

UN ESTUDIO SOBRE EL USO DE CHATGPT PARA RESOLVER TAREAS ESCRITAS EN LA CLASE DE ELE

This study examines how six B1-level adult learners of Spanish use ChatGPT to complete a monologic task. The results indicate that while students generally follow the structure suggested by GPT, they also adapt its recommendations to suit both the context and their language proficiency. Additionally, individual variation in responses is observed, reflecting students' personal engagement with the communicative situation. The findings suggest that the six students critically utilize AI tools in foreign language learning, highlighting the importance of supporting and encouraging this habit in educational settings.

● Nicola Brocca
| University
of Innsbruck
Rebeca Iniesta
Jiménez
| University
of Innsbruck

1. Introducción

Utilizar modelos grandes de lenguaje (LLMs, *large language models* en inglés) para escribir en una lengua extranjera es una práctica cada vez más habitual en la vida cotidiana (por ejemplo, para corregir un texto breve en una lengua extranjera). Las opiniones de los expertos en la enseñanza de lenguas extranjeras ante esta evolución son diversas: si bien los profesores consideran útiles estas herramientas, les preocupa que los alumnos confíen excesivamente en ellas (Ohashi, 2022; Berthele & Udry, 2023b) y afirman que los aprendientes no pueden desarrollar su competencia comunicativa si delegan todo el trabajo en la máquina y no realizan ellos mismos un esfuerzo cognitivo (Berthele & Udry, 2023a). Otros son de la opinión de que los LLMs no se pueden ignorar (Poole & Polio, 2023) y los profesores deberían integrarlos en el aula de idioma extranjero para ayudar a los alumnos a realizar tareas comunicativas auténticas.

Este estudio de caso investiga cómo estudiantes universitarios de español como lengua extranjera de nivel B1 utilizan un LLM para resolver tareas comunicativas y cómo esto influye en sus producciones escritas desde el punto de vista de la cortesía lingüística. En concreto, se examinan los resultados de 6 estudiantes que han producido dos tareas monológicas (petición y cancelación) utilizando GPT-4 (OpenAI, 2024). Dada su naturaleza exploratoria, la presente investigación se basa en una muestra pequeña de datos y no pretende ofrecer resultados generalizables. Más bien, representa un primer paso hacia una mejor comprensión de la enseñanza del español como segunda lengua utilizando IA, así como una posible fuente de sugerencias para los profesionales.

Los estudios científicos hasta la fecha se centran predominantemente en el inglés, descuidando el uso de los LLMs para la enseñanza de otras lenguas como el español, con pocas excepciones (p. ej., Fredholm, 2019). Algunas propuestas

didácticas para la clase de ELE plantean actividades para utilizar la IA en un entorno de enseñanza, siguiendo un protocolo específico para su uso (Arriagada, 2023; Hagenhof y Leienbach, 2020). En nuestro estudio examinamos específicamente cómo los estudiantes utilizan los LLM de forma espontánea cuando deben completar una tarea.

Por otra parte, este estudio explora el impacto del uso de los LLM sobre la cortesía lingüística. Usaremos el término *cortesía lingüística* para referirnos a la cortesía incorporada en el discurso que puede plasmarse en el uso de términos específicos de tratamiento, honoríficos, expresiones formulaicas convencionales (“gracias”, “perdón”) y varios recursos lingüísticos, como los empleados para mitigar la fuerza ilocutiva directa de una petición (p. ej. el uso de la expresión “por favor” o el uso del condicional) o para reducir los efectos negativos de un rechazo (Félix-Bresdefer, 2024).

Este estudio plantea la siguiente pregunta de investigación: (P) ¿Cómo realizan los estudiantes una tarea monológica desde el punto de vista de la cortesía lingüística cuando utilizan un LLM?

De manera más precisa, esta cuestión se abordará respondiendo a las siguientes preguntas:

P1: ¿Cómo utilizan los estudiantes GPT para producir sus textos?

P2: ¿Qué sugerencias de GPT incorporan los estudiantes en sus textos finales?

