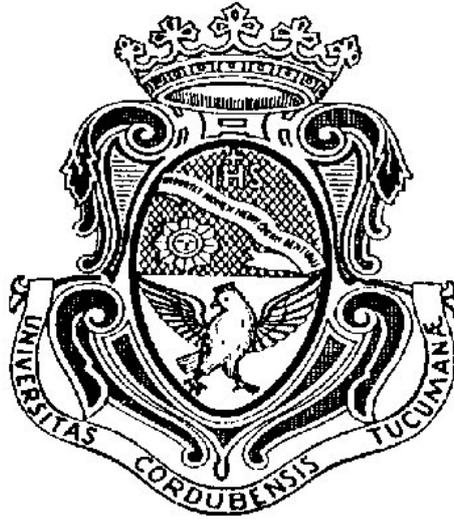


UNIVERSIDAD NACIONAL DE CÓRDOBA
FACULTAD DE LENGUAS



Licenciatura en Español Lengua Materna y Lengua Extranjera

Dirección: María Fernanda Freytes
Codirección: Laura Cecilia Díaz Dávila

Proyecto de investigación: Integración de *Large Language Models* en el proceso de producción de textos con fines educativos de español como lengua extranjera.

Alumnas: Fernández Iris
Gómez Sofía
Córdoba, 2023

Planteamiento del problema

En el contexto de la enseñanza de idiomas, una de las tareas fundamentales del docente es seleccionar cuidadosamente el material y los textos para potenciar las habilidades lingüísticas de sus alumnos. Sin embargo, esta tarea se vuelve compleja debido a la diversidad de factores que intervienen en ella. Actualmente, el mundo está marcado por una creciente digitalización y por el avance acelerado de la inteligencia artificial (IA) lo que vislumbra, sin dudas, un gran potencial para mejorar la educación a través de la personalización del aprendizaje (Ocaña-Fernández *et al.*, 2019). En este contexto, los *Large Language Models*, modelos de lenguaje basado en inteligencia artificial, emergen como una herramienta novedosa.

Como futuros docentes, enfrentamos el desafío de integrar los avances tecnológicos en nuestras prácticas pedagógicas para lograr clases más reveladoras y atractivas para los estudiantes.

Al centrarnos en el proceso de producción de textos con *Large Language Models* es imprescindible aclarar cómo funcionan estos modelos de lenguaje extensos. Estos, son un sistema de lenguaje generativo que es capaz de procesar grandes cantidades de textos y aprender a realizar tareas de procesamiento de lenguaje natural de manera muy efectiva. Posee una arquitectura de *transformadores* que se basa en el uso de atención y se ha entrenado usando el *aprendizaje por refuerzo* a partir de la retroalimentación humana. Esta técnica permite que el modelo se centre en diversas partes de la secuencia de entrada en diferentes momentos mientras la procesa y que aprenda a partir de su propia experiencia, interaccionando con el entorno hasta dar con el comportamiento ideal (Ortiz, 2023). Esto crea modelos como el *Chat GPT* de la empresa OpenIA, Perplexity IA fundada en 2022, Copilot IA de Microsoft, entre otros. En el contexto de esta investigación se ha optado por utilizar el *Chat GPT* de la empresa OpenIA, ya que ha ganado gran popularidad en los últimos meses. Sin embargo, es necesario mencionar que se alienta a la investigación y exploración continua de otras IA para su integración dentro de cualquier campo. Por ello, los precedentes escogidos se basan en el modelo de *chat GPT*.

En primer lugar, el artículo publicado de Jamal Kaid Mohammed Ali *et al.* (2023) aborda una investigación de tipo cualitativa que da cuenta, a través de un cuestionario en línea, que el *Chat GPT* motiva a los estudiantes a desarrollar habilidades de lectura y escritura. Los hallazgos se organizan en dos partes secuenciales: por un lado, aborda los

impactos del Chat GPT en la motivación de los estudiantes y, por otro, las diferencias estadísticas en las respuestas de los participantes. En ambas partes los resultados han sido efectivos, es decir, la mayoría de educadores coinciden en el *efecto positivo* del uso del Chat GPT en la motivación y respuestas de los estudiantes.

Un trabajo similar es el de Leunard *et al.* (2023) donde se investiga el uso de la aplicación *Chat GPT* en la enseñanza del árabe en instituciones terciarias. Los resultados muestran una respuesta positiva de los estudiantes, quienes experimentaron mejoras en sus habilidades de lectura, motivación y facilidad para completar las tareas.

