# INTELIGENCIA ARTIFICIAL Y ENSEÑANZA DE E/LE: POSIBILIDADES DEL PROGRAMA CHATGPT PARA LA PRÁCTICA DE LA EXPRESIÓN E INTERACCIÓN ESCRITAS

José Miguel Blanco Pena
*Universidad de Tamkang (Taiwán)*
*jmblanco@mail.tku.edu.tw*

**Resumen**

En los últimos años, la inteligencia artificial (IA) está revolucionando la forma en que se enseñan y aprenden idiomas en todo el mundo, entre otras razones porque esta tecnología puede ayudar a personalizar el aprendizaje del alumnado, adaptarse a sus necesidades individuales y proporcionarle retroalimentación en tiempo real. Una de las formas en que la IA puede aplicarse a la educación es mediante el uso de los *chatbots*, entre los que últimamente destaca *ChatGPT* por su capacidad para generar textos de manera automática. Ahora bien, ¿qué ventajas ofrece esta tecnología para el entorno de enseñanza/aprendizaje de E/LE?

Presentamos aquí los resultados de un estudio preliminar sobre las posibilidades del programa ChatGPT para la enseñanza y el aprendizaje de la escritura en clase de E/LE. Para ello, hemos probado este programa, por un lado, con tareas de expresión e interacción escritas tomadas del modelo oficial de DELE B1 del Instituto Cervantes y, por otro lado, con textos escritos por nuestros estudiantes. Los hallazgos obtenidos nos llevan a pensar que, utilizada adecuadamente, ChatGPT puede ser una herramienta valiosa, aunque no sustitutoria, tanto para el profesorado como para el alumnado por su capacidad para crear tareas, ofrecer modelos de escritura, elaborar rúbricas y corregir textos.

El estudio que se presenta forma parte de un proyecto de investigación-acción, desarrollado durante el curso 2023-24, que contempla la incorporación de la inteligencia artificial en la clase de Composición en Taiwán, y constituye una nueva aportación a una línea de investigación que llevamos desarrollando en los últimos años sobre la innovación educativa en entornos de E/LE sinohablantes.

## 1. INTRODUCCIÓN

La incorporación de las nuevas tecnologías (NNTT) en la enseñanza del español como lengua extranjera (E/LE) ha transformado significativamente la manera en que se aprende y enseña este idioma (Cruz Piñol, 2014; Volpicelli, 2018). Estas tecnologías ofrecen una variedad de herramientas y recursos que pueden mejorar la eficacia de la enseñanza y hacerla más accesible a una audiencia global. Una de las principales ventajas de las nuevas tecnologías es la posibilidad de acceder a una amplia variedad de recursos multimedia, como videos, audio, imágenes y juegos interactivos. Estos recursos pueden utilizarse para mejorar la comprensión auditiva y la pronunciación, así como para enriquecer el vocabulario y la gramática. Además, los estudiantes pueden acceder a estos recursos en cualquier momento y desde cualquier lugar, lo que les permite aprender a su propio ritmo y de acuerdo con sus necesidades individuales. Otra ventaja de las nuevas tecnologías es la posibilidad de conectarse con estudiantes y profesores de todo el mundo mediante plataformas de aprendizaje en línea, redes sociales y aplicaciones de mensajería. Esto permite a los estudiantes practicar su español con hablantes nativos y mejorar su competencia comunicativa. También pueden recibir retroalimentación y corrección de su español en tiempo real, lo que les ayuda a identificar y corregir errores.

Ahora bien, a pesar de las ventajas de las NNTT en la enseñanza del E/LE, también existen algunos desafíos asociados a la innovación en la docencia universitaria (Reinhardt, 2022). Uno de los mayores retos en este sentido es asegurar que las nuevas estrategias pedagógicas sean efectivas y adoptadas a gran escala. También es importante considerar cómo la innovación puede afectar a los estudiantes con necesidades educativas especiales y cómo se puede garantizar que el acceso a nuevas tecnologías sea equitativo. Por ejemplo, el acceso a Internet y dispositivos tecnológicos no es universal, lo que significa que algunos estudiantes pueden tener dificultades para acceder a los recursos en línea. Además, es posible que algunos estudiantes se distraigan con las tecnologías y pierdan el interés en el aprendizaje (Kuure et al., 2016).

Entre los terrenos educativos en los que se están llevando a cabo innovaciones educativas relacionadas con las NNTT, que además pueden combinarse entre sí para lograr un aprendizaje más efectivo y enriquecedor, podemos señalar los siguientes (Murillo, 2017; López y Heredia, 2017):