2. Revisión de la literatura

En el debate actual sobre la integración de herramientas basadas en inteligencia artificial en la enseñanza de idiomas, a menudo se asume que los estudiantes podrían usar estas herramientas de manera acrítica, lo cual podría tener efectos adversos en el proceso de adquisición del idioma (Berthele & Udry, 2023a). Muchos estudios, por lo tanto, abogan por realizar actividades que fomenten la interacción activa con los resultados generados por la IA —como releer, reinterpretar o modificar las sugerencias de la IA— reinvirtiendo así estas herramientas como recursos de apoyo en el aprendizaje. Por

ejemplo, Hagenhof y Leienbach (2020) proponen un ejercicio práctico en el que los estudiantes comparan traducciones producidas por dos herramientas de traducción basadas en IA y eligen la opción más precisa o adecuada al contexto. Arriagada (2023) destaca además el valor de las aplicaciones de traducción de voz para fomentar la conciencia pragmática intercultural en la clase de ELE, promoviendo el uso de herramientas de IA bajo un protocolo específico. Aunque en estos trabajos las interacciones de los estudiantes con LLMs ocurren en contextos espontáneos y no estructurados, los estudios ofrecen directrices muy específicas para los estudiantes.

Además, algunos estudios sugieren que la producción de formas de cortesía en el lenguaje por parte de los estudiantes no siempre está directamente influenciada por el input o la instrucción explícita. Por ejemplo, Siegal (1996) observó que algunas mujeres occidentales, a pesar de comprender las normas de cortesía en japonés, optaron por no ceñirse completamente a estas convenciones por no ajustarse a su identidad. Asimismo, los hablantes nativos pueden experimentar incomodidad cuando los hablantes no nativos se ajustan demasiado a las normas nativas, percibiéndolo como algo excesivamente formal o incluso pretencioso (Blum-Kulka, 1991: 269).

Basándonos en este contexto, planteamos la hipótesis de que, al redactar mensajes personales, los estudiantes pueden optar activamente por no replicar de forma literal las sugerencias generadas por la IA, lo que indica una interacción más compleja entre la asistencia de la IA y la autonomía del estudiante en el uso del idioma.

3. Metodología

3.1 Participantes

En este estudio participaron 6 personas matriculadas en un curso de español de nivel intermedio (B1) en una universidad en Austria. Todos los participantes eran hablantes nativos de alemán o poseían un dominio del idioma equiparable al de estos. Las edades de los participantes oscilaban entre los 21 y los 31 años. En las semanas previas a la realización de las tareas asignadas, los participantes recibieron clases sobre cómo expresar peticiones corteses utilizando el condicional,



Nicola Brocca is a post-doctoral fellow at the University of Innsbruck. His research interests include pragmatics in the context of speech-acts, TBLT and LLMs. nicola.brocca@uibk.ac.at



Rebeca Iniesta Jiménez is a doctoral candidate at the University of Innsbruck. Her research interests include the use of AI in high school Spanish classes. rebeca.iniesta-jimenez@uibk.ac.at

el uso del subjuntivo para expresar deseos y la redacción de mensajes breves. También recibieron input de saludos y expresiones formulaicas convencionales (“gracias”, “disculpa”) a través del libro de texto y de la práctica en clase.

3.2 Recolección de datos y análisis

A los participantes se les pidió que resolvieran en línea las siguientes tareas (Brocca 2021; Cortés Velásquez, Nuzzo 2022):

- Una cancelación (informal)

Un/a amigo/a te ha invitado a cenar esta noche y has aceptado la invitación. En el último momento, tienes que cancelar la cita porque tienes otro compromiso. Escribe un mensaje por WhatsApp a tu amigo/a comunicando tu ausencia.

- Una petición (formal)

Hoy tienes que enviar tus deberes al/a profesor/a de español. No has tenido tiempo porque has tenido que prepararte para otro examen. Escribe un e-mail a tu profesor/a de español solicitando más tiempo.

Los participantes realizaron las tareas en un formulario en línea de Google. En las instrucciones de la tarea se indicaba a los estudiantes que utilizaran GPT para prepararla. Para monitorizar este proceso, en el formulario se pedía a los estudiantes que copiaran y pegaran el *prompt* que habían insertado en GPT, así como la respuesta de la máquina. En la página siguiente del cuestionario se pedía a los estudiantes que escribieran sus mensajes finales.

Para el análisis de las estrategias de apropiación de las sugerencias de GPT, nos basamos en un cotejo entre el *prompt*, la respuesta de GPT y el mensaje final. Para el análisis de las tareas, utilizamos un

análisis de los actos de habla (Trosborg, 1995); este análisis permite examinar los movimientos de apoyo, como la justificación de la petición o de la cancelación, los movimientos de reparación y los marcadores de cortesía.