Otro trabajo relevante que hemos utilizado para respaldar nuestro estudio es el de Wilson Cheong Hing Hong (2023). En su trabajo expone los mecanismos, las funciones y concepciones que se tienen acerca del *Chat GPT* en relación a su uso y aplicación dentro del ámbito educativo considerando tanto sus beneficios como sus problemas y riesgos a la hora de implementarlo en la actividad escolar. Se argumenta que el *Chat GPT* ofrece importantes oportunidades para que los profesores y las instituciones de educación mejoren la enseñanza y evaluación de segundas lenguas.

Por último, como antecedente más directo, el estudio de Ariyaratne *et al.* (2023) se centra en la evaluación de la precisión y calidad de los artículos académicos generados por *Chat GPT*. Para ello, compararon estos textos generados por *Chat GPT* con artículos escritos por autores humanos en el campo de la radiología. Los resultados mostraron que la mayoría de los artículos generados contenían información inexacta y referencias ficticias.

La investigación de estos autores nos alerta sobre la necesidad de ser cautelosos al utilizar textos generados por estas herramientas en un contexto educativo y nos ayuda a entender los posibles riesgos y limitaciones de utilizar los *Large Language Models* como recurso para la enseñanza. Resulta importante señalar la escasez de investigaciones en el idioma español, especialmente en la creación didáctica de textos. Es por esta razón que esta investigación buscará explorar el potencial de los *Large Language Models* (el Chat GPT) en la creación de textos aplicables en el aula de ELE (Español como Lengua Extranjera) tomando como base los hallazgos previos.

Motivadas por lo expuesto, nos enfocamos en explorar cómo los *Large Language Models* podrían contribuir a la producción de textos. Para lograr nuestros objetivos, nos planteamos las siguientes preguntas preliminares que guiarán nuestra investigación: ¿Cuáles son las dimensiones clave que deben contemplarse en la formulación de los

prompts (indicaciones o instrucciones) para el proceso de producción de textos narrativos del nivel B1 con la IA?; ¿En qué medida la IA logra completar eficientemente las dimensiones contempladas en los *prompts* y generar textos cooperativos en términos de Charolles?; ¿Qué ajustes pueden proponerse en función de las respuestas obtenidas para mejorar la calidad y pertinencia de los textos generados para su aplicabilidad en el aula?; ¿Qué recomendaciones se pueden derivar de los hallazgos para una implementación efectiva de la IA como recurso para crear textos adecuados a los propósitos de enseñanza? Por último, y no menos importante, preguntarnos por ¿Cuáles son las implicaciones éticas específicas que se deben contemplar al utilizar la IA para generar textos en el aula de ELE, considerando la autoría de los textos, la autenticidad de la información proporcionada a la IA y la labor docente?

Objetivos generales y específicos

Como objetivo general nos hemos propuesto indagar en la viabilidad de utilizar los *large language models* desde el lugar de docentes de la lengua española en la producción de textos para trabajar la habilidad de comprensión lectora en la enseñanza de ELE según los objetivos específicos que se proponen en la clase y el nivel B1. Los objetivos específicos que guiarán el trabajo son los siguientes:

1. Describir las dimensiones contempladas en la formulación de los *prompts* para el proceso de producción de textos narrativos con la IA, para trabajar en profundidad la competencia lectora en la enseñanza de ELE.
2. Analizar rigurosamente cuáles de estas dimensiones han sido completadas por la IA.
3. Proponer ajustes en función de las respuestas obtenidas por la IA.
4. Identificar las implicaciones éticas en el uso de IA en el ámbito educativo desde la labor docente de ELE.

Conjeturas

Este trabajo de investigación surge de dos premisas iniciales. Primero, la IA logrará, en términos de Charolles (1983), textos altamente *cooperativos* que le servirán al docente para los propósitos de la clase. Segundo, el nivel de precisión y eficacia de la IA aumentará de manera proporcional a la cantidad y calidad de la información facilitada a través de las interacciones con el sistema y, por ende, disminuirá la cantidad de errores.

Variables y su justificación

En el contexto de nuestra investigación es fundamental establecer un marco de análisis sólido para comprender y evaluar la calidad de los textos generados por esta inteligencia. Para lograrlo, hemos definido una serie de elementos esenciales que constituyen nuestro enfoque metodológico. Estas categorías se justifican por su relevancia para garantizar que los *prompts* generen textos narrativos ficcionales coherentes con los objetivos de la clase y que respondan a las necesidades y preferencias de los estudiantes de nivel B1 en la enseñanza de ELE. Las variables abarcan desde aspectos lingüísticos hasta elementos de contenido y coherencia. Estas son:

- a. *Precisión lingüística*: evalúa la corrección gramatical y léxica de los textos generados; la coherencia en la estructura de las oraciones.
- b. *Relevancia temática*: refiere a la relación entre el texto generado y el prompt proporcionado, se evalúa la capacidad de la IA para abordar adecuadamente el tema propuesto.
- c. *El nivel de complejidad*: se centra en la adaptación del nivel de lenguaje y complejidad del texto al nivel de los estudiantes B1, donde se apunta a que el texto sea desafiante pero comprensible.
- d. *Adecuación cultural*: calcula si el texto generado incluye expresiones y referencias culturales apropiadas para el contexto de la enseñanza de ELE, que puede garantizar una experiencia culturalmente apropiada de lectura.
- e. *Coherencia con objetivos*: se refiere a la alineación del texto generado con los propósitos educativos y se evalúa si el contenido del texto contribuye directamente a los objetivos de la lección.
- f. *Quantum informativo*: esta categoría, según Viramonte de Avalos (1997), se relaciona con la cantidad y relevancia del contenido informativo en el texto. Esta variable busca garantizar que los textos sean valiosos y enriquecedores para los estudiantes.
- g. *La adecuación estructural del género narrativo*: determina si el texto sigue una estructura y estilo propio del género, se analiza la presencia y coherencia de elementos narrativos para asegurar una lectura pertinente.

En el siguiente cuadro sistematizamos las variables y los indicadores que tendremos en cuenta para determinar cuáles de estas fueron logradas en el texto generado por la IA.

Los índices correspondientes a cada variable se evalúan según la escala de valor del 1 (menos logrado) al 5 (altamente logrado).

Cuadro 1. Variables e indicadores de cumplimiento.

Variable	Indicadores	Valores o índices
Precisión lingüística	Correctitud gramatical y léxica.	Escala del 1 al 5
	Coherencia en la estructura de las oraciones.	
	Uso adecuado de tiempos verbales.	
Relevancia temática	Relación del texto con el prompt proporcionado.	Escala del 1 al 5
	Coherencia con el tema propuesto.	
Nivel de complejidad	Vocabulario y estructuras acordes al nivel B1.	Escala del 1 al 5
	Evita el uso de estructuras muy complejas o muy básicas.	
Adecuación cultural	Expresiones y referencias culturales adecuadas.	Escala del 1 al 5
	Sensibilidad y respeto cultural.	
Coherencia con los objetivos	Relación del texto con los objetivos de la clase.	Escala del 1 al 5
	Alineación con los propósitos de la actividad.	
Quantum Informativo	Contenido relevante y significativo.	Escala del 1 al 5
	Texto informativo y útil para los estudiantes.	
Adecuación a la estructura del género	Adecuada introducción del conflicto y	Escala del 1 al 5

narrativo	desafío. Desarrollo efectivo de la trama y los eventos. Caracterización de los personajes. Creación de un ambiente temporal y espacial. Coherencia con la perspectiva narrativa.	
------------------	--	--

Marco teórico

En consonancia con nuestra formulación del problema, concebimos la comprensión lectora desde una perspectiva contextualizada. De acuerdo con la perspectiva de Vigotsky, el aprendizaje se configura como un proceso social arraigado en el entorno cultural y social del individuo. Los elementos contextuales como las referencias culturales, los conceptos socio históricos y las particularidades de un entorno cultural son fundamentales para comprender y asimilar la información. Por lo que, a la hora de aprender una lengua, no solo nos centramos en el aprendizaje de aspectos lingüísticos sino también en los conceptos culturales significativos. Al considerar estos elementos contextuales en la generación de textos se fomentan ciertas habilidades cognitivas superiores como el razonamiento, la interpretación y la comprensión contextual, dado que el aprendizaje efectivo trata, por un lado, de dominar gramaticalmente un idioma y, por otro, comprender cómo se usa en diferentes contextos sociales y culturales (Vigotski, 1978, como se citó en Mota de Cabrera, C.; Villalobos, J. 2007).

Por esta razón, en este proyecto de investigación se concibe tanto la producción escrita como la habilidad de comprensión lectora como fenómenos intrínsecamente ligados a su entorno. Esta perspectiva reconoce que la capacidad de expresarse por escrito y comprender textos no puede separarse de las influencias contextuales que rodean al individuo. Así, al abordar tanto la producción escrita como la comprensión lectora desde esta perspectiva contextualizada, se considera fundamental explorar y comprender las complejas interconexiones entre el individuo y su entorno en el proceso de desarrollo de estas habilidades.