1. Aprendizaje personalizado: la tecnología permite adaptar el contenido y el ritmo de aprendizaje a las necesidades individuales de cada estudiante.
2. Aprendizaje colaborativo: las nuevas tecnologías facilitan la colaboración entre estudiantes y profesores, tanto dentro como fuera del aula.
3. Aprendizaje en línea: el aprendizaje en línea permite a los estudiantes acceder a cursos y materiales educativos desde cualquier lugar y en cualquier momento.
4. Aprendizaje basado en proyectos: en lugar de enseñar contenido de manera tradicional, las NNTT favorecen el aprendizaje basado en proyectos enfocados en problemas y desafíos reales que fomentan el pensamiento crítico y la creatividad.
5. Aprendizaje móvil: el aprendizaje móvil permite a los estudiantes acceder a cursos y materiales educativos desde dispositivos móviles, como smartphones y tabletas.
6. Realidad virtual y aumentada: estas tecnologías permiten a los estudiantes experimentar con escenarios y situaciones de manera inmersiva, lo que mejora la comprensión y la retención del conocimiento.
7. Inteligencia artificial y aprendizaje automático: la inteligencia artificial y el aprendizaje automático pueden ayudar a los estudiantes a identificar sus debilidades y proporcionar retroalimentación personalizada.
8. Gamificación: la gamificación consiste en utilizar elementos de juego para motivar y fomentar el aprendizaje.

De estos diferentes entornos, aquí nos interesa especialmente la inteligencia artificial (IA), por su un gran potencial para transformar la educación y mejorar la forma en que se aprende (Roll y Wylie, 2016; Kannan, 2018; Chen et al., 2020; Alam, 2021; Fan y Jiao, 2021; Huang et al., 2021; Hunag et al., 2023). Entre otras áreas educativas, en los últimos años, la IA está revolucionando la forma en que se enseñan y aprenden idiomas en todo el mundo, también en el ámbito del E/LE (Escobar Hernández, 2021; Divekar et al., 2022; Muñoz-Basols et al., 2024). Entre otras razones, porque la IA puede ayudar a personalizar el aprendizaje para cada estudiante, adaptarse a sus necesidades individuales y proporcionar retroalimentación en tiempo real, como se ha dicho más arriba. No obstante, también es importante abordar los desafíos y preocupaciones asociados con su uso para asegurar que se utilice de manera ética y efectiva.

Una de las formas en que la IA puede ser utilizada en la educación es mediante el uso de *chatbots*, es decir, asistentes virtuales y sistemas de tutoría automatizados, un tipo de software que puede proporcionar apoyo personalizado y ayudar a los estudiantes a navegar por el contenido y las tareas (Jeon et al., 2023; Jeon y Lee, 2023; Kasneci et al., 2023). Uno de los programas de IA más recientes, conocidos y avanzados es ChatGPT, un modelo de lenguaje desarrollado por la compañía OpenAI. Creemos que este programa, del que últimamente ser habla mucho en los medios y que a veces se presenta más como una amenaza que una oportunidad para la educación, ofrece numerosas posibilidades para la práctica de la expresión y la interacción escrita en el aula de E/LE, como vamos a mostrar más abajo.

ChatGPT es un modelo lingüístico relativamente nuevo y la investigación sobre su aplicación en la enseñanza de idiomas es actualmente escasa. Sin embargo, podemos encontrar información sobre cómo se pueden utilizar en la enseñanza de idiomas otros modelos similares, como GPT-4. Tanto ChatGPT como GPT-4 son modelos de lenguaje desarrollados por OpenAI, pero hay algunas diferencias clave entre ellos. GPT-4 es el modelo de lenguaje más reciente y avanzado de OpenAI, que ha sido entrenado con una gran cantidad de datos y es capaz de generar texto de manera natural y coherente. Es capaz de generar una variedad de textos que incluye noticias, poesía, código, entre otros. Es también capaz de responder preguntas, traducir idiomas, y realizar tareas de programación. ChatGPT es una versión modificada de GPT-4 que está específicamente diseñada para simular conversaciones en tiempo real. Es capaz de mantener un contexto y responder de manera coherente y natural a preguntas y comentarios en una conversación. A diferencia de GPT-4, ChatGPT es especialmente útil para la interacción escrita y la generación automática de textos, como correos electrónicos, cartas, y mensajes de texto. En resumen, GPT-4 es un modelo de lenguaje más general que puede ser utilizado para una variedad de tareas, mientras que ChatGPT es específico para la interacción en tiempo real y la generación de textos. Aunque estas referencias no son específicamente sobre ChatGPT, dan una idea de las posibilidades y desafíos de usar modelos de lenguaje previamente entrenados en la educación de idiomas extranjeros.

Una de las principales ventajas de ChatGPT es su capacidad para generar textos de manera automática. Esto permite a los estudiantes practicar la escritura de diferentes tipos de textos, como cartas, correos electrónicos, artículos, ensayos, etc. Además, ChatGPT puede ayudar a los estudiantes a mejorar su gramática y su vocabulario al proporcionarles retroalimentación inmediata sobre su escritura mediante correcciones automáticas y sugerencias de palabras. Por un lado, el programa tiene acceso a una base de datos masiva de textos escritos en español, lo que significa que los estudiantes pueden obtener ejemplos de cómo se utilizan correctamente las diferentes estructuras gramaticales y palabras en contexto. Esto es especialmente útil para los estudiantes que

tienen dificultades para comprender cómo se utilizan las palabras y las estructuras gramaticales en contextos reales. Por otro lado, ChatGPT es capaz de reconocer y corregir errores gramaticales y de ortografía en el texto escrito por el estudiante, lo que permite que los estudiantes vean inmediatamente dónde cometieron errores y cómo corregirlos, lo cual es especialmente útil para los estudiantes que tienen dificultades para identificar y corregir sus propios errores.