4. Resultados

Para responder a la P1, observaremos cómo formularon los estudiantes el *prompt* en GPT. Los estudiantes utilizaron dos estrategias distintas: i) pedir a GPT que resolviera la tarea desde cero y ii) utilizar GPT para mejorar el borrador del mensaje formulado por ellos mismos. La mayoría (4 de 6) utilizó el enunciado de la tarea como *prompt*, insertándolo en GPT. Solo dos estudiantes utilizaron GPT para mejorar el borrador de su mensaje.

Para responder a la P2, analizaremos ahora cómo los estudiantes pasaron de la propuesta de GPT a la versión final. A la hora de escribir el mensaje final (producción), los alumnos emplearon dos estrategias diferentes: iii) adaptar la respuesta de GPT o iv) ignorar prácticamente la respuesta de GPT. Al escribir el mensaje final, 9 mensajes de GPT fueron adaptados y tres sirvieron únicamente para tomar prestados algunos elementos léxicos y fraseológicos.

Las adaptaciones consistieron básicamente en ajustes menores al acto principal (petición o cancelación) siguiendo la estrategia de GPT, e incluyeron frecuentemente movimientos de apoyo sugeridos por GPT. Los estudiantes tendieron a adaptar elementos léxicos y morfosintácticos para ajustarlos a su nivel de competencia. En el siguiente ejemplo, S2 expresó la petición en indicativo en lugar de en condicional, como sugería GPT, mientras que S4 reemplazó palabras menos comunes por vocabulario más

básico. S4 también añadió un marcador de cortesía (“por favor”).

S2 (GPT): “¿Podría concederme unos días más para entregarlos?”

S2 (Estudiante): “¿Me *puede* dar unos días más para entregarlos?”

S4 (GPT): “Necesito unos días adicionales para completar las tareas adecuadamente.”

S4 (Estudiante): “Necesito dos días *más, por favor*, para completar los deberes.”

En algunos casos, las adaptaciones sirvieron para hacer el texto más adecuado al contexto, como añadir el nombre real del destinatario en el saludo o elegir un saludo más informal para el profesor, como «Hola», en lugar de «Estimado».

En raras ocasiones, algunos movimientos de apoyo sugeridos por GPT fueron incorporados en las versiones finales de los estudiantes, típicamente expresiones de gratitud o de arrepentimiento. Por ejemplo:

S1 (GPT): “¡Me puede hacer un gran favor!”

S4 (GPT): “Agradezco de antemano su comprensión y consideración.”

S2 (GPT): “Lamento mucho avisarte en el último momento.”

Aunque la adaptación es la estrategia más común, dos estudiantes personalizaron más sus textos, tomando solo algunos elementos de GPT. Esta estrategia es evidente comparando la siguiente sugerencia de GPT y el mensaje final del estudiante:

S3 (GPT): “Lo siento mucho, pero me ha surgido un compromiso de última hora y no podré ir a la cena de esta noche. 😞 ¿Podemos quedar otro día? Mil disculpas por avisar tan tarde.”

Tabla 1

Distribución de estrategias utilizadas para resolver la tarea utilizando GPT

		Preparación	
		Pedir a GPT resolver la tarea desde cero	Introducir un borrador y pedir a GPT que lo corrija y mejore
Producción	Copiar la respuesta de GPT	–	–
	Adaptar la respuesta de GPT	S2, S4, S5 (petición)	S1, S6
	Tomar elementos de la respuesta de GPT	S3, S4 (cancelación)	--

S3 (Estudiante): “Hola, Seb. Me sabe fatal, pero tengo que cancelar la cena de esta noche contigo. ¿Piensas que podemos quedar otro día? Disculpas por avisar más tarde. 🙄”

En el mensaje de S3, se añade un saludo y el nombre del destinatario, mientras que la expresión inicial de disculpa, “lo siento mucho,” se reemplaza por la expresión más idiomática “me sabe fatal,” que intensifica la disculpa. Aunque se omite la explicación sugerida por GPT, la oferta de reprogramar (“¿Podemos quedar otro día?”) se suaviza mediante una estructura incrustada (“¿Piensas que...?”). Se añade también un emoticono de cara de mono, que introduce un tono de humor auto-crítico y agrega un toque personal más allá de la carita triste sugerida por GPT.

La tabla 1 muestra la distribución de los estudiantes según las estrategias que utilizaron.