El texto narrativo, según como lo introduce Carrera (2010) y como piedra angular de este análisis, se caracteriza por contar una historia a través de una secuencia de eventos

relacionados. Estos eventos involucran a personajes que interactúan en un entorno temporal y espacial específico. En esencia, un texto narrativo se compone de elementos como la complicación, que introduce un conflicto y desafío en la historia, seguido de la resolución, que representa cómo se aborda este conflicto. Cada uno de estos eventos se sitúa en un marco contextual y esta colección de eventos y contextos se agrupa en episodios, que en conjunto conforman la trama de la historia. Finalmente, la trama se combina con la perspectiva del narrador ya que añade un matiz subjetivo a la historia y es lo que, en última instancia, da forma a la historia en sí. (Van Dijk, 1983, como se citó en Carrera; Vázquez, 2013). Para este trabajo de investigación es esencial la comprensión profunda del texto narrativo porque al analizar estos textos necesitamos considerar cómo se desarrolla la narrativa.

Además, esta comprensión del texto narrativo se relaciona directamente con el Marco Común Europeo de Referencias (MCER) para las lenguas: aprendizaje, enseñanza, evaluación Consejo de Europa (2020). Este, como se sabe bien, es un estándar internacional europeo que no solo define la competencia lingüística en una determinada lengua, sino que también establece parámetros, categorías, criterios y escalas que pueden utilizar los usuarios. En uno de sus capítulos se analiza las formas de desarrollar las capacidades necesarias para la comunicación y de facilitar el desarrollo de la comprensión lectora, por lo que trata en específico actividades comunicativas. Por tanto, en las actividades de comprensión lectora el sujeto como lector “recibe y procesa como información de entrada textos escritos producidos por uno o más autores” (p. 71). El MCER propone aspectos específicos de la comprensión lectora para cada nivel de dominio de la lengua. En el caso del nivel B1 el usuario de la lengua “lee textos sencillos sobre hechos concretos que tratan sobre temas relacionados con su especialidad con un nivel de comprensión satisfactorio” (p. 71). Resulta coherente la elección de este nivel, ya que en este los estudiantes corresponden a un nivel intermedio según el MCER y ya han superado el nivel básico y adquirido cierta competencia en el idioma que les permita comunicarse en situaciones relacionadas de su interés. Y por último, esto implica que los estudiantes deben ser capaces de analizar la estructura narrativa de los textos que leen y comprender cómo se desarrollan las historias.

El interés por investigar los textos generados por los *large language models* se enmarca dentro de lo que se denomina el método de *aprendizaje adaptativo* entendiendo

este como el sistema que se adapta al progreso, fortalezas y debilidades de los estudiantes. Morillo Lozano (2016) lo define como “el aprendizaje adaptativo nace para personificar el aprendizaje y ofrecer a los alumnos un recurso que les permita mediante una plataforma tecnológica aprender la materia sin dejar nada ni a nadie por el camino. [...] utiliza las nuevas tecnologías y las herramientas digitales para personalizar el proceso de enseñanza-aprendizaje y adaptar la propuesta de trabajo a las necesidades y características de los alumnos” (p. 11). Este enfoque no es solo una herramienta para el estudiante, sino también para el docente, ya que le permite realizar un seguimiento detallado del proceso de cada uno de ellos, complementando con tareas adicionales si así fuese necesario para un mejor avance y desarrollo personal. Asimismo, Marcin Frąckiewicz (2013) destaca que los educadores, a través de la IA, pueden crear entornos más eficientes y prácticos a los intereses de cada alumno. Se adapta el contenido educativo a las preferencias y niveles, necesidades individuales, estilos de los estudiantes con la finalidad de acompañar sus procesos de aprendizaje. Por otro lado, este proyecto de investigación se posiciona desde la perspectiva del *conectivismo*, una teoría de aprendizaje para la era digital. Si bien las teorías tradicionales como por ejemplo el conductismo o el constructivismo han sido fundamentales en la educación, el conectivismo reconoce que estas teorías tienen limitaciones en el entorno actual ya que la tecnología desempeña un papel fundamental en el proceso de aprendizaje. Según Siemens (2004), el conectivismo se basa en la idea de que los estudiantes deben seleccionar y filtrar la información entrante en un entorno en constante cambio.

En la creación de los *prompts* se han tomado diferentes categorías teóricas: pertinencia y adecuación, claridad y especificidad, enfoque comunicativo, enfoque cultural y vínculo con la experiencia y relevancia. Estos criterios se pueden descartar, mantener o, incluso, generar nuevos a medida que avance la investigación. Las categorías de pertinencia y adecuación, así como claridad y especificidad, se justifican a partir de la propuesta de R. Beaugrande y W. Dressler (1981), que plantea los siete rasgos de textualidad: cohesión, coherencia, intencionalidad, informatividad, aceptabilidad, situacionalidad e intertextualidad. Estas normas están estrechamente ligadas a la producción e interpretación de textos y son fundamentales para lograr una comprensión completa de los mismos. La categoría de pertinencia se vincula con la situacionalidad, la cual remite a las coordenadas espaciotemporales en que tiene lugar el texto. Al seleccionar los *prompts*, es esencial

