confusiones o inconsistencias.</p>
--	---

<p>ERRORES DE CONTENIDO Y ALUCINACIONES</p> <p>No entiendo la regla de la doble acentuación, ¿existen palabras con dos acentos? Que reglas siguen estas palabras?</p>	<p>No existen palabras en español con dos acentos escritos en la misma sílaba. Sin embargo, existen palabras que llevan acento en más de una sílaba, lo que se conoce como "acentuación diacrítica" o "doble acentuación". Esto se hace para diferenciar palabras que tienen la misma forma pero que tienen significados diferentes, como "tú" y "tu", "él" y "el", "más" y "mas", entre otras.</p> <p>La regla general de la acentuación diacrítica es que el acento se coloca en la sílaba que corresponde según las reglas generales de acentuación, pero solo se coloca en la palabra que lo requiere para diferenciarse de otra palabra con la misma forma. Por ejemplo, la palabra "tú" lleva acento en la segunda sílaba según las reglas generales de acentuación, pero la palabra "tu" no lo lleva porque es una palabra monosílaba.</p> <p>Otro ejemplo es la palabra "él", que lleva acento en la primera sílaba para diferenciarse del artículo "el". La palabra "el" no lleva acento porque es una palabra monosílaba.</p> <p>Es importante recordar que la acentuación diacrítica se utiliza para diferenciar palabras con la misma forma, pero con significados diferentes, y que las reglas generales de acentuación se aplican para todas las palabras en español.</p>
--	---

<p>ERRORES DE CONTENIDO Y ALUCINACIONES</p> <p>Me gustaría aprender la diferencia entre las frases "antes que" y "antes de que". Usualmente yo leo "antes que". ¿Significa algo diferente si lleva la palabra "de"?</p>	<p>Sí, hay una diferencia significativa en el significado de estas dos frases. "Antes que" es una conjunción que se usa para expresar una acción que ocurre antes de otra acción, pero sin la necesidad de un verbo conjugado en el subjuntivo. Por</p>
--	---

	<p>ejemplo: "Voy a comer antes que salga el sol".</p> <p>En cambio, "antes de que" se usa con un verbo conjugado en el subjuntivo y significa "antes de que ocurra algo". Por ejemplo: "Voy a comer antes de que salga el sol".</p> <p>Entonces, la diferencia entre las dos frases es que "antes que" no necesita el subjuntivo, mientras que "antes de que" siempre va seguido de un verbo conjugado en el subjuntivo.</p>
--	--

<p>ERRORES DE CONTENIDO Y ALUCINACIONES</p> <p>La lectura habla de la coma decimal la cual se dice sigue reglas opuestas al inglés. El texto dice: "así como 2 827 232 en español, se escribe 2,827,232 en inglés". Mi pregunta es, ¿los montos en español solo usan espacios para separar los números y no lleva ningún símbolo de puntuación, ya sea punto o coma?</p>	<p>En español, se utiliza la coma para separar los miles y el punto decimal para separar los decimales en números con cifras decimales. Por lo tanto, el número que mencionas, "2 827 232", debería escribirse en español como "2.827.232" con puntos separando los miles y no utilizando espacios.</p>
---	---

Las alucinaciones que se observan en este segundo corpus son más frecuentes y llamativas que en el primero ya que el tipo de *prompts* utilizado incluía dudas sobre aspectos lingüísticos, estilísticos y técnicos. El primer corpus, por el contrario, estaba formado por enunciados muy básicos, típicos de un curso para principiantes, que exigían, por lo general, respuestas muy poco elaboradas.

3. ANÁLISIS

El análisis de los dos corpus elaborados para este estudio permite afirmar que ChatGPT "interpreta" siempre las preguntas que se le hacen ya sean en inglés o español, o contengan lo que, desde una perspectiva normativista, podría considerarse errores. ChatGPT

siempre genera algún texto como contestación, excepto en el caso de cuestiones de índole personal, para cuya respuesta no siempre va más allá de la generación de fórmulas estandarizadas que alegan que una inteligencia artificial no puede responder preguntas de carácter personal. El único caso observado en el que pidió aclaración fue el de un mensaje truncado que resultaba completamente ininteligible. Por lo general, las interacciones analizadas se asemejan hasta cierto punto a las que el alumnado podría tener con hablantes reales de la lengua, aunque con la gran salvedad de que, en el mundo real, es muy probable que el interlocutor o interlocutora tomara la iniciativa y no fuera siempre capaz de contestar sin haber formulado algunas preguntas de clarificación antes de proceder a responder. El hecho de que ChatGPT conteste una pregunta del corpus que era incomprendible dice mucho sobre su falta de "capacidad para callarse" cuando es necesario.