Otra ventaja de ChatGPT es su capacidad para mantener una conversación con el estudiante en español mediante un "interlocutor" virtual en tiempo real, proporcionando respuestas adecuadas y naturales en una conversación. Esto puede ayudar al alumnado a aprender nuevo vocabulario y expresiones y, en definitiva, a practicar y mejorar su capacidad para interactuar en español escrito de manera fluida, y resulta especialmente útil para los estudiantes que tienen dificultades para interactuar en español escrito con otros hablantes nativos. En cuanto a los profesores, ChatGPT puede ser de gran ayuda en tareas como preparar ejercicios de expresión e interacción escrita, elaborar rúbricas para los alumnos, o corregir los textos de los estudiantes y ofrecerles la correspondiente retroalimentación.

En definitiva, partiendo de nuestra situación de enseñanza-aprendizaje particular, por un lado, y de una noción de innovación educativa centrada en las NNTT y la IA, por otro, presentamos, a continuación, un estudio de las posibilidades que ofrece ChatGPT al profesorado –y por extensión, al alumnado– de E/LE para la clase de expresión e interacción escritas. Se trata de un análisis preliminar de un proyecto de investigación-acción desarrollado en un departamento de español de una universidad taiwanesa durante el curso 2023-2024. Constituye, además, una ampliación de otra investigación realizada anteriormente bajo el título: *El aprendizaje del español centrado en las pautas del DELE y los criterios de puntuación para la prueba de expresión e interacción escrita*, financiado por el Consejo Nacional de Ciencia y Tecnología de Taiwán (Ref. 109-2410-H-032-062-) durante el curso académico 2019-2020 (Blanco Pena, 2024).

## 2. METODOLOGÍA

Como indagación preliminar para nuestra investigación, probamos el programa web ChatGPT con materiales vinculados con la prueba de expresión e interacción escritas de una muestra oficial del examen DELE B1, administrado en 2013 (colgadas en la web del Instituto Cervantes), por un lado, y con textos correspondientes a tareas de DELE B1 escritos por varios de nuestros alumnos de E/LE universitarios taiwaneses de tercer año en junio de 2021, por otro. Aunque experimentamos con las dos tareas, esto es, Tarea 1 y Tarea 2, aquí presentamos solo las pruebas realizadas en relación con la Tarea 1, obligatoria para todos los candidatos de la prueba DELE B1.

En total llevamos a cabo catorce experimentos diferentes. Creemos que se trata de un número suficiente para nuestro objetivo; de hecho, así nos lo han confirmado los sustanciosos resultados obtenidos, como se podrá comprobar más adelante. Decidimos dividir las pruebas en tres tipos, correspondientes a sendas operaciones pedagógicas: A) Responder a la Tarea 1; B) Crear la Tarea 1; C) Corregir la Tarea 1. Concretamente, los experimentos y las instrucciones fueron los siguientes:

A) Respuesta a la Tarea 1

*Prueba 1*

Introducir en ChatGPT las instrucciones de la tarea 1 de la muestra oficial del IC tal cual, es decir, sin añadir ninguna indicación por nuestra parte.

B) Creación de la Tarea 1

*Prueba 2*

Redacta "una tarea 1 de expresión e interacción escritas para una prueba de DELE B1".

*Prueba 3*

Redacta "una tarea 1 de expresión e interacción escritas para una prueba de DELE B1, de tal modo que incluya el texto de interacción".

*Prueba 4*

Redacta las "instrucciones de una tarea 1 de expresión e interacción escritas para una prueba de DELE B1, de tal modo que incluya el texto de interacción, sobre el tema La ciudad".

C) Corrección de la Tarea 1

*Prueba 5*

Corrige "esta tarea de escritura".

*Prueba 6*

Corrige "esta tarea de escritura añadiendo explicaciones de corrección".

*Prueba 7*

Corrige la respuesta (texto) de un alumno a una tarea 1 de una prueba DELE B1 “marcando los errores”.

*Prueba 8*

Corrige la respuesta (texto) de un alumno a una tarea 1 de una prueba DELE B1 “clasificando los errores”.

*Prueba 9*

Corrige la respuesta (texto) de un alumno a una tarea 1 de una prueba DELE B1 “clasificando los errores de acuerdo con las escalas de calificación del DELE B1”.

*Prueba 10*

Corrige la respuesta (texto) de un alumno a una tarea 1 de una prueba DELE B1 “clasificando los errores de acuerdo con la escala analítica del DELE B1, es decir, Adecuación al género, Coherencia, Gramática y Alcance (Vocabulario), dicho de otro modo, según la escala analítica del Instituto Cervantes”.

*Prueba 11*

Corrige la respuesta (texto) de un alumno a una tarea 1 de una prueba DELE B1, ofreciéndole 1) las instrucciones de la tarea, 2) la tarea de alumno (el texto a corregir), de acuerdo con las escalas de calificación de la prueba de expresión e interacción escritas del DELE B1.