asegurarse de que estén contextualizados y sean relevantes para el contexto en el que se utilizarán. La pertinencia garantiza que los textos generados sean apropiados y adecuados para la enseñanza de ELE en un nivel B1, tomando en cuenta intereses y necesidades específicas de los estudiantes. La adecuación, por su parte, se relaciona con la intencionalidad y aceptabilidad. La intencionalidad se refiere al objetivo propuesto por el emisor al crear un texto, en este caso, los docentes utilizan la IA como herramienta para generar textos narrativos. Los *prompts* deben estar alineados con los objetivos de la clase y la intención de mejorar la comprensión lectora. Por otro lado, la claridad y especificidad de los *prompts* se enlazan con las normas de coherencia y cohesión. Estos deben redactarse de manera clara y coherente para evitar ambigüedades y garantizar que los estudiantes puedan leer textos narrativos con una estructura lógica y cohesiva. En los años 80, el desarrollo del concepto de *competencia comunicativa* incorporó en las clases de lenguas aspectos vinculados a lo sociolingüístico y sociocultural “basado en tres ejes: a. reglas de comportamiento comunicativo, b. variedades sociolingüísticas, c. el sentido se construye en la interacción” (Areizaga, Gómez e Ibarra, 2005). El *enfoque intercultural* se vincula intrínsecamente a esto ya que, bien como definen estos autores, “aprender otra lengua es una cuestión esencialmente cultural” (p. 28). La categoría de vínculo con la experiencia y relevancia mantiene una estrecha relación con el enfoque intercultural. Desde la noción de *aprendizaje situado* y de su definición: “la educación es el producto de procesos conformados en actividad [...] Estos son el resultado de la relación dinámica que se establece entre quien aprende y el entorno sociocultural en el que ejerce su acción” (Sagastegui, 2004), entendemos que este aprendizaje basado en una situación específica y real de la vida cotidiana es clave para la adquisición de habilidades y competencias.

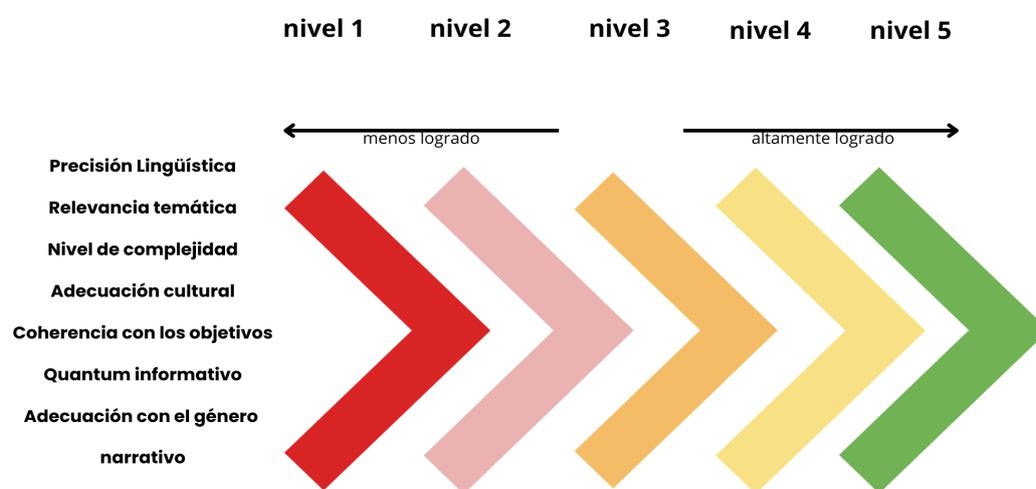
Metodología

En lo que respecta al tipo de estudio, se llevará a cabo un estudio de alcance exploratorio - descriptivo. Se trata de un estudio observacional en donde no se interviene ni manipula el fenómeno, es decir, se examina lo que ocurre en la realidad, en este caso, a partir del uso del *Chat GPT*.

Para abordar de manera integral la investigación se optará por una metodología mixta que combine enfoques cuantitativos y cualitativos. Por un lado, la metodología cuantitativa permitirá recopilar datos numéricos y medir las diferentes pautas propuestas. Por otro lado, la metodología cualitativa se utiliza para comprender a profundidad las

dimensiones contempladas en la formulación de los *prompts*, así como analizar detalladamente cuáles de estas dimensiones han sido completadas por la IA. La combinación de ambas metodologías permitirá obtener un panorama completo y detallado sobre la utilidad y efectividad de los textos generados por esta inteligencia para la enseñanza de ELE. Hemos optado por seleccionar un conjunto de 15 *prompts* que servirán como base para obtener un total de 15 textos producidos. Es importante destacar que esta selección no representa una limitación definitiva en el volumen de texto, ya que se podrá ajustar el número de *prompts* en función de la relevancia y calidad de los resultados obtenidos durante el desarrollo de la investigación. En cada elaboración de *prompts* y en cada texto obtenido se tendrá en cuenta el cumplimiento de los siguientes criterios presentados en el cuadro 2. Cada uno de estos criterios posee una especificidad por nivel, es decir, se deben cumplir ciertos requisitos para comparar que haya un progreso o no en los textos generados.