5. Discusión

El análisis indica que los textos finales de los estudiantes están bastante influenciados por las sugerencias de GPT. Sin embargo, tanto las estrategias de adaptación como de selección indican que los estudiantes interactúan críticamente con las sugerencias de GPT, prefiriendo expresiones que pueden producir de forma independiente dentro de su nivel de competencia lingüística. Además, los estudiantes en ocasiones excluyen expresiones que no se alinean con sus preferencias de cortesía, introducen elementos como el nombre del destinatario o utilizan el saludo “Hola” en lugar del más formal “Estimado”. Aunque la estructura de los mensajes es en su mayoría similar, las adaptaciones tienden a reforzar una mayor informalidad entre los estudiantes y el destinatario. Otro hallazgo es la variabilidad personal en las estrategias utilizadas por los estudiantes: por ejemplo, S3 mostró una creación de tareas divergente y más autónoma, mientras que los otros estudiantes siguieron de manera más sistemática las sugerencias de GPT.

Podemos proponer dos hipótesis para explicar este fenómeno: copiar un texto se percibe como algo éticamente incorrecto, especialmente en un contexto didáctico. Otra hipótesis está relacionada con el tipo de tarea: escribir un mensaje personal.

Los estudiantes creen que aceptar las sugerencias de GPT hace que el mensaje no sea auténtico o creíble, por lo que prefieren ceñirse a la versión que escribirían espontáneamente. Esto podría explicar el rechazo de los estudiantes a las fórmulas de cortesía sugeridas por GPT, ya que, aunque correctas, las consideran alejadas de su forma de expresarse (Siegal, 1996).

Bibliografía

- Arriagada, M.** (2023). Pragmática práctica para una clase intercultural II/II: Límites y posibilidades de los traductores automáticos de voz para el aprendizaje de los actos de habla en español-alemán. *Hispanorama*, 180(2), 8-11.
- Berthele, R. & Udry I.** (2023a). Welche fremdsprachlichen Lernziele der Schulen können mit digitalen Tools aber ohne Lernen erreicht werden? Simply playing the ostrich will not help in the long run. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 28.1. doi: 10.48694/zif.3519 (recuperado el 8 de agosto de 2024).
- Berthele, R. & Udry I.** (2023b). Digitale Übersetzungsprogramme und Online Wörterbücher im Fremdsprachenunterricht: Erkenntnisse aus einer Umfrage bei Lehrpersonen und Lernenden der obligatorischen, post-obligatorischen und tertiären Bildung. *Linguistik Online*, 120.2, 145-167. doi: 10.13092/lo.120.9720.
- Blum-Kulka, S.** (1991). Interlanguage Pragmatics: The Case of Requests. In R. Phillipson et al. (eds.), *Foreign/second Language Pedagogy Research*. Clevedon: Multilingual Matters, 255-72.
- Brocca, N.** (2021). LADDER: Un corpus di scrittura digitali per l'insegnamento della pragmatica in L2. Un esempio di analisi di disdette in WhatsApp. *Italiano LinguaDue*, 13(1), 241-259.
- Félix-Brasdefer, J. C.** (2024). Politeness 1. *Pragmatics at Indiana University*. <https://pragmatics.indiana.edu/politeness/index.html> (recuperado el 6 de agosto de 2024).
- Fredholm, K.** (2019). Effects of Google translate on lexical diversity: vocabulary development among learners of Spanish as a foreign language. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, 13(26), 98-117.
- Hagenhoff, T. & Leienbach, S.** (2022). Übersetzungstools im Spanischunterricht - ein Plädoyer für einen kritischen Umgang. *Hispanorama*, 72-82.
- Ohashi, L.** (2022). The use of machine translation in L2 education: Japanese university teachers' views and practices. En: B. Arnbjörnsdóttir, B. Bédi, L. Bradley, K. Friðriksdóttir, H. Garðarsdóttir, S. Thouèsny, & M. J. Whelpton (eds.): *Intelligent CALL, granular systems and learner data: short papers from EUROCALL 2022*, 308-314.
- OpenAI.** 2024. ChatGPT (Version GPT-4o) [Computer software] <https://openai.com/> (recuperado el 6 de agosto de 2024).
- Poole, F. J. y Polio. C.** (2024). From sci-fi to the classroom. Implications of AI in task-based writing. *TASK Vol. 3 (2)*, 243-272. <https://doi.org/10.1075/task.00022.poo>.
- Siegal, M.** (1996). The Role of Learner Subjectivity in Second Language Sociolinguistic Competence: Western Women Learning Japanese. *Applied Linguistics*, 17(3), 356-382. <https://doi.org/10.1093/applin/17.3.356>
- Trosborg, A.** (1995). *Interlanguage pragmatics. Requests, Complaints and Apologies*. Berlin: Mouton de Gruyter.