*Prueba 12*

Corrige la respuesta (texto) de un alumno a una tarea 1 de una prueba DELE B1, ofreciéndole 1) las instrucciones de la tarea, 2) la tarea de alumno (el texto a corregir), de acuerdo con las escalas de calificación de la prueba de expresión e interacción escritas del DELE B1, que se indican expresamente con sus correspondientes porcentajes.

*Prueba 13*

Corrige la respuesta (texto) de un alumno a una tarea 1 de una prueba DELE B1, ofreciéndole 1) las instrucciones de la tarea, 2) la tarea de alumno (el texto a corregir), de acuerdo con las escalas de calificación de la prueba de expresión e interacción escritas del DELE B1, que se indican expresamente con sus correspondientes porcentajes y puntuaciones de 0 a 4 (No Apto y Apto).

*Prueba 14*

Corrige la respuesta (texto) de un alumno a una tarea 1 de una prueba DELE B1, ofreciéndole 1) las instrucciones de la tarea, 2) la tarea de alumno (el texto a corregir), de acuerdo con las escalas de calificación de la prueba de expresión e interacción escritas del DELE B1, que sin indicar los porcentajes ni las puntuaciones de 0 a 4 (No Apto y Apto), sino solo la puntuación máxima por escala o criterio.

Técnicamente hablando, una vez seleccionadas las muestras con los criterios señalados más arriba, para cada una de las pruebas realizadas, excepto para la primera, hemos utilizado exactamente las instrucciones antes indicadas para insertarlas en el cuadro de diálogo de ChatGPT, con las siguientes diferencias: en los test sobre creación de tareas, el input ha consistido solo en nuestras propias indicaciones; en las pruebas sobre corrección de determinadas tareas, nuestras indicaciones han sido antepuestas o pospuestas a los materiales empleados para el input. En el caso de la prueba de respuesta a una tarea, excepcionalmente y por la naturaleza del experimento, nos hemos limitado a trasladar la muestra tal cual al cuadro de envío de mensajes de la página web. Tras redactar o copiar y pegar de Word el texto completo la prueba en el cuadro de diálogo, y cliquear el icono de “Enviar un mensaje”, el sistema ChatGPT nos respondió con su output en segundos y sin interrupciones.

Finalmente, para valorar los resultados y el grado de cumplimiento de las diferentes tareas llevadas a cabo por ChatGPT (respuesta, creación y corrección de tareas), nos hemos fiado de nuestro conocimiento profesional y nuestra dilatada experiencia docente en Taiwán, que, entre otros aspectos, incluye, desde hace tiempo, la utilización de modelos DELE del Instituto Cervantes, de sus instrucciones y escalas de calificación, en las clases de expresión escrita.

## 3. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS

En este apartado, para una mayor claridad expositiva de los fenómenos observados, organizamos nuestra discusión de los resultados en cuatro secciones: 1) Interacción con el interlocutor, donde valoramos la capacidad de ChatGPT para responder dirigiéndose expresamente a la persona que le ha mandado el mensaje, es decir, para hacer más humana e interactiva su contribución; 2) Realización de tareas, en que analizamos la habilidad del sistema para actuar como candidato al DELE que se somete de la prueba escrita; 3) Creación de tareas, en que se examina la competencia de ChatGPT a la hora de elaborar pruebas de expresión e interacción

escrita tal como son concebidas por el IC; 4) Corrección de errores y comentario, donde analizamos la aptitud del programa de IA como "profesor" de E/LE en su faceta de corregir y comentar textos escritos por los alumnos.

### 3.1. Interacción con el locutor

Después de probar el sistema ChatGPT, observamos varios fenómenos interesantes relacionados con el "talento dialógico" de este programa. Así, por un lado, mientras que en algunos casos (pruebas 1, 5, 6, 7, 8 y 10) el asistente virtual se limita a emitir directamente su respuesta sin más, esto es, sin dirigirse de algún modo a quien le ha enviado el mensaje, en otros sí se da una retroalimentación pragmática que no solo tiene en cuenta las indicaciones de su contraparte, sino que, además, los anima. Por ejemplo, en la Prueba 2, y por partida doble, una vez al principio: "Claro, aquí tienes una posible tarea 1 de expresión e interacción escritas para una prueba de DELE B1"; y de nuevo al final: "Espero que te sea útil. ¡Buena suerte!". Esta tendencia a interacción positiva y empática con quien le escribe unas veces se da al comienzo del output, como el caso ya comentado de la Prueba 2 o, mientras que otras veces surge solo hacia el final, como en la prueba 11: "Espero que esta retroalimentación sea útil. ¡Mucho éxito en tus estudios!".