Cuadro 2. Criterios y escala de índice.



Con el propósito general de esta investigación en mente, se plantea una situación hipotética que representa a un estudiante ficticio con objetivos específicos. Esta estrategia es fundamental para poder elaborar los prompts de manera precisa, ya que nos proporciona una base para suponer metas específicas en el contexto de la clase.

Situación hipotética

Un estudiante de intercambio académico, proveniente de los EE.UU, ha llegado a la Argentina para estudiar en la provincia de Córdoba. Cuenta con un nivel B1 de español, lo que significa que puede establecer una comunicación efectiva en situaciones cotidianas y posee un conocimiento básico de gramática y vocabulario. Sin embargo, desea mejorar su habilidad para comprender y producir textos narrativos en español.

Objetivos de la clase:

- Comprender el uso variado de los tiempos verbales en pasado.
- Reconocer los componentes del género narrativo.
- Adquirir vocabulario específico del pasado.
- Identificar elementos culturales Argentinos no estereotipados (particularmente de Córdoba).

Contenidos de la clase:

- Reconocimiento de elementos narrativos: estructura, personajes, argumento, tiempo y espacio.
- Desarrollo de vocabulario específico relacionado con la narrativa literaria, como términos para describir personajes, emociones, ambientes y eventos en el pasado.
- Identificación de elementos culturales de Argentina.

Objetivos del alumno: A través de la comprensión de textos narrativos literarios, el alumno busca ampliar su conocimiento del idioma y su capacidad para expresar emociones y describir situaciones de manera más detallada y vívida. Esto le permitirá acceder a una mayor variedad de textos escritos u orales en español.

Técnica de recolección de datos

La técnica de recolección de datos en esta investigación tendrá como primer paso la apertura de la plataforma de los *Large Language Models* en donde se deberá seguir los siguientes pasos:

1. Diseño de los *prompts*: se diseñan prompts que siguen criterios específicos. Estos prompts son fundamentales para guiar a la IA hacia la generación de textos que sean adecuados para la enseñanza de español como lengua extranjera.
2. Generación de textos: se introduce cada prompt y se genera un texto narrativo como respuesta.
3. Recopilación de datos según la matriz.

4. Análisis de textos: los textos se evalúan en función de los criterios mencionados previamente y detallamos qué criterios han sido logrados o no por la IA.
5. A partir de esto, proponemos mejoras en la elaboración de los prompts a tener en cuenta para mejorar la producción.

Nuestra técnica de recolección se alinea de manera efectiva con lo que se conoce como *Modelo CDO* propuesto por Bereiter y Scardamalia (1987). Aunque este modelo se desarrolló originalmente en el ámbito de la escritura y el aprendizaje, su aplicación a nuestra investigación resulta pertinente. El *Modelo CDO* (comparar, diagnosticar y operar) se enfoca en el proceso de revisión y mejora de un texto. En un primer paso, se realiza una comparación entre la representación mental que el autor tiene de su texto con el texto real que quiere lograr, identificando discrepancias o desajustes. Cuando se detectan estos desajustes, se procede a la etapa de diagnóstico, donde el autor se esfuerza por identificar la causa de estos desajustes haciendo uso de su conocimiento retórico y lingüístico. Una vez que se ha diagnosticado la causa, se toma una decisión: cambiar el plan del texto o modificar el texto existente, lo que nos lleva a la fase de operar, en la que se eligen tácticas para implementar los cambios necesarios (Cassany, 1999).

Este modelo guarda una estrecha relación con nuestra técnica de recolección de datos ya que en nuestro proceso de investigación, es la IA la que se encarga de elaborar y producir los textos a partir de los *prompts* proporcionados por parte de los investigadores. A través de esta metodología buscamos perfeccionar los textos generados por esta misma. Este enfoque de revisión y mejora contribuye de manera significativa a nuestros objetivos de investigación y a la calidad de los textos generados.

Matriz de análisis

En esta investigación se ha optado por utilizar la siguiente matriz (cuadro de análisis) para poder dar cuenta, de una manera más precisa, el cumplimiento del texto obtenido por la IA con las variables mencionadas anteriormente. En el primer recuadro de unidad de análisis se colocará el texto al que corresponde esa evaluación. Este será determinado a partir de las siete variables que aparecen explícitas y se marcará un índice de valor que parte de 1 (menos logrado) a 5 (altamente logrado).