Otro contraste entre las interacciones reales y las de los corpus de este estudio está relacionado con el nivel del español generado por ChatGPT. Si bien es cierto que el alumnado de los niveles iniciales se va a enfrentar siempre a situaciones donde no comprenda todo lo que le dicen sus interlocutores, lo normal es que estos intenten adaptarse al nivel del hablante de L2 en cuanto perciben que puede que no esté entendiendo todo. Por este motivo, las reformulaciones por parte de los L1 son muy típicas en intercambios donde hay una gran diferencia de nivel. Sin embargo, para conseguir que una conversación con ChatGPT sea más parecida a la realidad no es suficiente introducir en los enunciados del L2 acotaciones del tipo "utiliza un español básico" y similares (algo que idealmente podría equivaler a la capacidad de adaptación que emplean los L1 cuando hablan con los L2). Para evitar que la interacción en los niveles iniciales se interrumpa por falta de comprensión por parte de los aprendientes, resulta más efectivo recurrir al uso de ChatGPT como traductor. De esta forma se garantiza la continuidad de la

conversación, pero el uso desmedido de la L1 puede afectar negativamente la adquisición de la L2. Por otra parte, recientemente se ha propuesto en las redes sociales que ChatGPT puede ser entrenado para que imite el estilo de una persona (IA Viajero, 2023). Sin embargo, las pruebas realizadas por la autora de este artículo no consiguieron que ChatGPT generara textos similares a los escritos por el alumnado en este nivel de español L2 o lengua de herencia.

En cuanto a las preguntas de contenido, queda claro que ChatGPT no explica la gramática bien. Tampoco es sistemático en su grado de fiabilidad en las respuestas sobre temas léxicos y ortográficos. Una forma de evitar estas alucinaciones es prescindir de toda clase de preguntas metalingüísticas, sobre todo a la vista de que ChatGPT genera con demasiada frecuencia contestaciones en las que solo una parte del contenido es correcto, pero lo hace con tal seguridad que resulta difícil sospechar de la fiabilidad del resto del contenido, a no ser que se sea experto en el contenido de la respuesta. Hacer que el alumnado entienda esta gran (y, al mismo tiempo, lógica) limitación de ChatGPT es fundamental, máxime en el clima de antropomorfización que rodea a todas las herramientas de inteligencia artificial generativa hoy en día.

El grado de fiabilidad de ChatGPT es mayor a la hora de responder preguntas sobre el estilo de un documento comercial, como cartas de presentación o *curriculum vitae*, pero se nutre excesivamente de modelos predictivos basados en la norma anglosajona, como queda patente en la ortotipografía de los saludos, las fórmulas de cortesía y la estructuración de los párrafos.

Por último, resulta interesante observar que, en algunas ocasiones, ChatGPT introduce fórmulas de desdoblamiento que podrían hacer pensar que está entrenado para ofrecer un discurso inclusivo. Sin embargo, se trata de casos muy aislados que incluso presentan

fórmulas no inclusivas en la misma oración, como el ejemplo recogido en la Tabla 3).

4. CONCLUSIÓN Y PROPUESTAS DE TRABAJO FUTURO

En palabras de ChatGPT (2023):

“El uso de texto generado por inteligencia artificial para enseñar español plantea una serie de oportunidades y desafíos para los profesores y los alumnos. Por un lado, el texto generado por IA puede ofrecer una fuente de input auténtico, variado y personalizado, que puede motivar y enriquecer el aprendizaje de la lengua. Por otro lado, el texto generado por IA puede presentar problemas de calidad, fiabilidad y sesgo, que pueden afectar negativamente al desarrollo de la competencia lingüística y comunicativa. Por lo tanto, es necesario que los profesores y los alumnos adopten una actitud crítica y reflexiva hacia el uso de texto generado por IA, y que lo integren de manera coherente y equilibrada con otras fuentes y actividades de aprendizaje. Así, el texto generado por IA puede ser una herramienta útil y complementaria, pero no sustitutiva, para la enseñanza y el aprendizaje del español”.

Sin embargo, como era de esperar de ChatGPT, es necesario hacer algunas matizaciones al texto anterior, generado en respuesta a la cuestión sobre cuál es su potencial para enseñar español. En primer lugar, “el *input* variado y personalizado” que afirma poder proporcionar no satisface las necesidades de los aprendientes de español como L2. Es cierto que los textos generados por ChatGPT son gramaticalmente correctos en la mayoría de las ocasiones, pero el contenido, sobre todo cuando se trata de dudas sobre la gramática, estilo, ortografía y diversidad cultural es erróneo en un alto porcentaje de casos. Además, se observa una clara influencia del inglés en la estructuración de los párrafos y una urgencia por contestar, aunque la pregunta no esté bien formulada. ChatGPT “entiende” *prompts* con graves errores gramaticales e interpreta con rapidez en qué lengua ha de contestar.