Sin embargo, en otras ocasiones, como en la Prueba 4, la respuesta se ofrece de manera despersonalizada, automatizada y, por lo tanto, más fría: "Instrucciones para la tarea 1 de expresión e interacción escritas para una prueba de DELE B1 sobre el tema 'La ciudad'". Esto mismo también ocurre en otras pruebas; por ejemplo, en la Prueba 9, no solo al principio: "La tarea de escritura tiene los siguientes errores clasificados según las escalas de calificación del DELE B1", sino también al final: "En general, la tarea de escritura tiene pocos errores y se entiende claramente el mensaje. La mayoría de los errores son leves y no afectan significativamente la comprensión del texto". De nuevo en las pruebas 12 y 14, también al principio y exactamente con las mismas palabras en ambos casos: "Aquí está la corrección del texto del alumno de acuerdo con las escalas de calificación del DELE B1". Y otra vez encontramos este automatismo en la Prueba 13: "Corrección del texto del alumno de acuerdo con la escala de calificación del DELE B1", así como en este comentario final de la Prueba 9: "En general, la tarea de escritura tiene pocos errores y se entiende claramente el mensaje. La mayoría de los errores son leves y no afectan significativamente la comprensión del texto".

### 3.2. Realización de la tarea modelo del IC

En cuanto a la respuesta a la muestra de tarea 1 del DELE B1 colgada en la web del IC, solo podemos calificarla de excelente, si no perfecta, ya que se produce con un nivel de adecuación al género, coherencia, gramática y alcance, que, a nuestro modo de ver, llega e incluso supera el nivel B2. Por lo demás, resulta baladí afirmar que, para un nivel de DELE B1, estaríamos ante un APTO clamoroso.

Algunas muestras de la suficiencia con la que ChatGPT resuelve la tarea, que reflejan un gran dominio de la gramática y el vocabulario, las encontramos, por ejemplo, en: 1) uso del subjuntivo con verbos de sentimiento ("me alegra mucho que hayamos decidido") y de influencia ("te recomiendo que trates de mantenerla", "te sugiero que te tomes", "te pediría que cuides", "Espero que esta información te sea"); 2) variedad en el empleo de marcadores pra introducir subtemas ("Con respecto a", "en cuanto a", "respecto a"); 3) uso de estructuras subordinadas, de causa ("ya que", en dos ocasiones), comparación ("como si fuera"), condicional (""Si tienes alguna que otra pregunta"); 5) utilización de conectores ordenadores u organizadores de la información ("además", "por último"); 6) amplitud de recursos léxicos ("acceso", "estacionar", espacio", "disponible", "transporte público", "eficiente", "parada de autobús", "centro de la ciudad", "mantenerla", "estancia", "muebles", "electrodomésticos", "luces encendidas", "vecinos", etc.).

Por lo demás, el texto elaborado por ChatGPT resulta perfectamente adecuado al género y coherente (si exceptuamos el error de usar coma en lugar de dos puntos en el encabezado del correo), tanto local como globalmente. Tan solo podríamos achacarle el hecho de que supera con creces las 120 palabras (se va hasta las 185), problema que quizá podríamos resolver enfatizando de algún modo en las instrucciones que le damos que no se pase de las 125-130 palabras.

### 3.3. Creación de tareas

Con respecto a la creación de tareas, el resultado también cabe calificarlo como de positivo, en todas y cada una de sus partes: contextualización inicial, texto intermedio para la interacción e instrucciones finales. No obstante, vemos que, en el caso de la Prueba 2, por un lado, ChatGPT "ha olvidado" incluir el texto para la interacción con el candidato al DELE, y por otro lado, se equivoca al ofrecer un número de palabras por encima de lo requerido (entre 120 y 150 en lugar de entre 100 y 120); en consecuencia, esta tarea resulta fallida e inútil para el fin con el que, en teoría, ha sido creada.

En cambio, si modificamos las indicaciones que le damos al programa, señalando expresamente la necesidad de que se incluya el texto para la interacción, el resultado vuelve a ser de matrícula de honor, tal como podemos comprobar en la Prueba 3, donde el texto incluso aparece entrecomillado, si bien es cierto que el número de indicaciones (tan solo tres) se queda algo corto para lo que es este modelo de tarea. Aparte, en esta

ocasión el número de palabras sí se ajusta a lo requerido. En la prueba 4, en cambio, con siete indicaciones, la respuesta de ChatGPT resulta más analítica y ajustada al formato oficial del examen DELE B1 y, por tanto, de más ayuda y orientación para el candidato.

Un aspecto curioso, pero al mismo interesante, que observamos en las pruebas 3 y 4 es que la tarea creada no termina con las instrucciones para realizarla, como ocurre normalmente en este tipo de tareas DELE, sino con un elemento adicional de eventual ayuda al candidato: se le recuerda los criterios de calificación con los que será evaluado su texto. Además, en la Prueba 3, ChatGPT incluso muestra capacidad de imaginación o invención para asignar puntuaciones a cada escala y criterio, que muestra conocer: 15 puntos a cada uno de los cuatro criterios de la escala analítica (60 en total) y 20, 10 y 10 (40 en total) a cada uno de los tres criterios que crea *ad hoc* para la escala holística ("Dominio del tema y habilidad para comunicar ideas", "Nivel de desarrollo del contenido" y "Uso efectivo del lenguaje", respectivamente). Sin embargo, en la Prueba 4, el sistema de IA se muestra más escueto, ya que se limita a señalar de manera general que "Se valorará la adecuación al género, la coherencia, la corrección y el alcance del vocabulario utilizado."