Unidad de análisis: Texto generado por la IA						
Variables						
Precisión lingüística	Relevancia temática	Nivel de complejidad	Adecuación cultural	Coherencia con los objetivos	Quantum informativo	Adecuación a la estructura del género narrativo
Índice						
1	1	1	1	1	1	1
2	2	2	2	2	2	2
3	3	3	3	3	3	3
4	4	4	4	4	4	4
5	5	5	5	5	5	5

Referencias:

Ali, JKM, Shamsan, MAA., Hezam, TA y Mohammed, AAQ (2023). Impact of ChatGPT on learning motivation: Teachers and students' voices. *Revista de estudios ingleses en Arabia Felix*, vol. 2 (1), pp. 41–49. <https://doi.org/10.56540/jesaf.v2i1.51>

Areizaga, E., Gómez, I., Ibarra, E. (2005) El componente cultural en la enseñanza de lenguas como línea de investigación. *Revista de Psicodidáctica*, vol. 10 (2), pp. 27-45 *Universidad del País Vasco*. <https://www.redalyc.org/articulo.oa?id=17510204>

Ariyaratne, S., Iyengar, KP, Nischal, N., Chitti Babu, N., & Botchu, R. (2023). A comparison of ChatGPT-generated articles with human-written articles. *Radiol esquelético* vol. 52 (9), 1755-1758. <https://doi.org/10.1007/s00256-023-04340-5>

Beaugrande, R., Dressler, W. (1997). Introducción a la lingüística del texto. *Editorial Ariel, S.A.* <https://es.scribd.com/document/405014696/BeauGrande-Dressler-Introduccion-a-la-Lingu-icharolles-stica-del-Texto-pdf>

Bereiter, C., Scardamalia, M. (1987). La psicología de la composición escrita. *Lawrence Erlbaum Associates. Ontario Institute for Studies in Education*, pp. 265-269. <https://books.google.com.ar/books?hl=es&lr=&id=zdP5AQAAQBAJ&oi=fnd&pg=PP1>

[&dq=Bereiter,+C.+y+M.+Scardamalia+\(1987\)+The+psychology+of+written+compositi
on.+Hillsdale,+N.J:+Erlbaum.&ots=d VWuGCGSR&sig=3fj3e9gIJRfzh2kusOmIJFCNJ5s
#v=onepage&q&f=false](#)

Charolles, M. (1983). Coherence as a principle in the interpretation of discourse. *Interdisciplinary Journal for the Study of Discourse*. vol. 3 (1), pp. 71-98. <https://doi.org/10.1515/text.1.1983.3.1.71>

Consejo de Europa (2020). Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación. *Volumen complementario. Servicio de publicaciones del Consejo de Europa: Estrasburgo*. www.coe.int/lang-cefr.

Leunard, H., Rachmawati, R., Zani, BN y Maharjan, K. (2023). GPT Chat: Opportunities and Challenges in the Learning Process of Arabic Language in Higher Education. *Revista Internacional de Lengua y Tecnología*, vol. 2 (1), pp. 11–23. <https://doi.org/10.55849/jiltech.v2i1.439>

Liu, Y., Zhang, Z., Zhang, W., Yue, S., Zhao, X., Cheng, X., Zhang, Y., & Hu, H (2023). ArguGPT: evaluating, understanding and identifying argumentative essays generated by GPT models. *Shanghai Jiao Tong University*. <https://doi.org/10.48550/arXiv.2304.07666>

Morillo Lozano, M. D. C. (2016). Aprendizaje adaptativo. Tesis de maestría. *Universidad de Valladolid, Escuela Técnica Superior de Ingeniería Informática*. <http://uvadoc.uva.es/handle/10324/21000>

Ocaña-Fernández, Y., Valenzuela-Fernández, L. A., y Garro-Aburto, L. L. (2019). Inteligencia artificial y sus implicaciones en la educación superior. *Revista Propósitos Y Representaciones*, vol. 7 (2), pp. 536–568. <https://doi.org/10.20511/pyr2019.v7n2.274>

Sagástegui, D. (2004). Una apuesta por la cultura: el aprendizaje situado. *Revista Electrónica Sinéctica*. vol. 24, pp. 30-39 *Instituto Tecnológico y de Estudios Superiores de Occidente Jalisco*. <https://www.redalyc.org/articulo.oa?id=99815918005>