¿Quiere decir todo esto que se debe abandonar la idea de utilizar ChatGPT como agente conversacional? En absoluto. La conclusión a la que se ha de llegar es más bien la contraria. El estudiantado necesita conocer las ventajas y limitaciones del uso de aplicaciones de generación de texto basadas en LLM y solo puede conseguirlo si las utiliza con el apoyo de un profesorado bien informado. No obstante, establecer guías de uso y buenas prácticas en un campo en continuo desarrollo y tan poco transparente como este es una tarea muy compleja. Tampoco ayuda el hecho de que, a nivel general, el debate sobre el uso de la inteligencia artificial se esté polarizando sin matices. A pesar de sus limitaciones, el análisis presentado en este artículo pone de manifiesto las posibilidades que presenta el uso de ChatGPT para el aprendizaje de segundas lenguas. También deja clara la dificultad que entraña establecer guías para el aprovechamiento de las herramientas de generación de texto basadas en la inteligencia artificial. Por todo ello, y como indican Darda, Carre y Cross (2023), van Dis et al., (2023) y Goodlad y Baker (2023), la conclusión más clara a la que se puede llegar en este momento radica en que resulta esencial establecer canales para el trabajo colaborativo que permitan realizar investigaciones de mayor profundidad en cortos plazos de tiempo, además de organizar y actualizar acciones de formación conjuntas. Asimismo, es indispensable que la comunidad educativa exija una mejor legislación y el compromiso ético de todos los actores implicados en el entrenamiento de los LLM, abogue por una inteligencia artificial libre y abierta, apoye la regulación y la verificación humana de sus aplicaciones, y supervise la implementación de todas las herramientas que se generen con ella.

5. BIBLIOGRAFÍA

- Amer, M. (2022). *Large Language Models and Where to Use Them*. 7 de julio.
<https://txt.cohere.ai/llm-use-cases/>
- Bajohr, J. (2023). *On Artificial and Post-Artificial Texts*. Blog de J. Bajohr, 11 de marzo.
<https://hannesbajohr.de/en/2023/03/11/on-artificial-and-post-artificial-texts/>
- Belda-Medina, J. y Calvo-Ferrer, J.R. (2022). "Using Chatbots as AI Conversational Partners in Language Learning". *Applied Sciences*, 12(17), 8427.
<https://doi.org/10.3390/app12178427>
- Bjork, C. (2023). "ChatGPT threatens language diversity. More needs to be done to protect our differences in the age of AI". *The Conversation*, 9 de febrero.
<https://theconversation.com/chatgpt-threatens-language-diversity-more-needs-to-be-done-to-protect-our-differences-in-the-age-of-ai-198878>
- Coniam, D. (2004) "Using language engineering programs to raise awareness of future CALL potential". *Computer Assisted Language Learning*, 17(2), 149–176.
- Coniam, D. (2008). "Evaluating the language resources of chatbots for their potential in English as a second language". *ReCALL: The Journal of EUROCALL*, 20(1), 98–116.
<https://doi.org/10.1017/S0958344008000815>
- Darda, K., Carre, M. y Cross, E. (2023). "Value attributed to text-based archives generated by artificial intelligence". *Royal Society Open Science*, 10(2).
<https://doi.org/10.1098/rsos.220915>

van Dis, E.A.M., Bollen, J., Rooij, R., Zuidema, V. y Bockting, C.L. (2023). ChatGPT: five priorities for research. *Nature*, 614.
<https://www.nature.com/articles/d41586-023-00288-7>

Eliot, L. (2022). "AI Ethics and the Future of Where Large Language Models are Heading", 30 de agosto *Forbes*.
<https://www.forbes.com/sites/lanceeliot/2022/08/30/ai-ethics-asking-aloud-whether-large-language-models-and-their-bossy-believers-are-taking-ai-down-a-dead-end-path/>

Feigenbaum, E.A. (2003). "Some Challenges and Grand Challenges for Computational Intelligence". *Journal of the ACM*, 50(1), 32–40.