### 3.4. Corrección de errores y comentarios

En relación con los resultados a la hora de corregir la tarea de nuestros alumnos, en general podemos señalar que el trabajo es muy bueno. Hemos probado expresamente, o ChatGPT nos ha ofrecido, varias modalidades de corrección, que pasamos a comentar.

Primero, con la indicación general de que corrija la tarea (Prueba 5), el asistente virtual ofrece directamente una versión corregida del texto del alumno, muy bien escrita, por cierto, tal como damos cuenta más abajo.

Segundo, con la indicación expresa de que añada explicaciones o marque los errores, el programa ofrece un apartado explicativo bajo el título "Corrección" (Prueba 6) o "Errores" (Prueba 7) después de cada párrafo, distinguiendo lo que no necesita corrección, de un lado, de las palabras (aunque en la prueba 6 se equivoca) que hay que utilizar en lugar de otras, justificándolo gramatical o pragmáticamente (pruebas 6 y 7). Curiosamente, ChatGPT señala más faltas tras indicarle que marque los errores (detecta ocho errores) que al decirle que añada explicaciones de corrección (detecta solo cuatro errores), además, como aclaración adicional, en el primer caso incluye al final una versión corregida del texto, por lo que parece que el uso de determinadas palabras en las instrucciones que le damos resulta clave.

Tercero, cuando le pedimos que corrija el texto clasificando los errores sin mayor detalle (Prueba 8), encontramos que los clasifica en: Ortografía, Puntuación, Gramática, Usos de las mayúsculas, Redundancia, Omisión de palabras. En cambio, al pedirle que corrija clasificando los errores de acuerdo con las escalas de calificación del DELE B1 en general (Prueba 9), encontramos que los divide en: Ortografía, Uso de mayúsculas y minúsculas, Puntuación, Gramática y estructura. Y cuando le pedimos (Prueba 10) que corrija más detalladamente clasificando los errores en 1) Adecuación al género, 2) Coherencia, 3) Gramática y 4) Vocabulario, es decir, indicándole los criterios de la escala analítica del IC (modificando Alcance por Vocabulario), la respuesta del sistema se ajusta a esta división.

Cuarto, cuando, previamente al copiado del texto del alumno, al programa se le ofrece todas las instrucciones de la tarea en cuestión y a continuación se le pide que corrija la respuesta del alumno de acuerdo con las escalas de calificación de la prueba de expresión e interacción escritas del DELE B1 (Prueba 11), ChatGPT da una respuesta de lo más que curiosa. Decimos esto porque, en lugar ofrecer una propuesta de corrección mediante una nueva versión del texto, en esta ocasión comienza con una carta de respuesta a Lucía –es decir, como si la firmante del texto fuera quien ha utilizado el cuadro de diálogo de ChatGPT–, donde le indica (erróneamente, por cierto) que el número de palabras no llega al mínimo exigido, por un lado, y que debe hacer algunas correcciones de gramática y ortografía, por otro. Además, el output concluye con una muestra adicional de interacción empática con Lucía: "Espero que esta retroalimentación sea útil. ¡Mucho éxito en tus estudios!".

Quinto, cuando se le da al sistema todas las instrucciones de la tarea en cuestión, previamente al copiado del texto del alumno, a continuación se le pide que corrija la respuesta de este de acuerdo con las escalas de calificación de la prueba de expresión e interacción escritas del DELE B1, y además se le indica expresamente los criterios y porcentajes generales de las dos escalas (Prueba 12), obtenemos una respuesta más detallada, tanto desde el punto de vista de la puntuación (con un baremo de 2/2 para cada uno de los criterios y 4/4 para la escala holística) como de las explicaciones de los errores. En cuanto a lo primero, al final incluso ofrece una puntuación total: 5/10. Adicionalmente, ChatGPT ofrece comentarios para cada uno de los criterios y escalas y señala varios errores gramaticales. Por lo demás, al final también ofrece una "versión mejorada y corregida del texto" que, incluso, autopuntúa con 8/10.

Sexto, cuando se le ofrece de nuevo todas las instrucciones de la tarea en cuestión, previamente al copiado del texto del alumno, a continuación se le pide que corrija la respuesta este de acuerdo con las escalas de calificación de la prueba de expresión e interacción escritas del DELE B1, pero esta vez indicando tanto los criterios y los porcentajes generales de las dos escalas de calificación, con el baremo 0 a 4 que va del No Apto

al Apto (Prueba 13), lo que nos devuelve el sistema es, primero, una versión del texto corregida; luego, una puntuación de Apto (4), tanto para los criterios de la escala analítica como para la escala holística; y al final, un comentario general de la tarea, que se remata con una puntuación total de 9 sobre 10.

Por último, cuando volvemos escribirle todas las instrucciones de la tarea en cuestión, previamente al copiado del texto del alumno, a continuación se le pide que corrija la respuesta de acuerdo con las escalas de calificación de la prueba de expresión e interacción escritas del DELE B1, añadiendo ahora los criterios concretos y el número máximo de puntos de las dos escalas –60 y 40 para las escalas, hasta 15 para cada criterio analítico– (Prueba 14), el output de ChatGPT se estructura de la siguiente manera: 1) escala analítica, 1.1.) criterio con puntuación parcial (0/15), 1.1.1.) comentarios, 2) escala holística, 2.1.) comentario con puntuación total (26/40). En esta ocasión sin dar una versión alternativa corregida.