Viramonte de Ávalos, M. (1997). Lengua, ciencias, escuela, sociedad. Para una educación lingüística integral. *I.S.B.N. 950-581-658-8. Colección nuevos caminos. Ediciones*

https://books.google.com.ar/books?id=0X2uP9rL6aQC&pg=PA3&hl=es&source=gbs_selected_pages&cad=1#v=onepage&q&f=false

Mota de Cabrera, C.; Villalobos, J. (2007). El aspecto socio-cultura del pensamiento y del lenguaje: visión Vygotskyana. *Vol. 11 n.38. Facultad de Humanidades y Educación. Universidad de Los Andes, ULA. Educere.*
https://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1316-49102007000300005

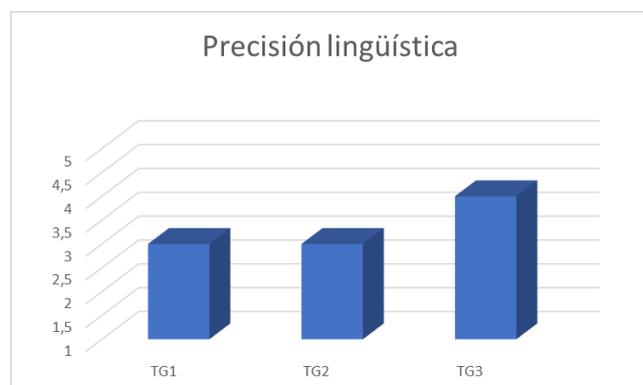
Wilson Cheong Hin Hong (2023). "The impact of Chat GPT on foreign language teaching and learning: Opportunities in education and research". *Centre for Teaching and Learning Enhancement, Revista de Tecnología e Innovación Educativa (JETI) vol. 5(1), pp. 37-45.* <https://jeti.thewsu.org/index.php/cieti/article/view/103>

Anexo:

Prueba piloto

1. Precisión lingüística

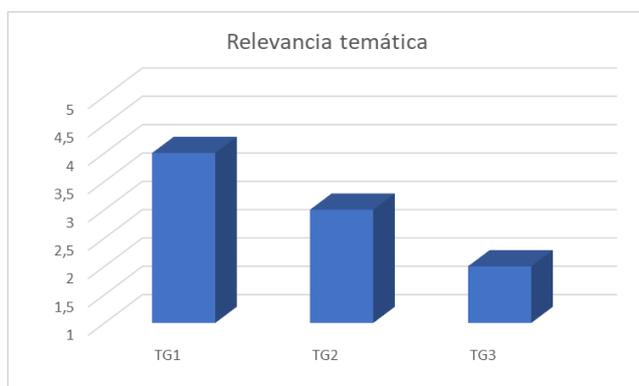
En todos los textos generados, la precisión lingüística se mantiene alta y aumenta progresivamente. La IA demuestra un desempeño sobresaliente en términos de corrección gramatical y léxica. Las estructuras de las oraciones son coherentes y se observa un uso adecuado de los tiempos verbales en español. Esta variable es la más lograda en los textos. Sin embargo, pudimos notar en los textos generados algunos errores en cuanto a la concordancia de número.



2. Relevancia temática

Esta variable está relacionada con la adecuación al prompt solicitado. Si bien en el TG1 cumple altamente, a medida que le brindamos más especificaciones vemos que la variable no es lograda con éxito.

En el segundo texto se incluye un personaje llamado Martina, lo que demuestra un intento de enriquecer la trama. Sin embargo, en el tercer texto, aunque se sigue la estructura temporal solicitada en el prompt, no se agrega un conflicto significativo. A pesar de los esfuerzos, sigue existiendo margen para mejorar esta variable. Se podría incorporar más detalles y elementos narrativos que fortalezcan la conexión entre el lector y la historia, lo que beneficiaría el ajuste al propósito de la lección y objetivos del alumno.



3. Nivel de complejidad, adecuación cultural, quantum informativo y adecuación al género narrativo.

En términos generales, estas variables se han mantenido en un nivel de menos logradas. A pesar de los pedidos exigidos en el prompt, la IA no ha podido cumplir adecuadamente con estas variables. El nivel de complejidad por ahora mantiene estructuras oraciones simples que no son adecuadas para el nivel B1. Por otro lado, la adecuación cultural es la variable que más alcance positivo ha logrado ya que no se han incorporado referencias culturales pertinentes para los propósitos de clase ni del alumno. El quantum informativo ha disminuido ya que si bien proporciona detalles e información sobre los hechos narrados, todavía se requiere de una estructuración pertinente para lograr una narración efectiva. Por último, la adecuación al género narrativo todavía no se perfila como una variable lograda. Los textos narrativos carecen de elementos esenciales de los textos narrativos como la introducción de conflictos y su resolución. En futuros prompts, se recomienda proporcionar instrucciones aún claras y específicas.