Godwin-Jones, R. (2023). "Emerging spaces for language learning: AI bots, ambient intelligence, and the metaverse". *Language Learning & Technology*, 27(2), 6–27.
<https://hdl.handle.net/10125/73501>

Goodlad, L.M.E. y Baker, S. (2023, 20 de febrero). *Now the Humanities can Disrupt "AI"*. Public Books.
<http://www.publicbooks.org/now-the-humanities-can-disrupt-ai/>

Heimans, S., Biesta, G., Takayama, K. y Kettle, M. (2023). "ChatGPT, subjectification, and the purposes and politics of teacher education and its scholarship". *Asia-Pacific Journal of Teacher Education*, 51(2), 105-112.
DOI: <https://doi.org/10.1080/1359866X.2023.2189368>

Huang, W., Hew, K.F. y Fryer, L.K. (2022). "Chatbots for Language learning - Are they Really Useful? A Systematic Review of Chatbot-supported Language Learning". *Journal of Computer Assisted Learning*, 38(1), 237–257.

IA Viajero [@IAViajero]. (2023). *El prompt mágico. Este prompt preparará ChatGPT para recibir textos de entrenamiento (prompt en el ALT de la imagen). Consejo: utilice GPT-4 o Bing Chat Creative Mode.* [vídeo e hipervínculo] [tuit inicial de hilo], 29 de mayo. Twitter.
<https://twitter.com/IAViajero/status/1663198271269945346>

Jia, J. (2003). *The Study of the Application of a Keywords-based Chatbot System on the Teaching of Foreign Languages*.
<https://arxiv.org/abs/cs/0310018v1>

Kerschner, C. y Ehlers, M.-H. (2016). "A framework of attitudes towards technology in theory and practice". *Ecological Economics*, 126(C), 139–151.
DOI: <https://doi.org/10.1016/j.ecolecon.2016.02.010>

Kim, H., Yang, H., Shin, D. y Lee, J. H. (2022). "Design principles and architecture of a second language learning chatbot". *Language Learning & Technology*, 26(1), 1–18.
<http://hdl.handle.net/10125/73463>

Mills, A. (2023). *AI Text Generators. Sources to Stimulate Discussion among Teachers*, febrero-presente.
<https://docs.google.com/document/d/1V1drRG1XlWTBrEwgGqd-cCySUB12JrcoamB5i16-Ezw/edit#heading=h.y7v1xxluoxbv>

Reber, P. (2010). "What Is the Memory Capacity of the Human Brain?" *Scientific American Mind*, 10 de mayo.
<https://www.scientificamerican.com/article/what-is-the-memory-capacity/>

Rettberg, J. W. (2022). *ChatGPT is multilingual but monocultural, and it's learning your values*, 6 de diciembre, jill/txt.
<https://jilltxt.net/right-now-chatgpt-is-multilingual-but-monocultural-but-its-learning-your-values/>

Roberts, I., Watumull, J. y Chomsky, N. (2023). "The False Promise of ChatGPT", 8 de marzo, *The New York Times*.

Román Mendoza, E. (2023). *ChatGPT Retos y oportunidades para el aprendizaje de lenguas*, 22 de febrero.

<https://t.co/2RKmcdCfXA>

Ruiz, A. (2022,). *Introduction to Large Language Models*, 12 de octubre NoCode.ai.

<https://www.nocode.ai/introduction-to-large-language-models/>

Russell, S.J. y Norvig, P. (2020). *Artificial Intelligence: A Modern Approach* (4ª ed.). Prentice Hall.

Silió, E. (2023). "Tus preguntas de Alemán en Selectividad las redactó ChatGPT". *El País*, 25 de mayo.

<https://elpais.com/educacion/2023-05-26/tus-preguntas-de-aleman-en-selectividad-las-redacto-chatgpt.html>

Thompson, A.D. (2022). *What's in my AI? A Comprehensive Analysis of Datasets Used to Train GPT-1, GPT-2, GPT-3, GPT-NeoX-20B, Megatron-11B, MT-NLG, and Gopher*.

<https://LifeArchitect.ai/whats-in-my-ai>

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59, 433–460.

Vázquez-Cano, E., Mengual-Andrés, S. y López-Meneses, E. (2021). "Chatbot to improve learning punctuation in Spanish and to enhance open and flexible learning environments". *International Journal of Educational Technology in Higher Education*, 18(1).

<https://doi.org/10.1186/s41239-021-00269-8>

Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q. y Tate, T. (2023). *The Affordances and Contradictions of AI-Generated Text for Second Language Writers*, marzo.

<http://dx.doi.org/10.2139/ssrn.4404380>

FECHA DE ENVÍO: 31 DE MAYO DE 2023

FECHA DE ACEPTACIÓN: 11 DE JUNIO DE 2023