Por lo demás, destacaremos que, de las pruebas realizadas, se deduce que ChatGPT es muy hábil a la hora de redactar en español, pero no perfecto. Aun con toda su capacidad para enfrentarse al reto de contestar, crear y corregir tareas de DELE B1, comete algunos errores, sobre todo a la hora de explicarlos o comentarlos. Como botón de muestra, he aquí una lista no exhaustiva:

1. uso de la coma (,) en lugar de los dos puntos (:) tras el tratamiento que encabeza una carta (Prueba 1), si bien lo hace correctamente en otros casos (Prueba 5);
2. repetición de algunas estructuras, como “en cuanto a” (Prueba 1), si puede considerarse esto un error (¿de estilo?);
3. ausencia de coma (,) en el encabezamiento “Hola amigo” (pruebas 3 y 13), pero no así en otros casos, como en la Prueba 7, donde, además, ChatGPT indica expresamente el error: “Se debe colocar una coma después de “Hola” para separar la interjección del nombre”;
4. dar como correcto “nos quedamos” en lugar de “quedamos” (pruebas 6, 7 y 8);
5. advertir que “organizárle” debe ser corregido a “organizarle con z” (prueba 7), que es correcto (Prueba 8);
6. proponer bastantes correcciones inapropiadas (pruebas 6, 7, 9);
7. falta de congruencia en casos como “a las 3 por la tarde” (Prueba 8) frente a “a las tres de la tarde” (Prueba 11);
8. mantener bastantes errores sin corregir, como “para tu amiga Elga, por supuesto, soy bienvenida”, entre otros, en la Prueba 8;
9. empleo de coma (,) en lugar de punto (.) al final de “Querido Javier” en la Prueba 12.

Antes de terminar este apartado queremos también resaltar, no sin gran sorpresa y admiración, la gran calidad de las diferentes versiones de la tarea 1 ofrecidas por ChatGPT, no solo por su adecuación al género, coherencia estructural y de sentido, gramática y alcance, sino también por la variedad de recursos lingüísticos que muestra en sus respuestas a la hora de plasmar una y otra vez los mismos significados. En este sentido, los textos que se nos da en las pruebas 11, 12 y 13 resultan sin lugar a duda modélicos.

## 4. CONCLUSIONES

Al principio nos preguntábamos qué ventajas puede ofrecer la IA, y más concretamente herramientas como ChatGPT, para la enseñanza y el aprendizaje del E/LE. Pues bien, tras las pruebas que hemos efectuado podemos confirmar que esta tecnología constituye un avance innovador que otorga enormes posibilidades y ventajas tanto para el profesorado como para el alumnado, no solo de cara a la práctica y mejora de la expresión escrita, sino también de otras destrezas, gracias, sobre todo, a su habilidad para interactuar con su usuario.

Así, hemos podido comprobar la capacidad, e incluso el “talento innato”, de ChatGPT a la hora de comprender, responder, crear y corregir ejercicios de escritura, en particular, tareas y textos de acuerdo con los estándares del DELE. La eficacia alcanzada por este programa frente a instrucciones complejas, que incluso pueden contener rúbricas para la calificación de las tareas, supone un gran avance tecnológico y una gran ayuda para el quehacer del profesor. De cara a su aplicación pedagógica en el aula, abre la posibilidad de trabajar en grupos de Composición comparando modelos “inteligentes” de redacciones o diferentes formas de corregir un mismo texto, por ejemplo. También ofrece al alumno la opción de realizar la autocorrección con retroalimentación de su “profesor virtual”, no ya a través de una corrección automática del texto sin más, sino a través de instrucciones concretas sobre determinados aspectos (gramática, léxico, coherencia, adecuación al género, etc.) o mediante rúbricas diseñadas *ad hoc* por su profesor o incluso por el propio estudiante. Además, creemos que las versiones alternativas correctas de textos originalmente anómalos, así como los comentarios y las explicaciones de los errores, que incluso podrían solicitarse en la lengua nativa del alumno, pueden contribuir a que este comprenda mejor la gramática y mejore su competencia discursiva en general y, por ende, su preparación para afrontar el DELE.

No obstante, de nuestros experimentos también se concluye que la IA, al menos la que sostiene a asistentes virtuales como ChatGPT, hoy en día todavía no ha alcanzado un grado de idoneidad suficiente como para confiar plenamente en ella. Al contrario, el sistema comete bastantes errores de apreciación ante determinadas instrucciones pensadas para corregir textos, lo cual implica que la forma en que se estas se dan y su refinamiento resultan clave para el éxito a la hora de manejar esta nueva tecnología. Dado que programas como ChatGPT tienen sus problemas y sus trucos de utilización, nos parece necesario trasladarlos a los estudiantes a fin de que hagan un buen uso de la retroalimentación que aquellos les proporcionan.

En definitiva, el uso de asistentes virtuales como ChatGPT sin duda facilita mucho el trabajo del profesorado y amplía enormemente las opciones pedagógicas de enseñanza de lenguas. Sin embargo, para sacarle el adecuado provecho, este tipo de herramientas debe emplearse con cuidado, solo como apoyo a la instrucción y siempre bajo la supervisión del profesorado, pero en ningún caso para sustituir por completo la enseñanza humana. En este sentido, animamos a los profesores involucrados en la enseñanza de la escritura en diferentes niveles a incorporar esta tecnología en los cursos de expresión escrita o de preparación de los DELE.

## 5. REFERENCIAS


Alam, A. (2021). Possibilities and Apprehensions in the Landscape of Artificial Intelligence in Education. *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)* (pp. 1-8). G. H. Raisoni College of Engineering, Nagpur, India. doi: 10.1109/ICCICA52458.2021.9697272.

Blanco Pena, J. M. (2024). Innovación en entornos educativos sinohablantes: una investigación-acción en un aula de Composición Escrita en Taiwán. *Círculo de Lingüística Aplicada a la Comunicación* (en prensa).

Chen, L., Chen, P. y Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access*, *8*, 75264-75278. doi: https://doi.org/10.1109/ACCESS.2020.2988510

Cruz Piñol, M. (2014). Veinte años de tecnologías y ELE. Reflexiones en torno a la enseñanza del español como lengua extranjera en la era de internet. *marcoELE. Revista de Didáctica Español Lengua Extranjera*, *19*.

Divekar, R. R., Drozdal, J., Chabot, S., Zhou, Y., Su, H., Chen, Y., Zhu, H., Hendler, J. A., y Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: design and evaluation. *Computer Assisted Language Learning*, *35*(9), pp. 2332-2360. https://doi.org/10.1080/09588221.2021.1879162

Escobar Hernández, J. C. (2021). La Inteligencia Artificial y la Enseñanza de lenguas: una aproximación al tema. *Decires*, *21*(25), pp. 29-44. https://doi.org/10.22201/cepe.14059134e.2021.21.25.3

Fan, O. y Jiao, P. (2021). Artificial intelligence in education: The three paradigms, *Computers and Education: Artificial Intelligence*, *2*, 100020. https://doi.org/10.1016/j.caeai.2021.100020

Huang, J., Saleh, S., y Liu, Y. (2021). A Review on Artificial Intelligence in Education. *Academic Journal of Interdisciplinary Studies*,10(206), 10.36941/ajis-2021-0077.

Huang, X., Zou, D., Cheng, G., Chen, X. y Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, *26*(1), pp. 112-131. https://www.jstor.org/stable/48707971

Jeon, J., Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Educ Inf Technol*, *28*, pp. 15873-15892. https://doi.org/10.1007/s10639-023-11834-1

Jeon, J., Lee, S. y Choe, H. (2023). Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers & Education*, *206*, 104898. https://doi.org/10.1016/j.compedu.2023.104898.

Kannan, J. y Munday, P. (2018). New trends in second language learning and teaching through the lens of ICT, networked learning, and artificial intelligence. En C. Fernández Juncal y N. Hernández Muñoz (Eds.), *Vías de transformación en la enseñanza de lenguas con mediación tecnológica* (*Círculo de Lingüística Aplicada a la Comunicación, 76*, pp. 13-30). http://dx.doi.org/10.5209/CLAC.62495

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J. y Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274, https://doi.org/10.1016/j.lindif.2023.102274.

Kuure, L., Molin-Juustila, T., Keisanen, T., Riekki, M., Iivari, N. y Kinnula, M. (2016). Switching perspectives: from a language teacher to a designer of language learning with new technologies. *Computer Assisted Language Learning*, *29*(5), pp. 925-941. https://doi.org/10.1080/09588221.2015.1068815

López, C. y Heredia, Y. (2017). *Marco de referencia para la evaluación de proyectos de innovación educativa - Guía de Aplicación*. Tecnológico de Monterrey. http://escalai.com/que_escalai/guia_app/

Muñoz-Basols, J., Gutiérrez, M. F. y Cerezo, L. (Eds.). (2024). *La enseñanza del español mediada por tecnología: de la justicia social a la Inteligencia Artificial (IA)*. Taylor & Francis.

Murillo, A. (3 de octubre de 2017). *¿Qué es innovación educativa?* Instituto para el futuro de la educación. https://observatorio.tec.mx/edu-news/innovacion-educativa

Reinhardt, J. (2022). Everyday technology-mediatized language learning. New opportunities and challenges. En C. Lütge (Ed.), *Foreign Language Learning in the Digital Age* (pp. 1-12). Routledge.

Roll, I. y Wylie, R. (2016). Evolution and Revolution in Artificial Intelligence in Education. *Int J Artif Intell Educ*, *26*, pp. 582-599. https://doi.org/10.1007/s40593-016-0110-3

Volpicelli, F. (2018). Las nuevas tecnologías en la enseñanza del español como lengua extranjera. En E. Cutillas Orgilés (Coord.), *Convergencia y transversalidad en humanidades. Actas de las VII Jornadas de Investigación de la Facultad de Filosofía y Letras de la Universidad de Alicante* (pp. 135-140). Universidad de Alicante.