

Análisis de la competencia léxica en ELE con Gemini: retroalimentación automatizada basada en IA

Análise da competência lexical em espanhol como língua estrangeira com o Gemini: retroalimentação automatizada baseada em IA

Analysis of Lexical Competence in Spanish as a Foreign Language (ELE) with Gemini: AI-Based Automated Feedback

María-Victoria Cantero-Romero ^{*1}, Ana María Ortiz-Colón ^{†2} y Salud María Jiménez-Zafra ^{‡3}

¹Universidad de Jaén, Facultad de Humanidades y ciencias de la educación, Departamento de Filología Española, Sistemas Inteligentes de Acceso a la Información, Jaén, España.

²Universidad de Jaén, Facultad de Humanidades y ciencias de la educación, Departamento de pedagogía, Jaén, España.

³Universidad de Jaén, Escuela politécnica superior de Jaén, Departamento de informática, Centro de Estudios Avanzados en Tecnologías de la Información y la Comunicación, Sistemas Inteligentes de Acceso a la Información, Jaén, España.

Resumen

El léxico es un componente esencial en el desarrollo de la competencia escrita en lenguas extranjeras. En el presente estudio se aborda la capacidad léxica de un grupo de estudiantes de español como Lengua Extranjera (ELE) de un centro de idiomas de una universidad española. El propósito del estudio es determinar si el conocimiento léxico permite nivelar al alumnado siguiendo el MCER. Para ello, se utiliza el modelo de lenguaje Gemini, al que se le solicita la nivelación de los estudiantes tras proporcionarle información léxica de sus textos extraída con la herramienta Textflow y se le pide, además, retroalimentación, con el objetivo de orientar al alumnado para alcanzar el siguiente nivel. Finalmente, se evalúa la capacidad de Gemini para nivelar de forma correcta y se analiza la calidad educativa de su retroalimentación. En definitiva, este estudio pretende comprobar el potencial de los modelos de lenguaje de gran tamaño (LLM) como apoyo a la labor docente y al aprendizaje autónomo.

Palabras clave: Español como lengua extranjera. Expresión escrita. Vocabulario. Procesamiento del lenguaje natural. Inteligencia artificial.

Resumo

O léxico é um componente essencial no desenvolvimento da competência escrita em línguas estrangeiras. O presente estudo aborda a capacidade lexical de um grupo de alunos de Espanhol como Língua Estrangeira (ELE) de um centro de idiomas de uma universidade espanhola. O objetivo do estudo é determinar se o conhecimento lexical permite classificar os alunos de acordo com o QECR. Para isso, utiliza-se o modelo de linguagem Gemini, ao qual é solicitada a classificação dos alunos após fornecer-lhe informações lexicais de seus textos extraídas com a ferramenta Textflow e, além disso, é solicitado feedback, com o objetivo de orientar os alunos para alcançarem o próximo nível. Por fim, avalia-se a capacidade do Gemini de classificar corretamente e analisa-se a qualidade educacional de seu feedback. Em suma, este estudo pretende verificar o potencial dos modelos de linguagem de grande porte (LLM) como apoio ao trabalho docente e à aprendizagem autônoma.

Palavras-chave: Espanhol como língua estrangeira. Expressão escrita. Vocabulário. Processamento de linguagem natural. Inteligência artificial.

Abstract

Vocabulary is an essential component in the development of written proficiency in foreign languages. This study examines the lexical ability of a group of students learning Spanish as a Foreign Language (ELE) at a

Textolivre
Linguagem e Tecnologia

DOI: 10.1590/1983-3652.2026.64168

Sección:
Artículos

Autor correspondiente:
María-Victoria
Cantero-Romero

Editor de sección:
Hugo Heredia Ponce
Editor de maquetación:
Rayany Chaves

Recibido el:
31 de enero de 2026
Aceptado el:
22 de marzo de 2026
Publicado el:
13 de octubre de 2026

Esta obra está bajo una
licencia «CC BY 4.0».
© ⓘ

*Email: vcantero@ujaen.es

†Email: aortiz@ujaen.es

‡Email: sjzafra@ujaen.es

language centre within a Spanish university. The aim of the study is to determine whether lexical knowledge enables students to be assessed according to the CEFR. To this end, the Gemini language model is used; it is asked to assess the students' proficiency levels after being provided with lexical information from their texts extracted using the Textflow tool, and is also asked to provide feedback, with the aim of guiding students towards reaching the next level. Finally, Gemini's ability to level students correctly is evaluated, and the educational quality of its feedback is analysed. In short, this study aims to assess the potential of large language models (LLMs) as a support for teaching and self-directed learning.

Keywords: Spanish as a foreign language. Written expression. Vocabulary. Natural language processing. Artificial intelligence.

1 Introducción

La competencia en expresión escrita es una de las competencias principales de la enseñanza de idiomas. Su evaluación es uno de los retos más complejos, puesto que supone más tiempo y dedicación que otras competencias por parte del docente. La producción escrita no sólo requiere dominio gramatical por parte del estudiante, sino también una gestión adecuada del léxico.

El léxico constituye la base de la competencia comunicativa, ya que determina la precisión, la cohesión y la riqueza expresiva de los textos producidos por el estudiante. En los últimos años, la enseñanza de lenguas ha empezado a beneficiarse de los desarrollos que han tenido lugar en el campo de la inteligencia artificial, particularmente por medio de los llamados "modelos de lenguaje de gran tamaño" (Large Language Models, LLMs). Estas herramientas son capaces de procesar grandes volúmenes de información lingüística y de generar textos coherentes, por lo que han abierto la puerta a nuevas oportunidades en la enseñanza y evaluación de las lenguas. Así, la Inteligencia Artificial Generativa (IAG) ha pasado a ser un recurso auxiliar, entendiendo que puede analizar, sintetizar, mostrar y ofrecer retroalimentación adaptada al estudiante. Desde una óptica pedagógica, estas herramientas no debe entenderse como un reemplazo del profesor, sino que se pueden observar como un tipo de apoyo alternativo que ayuda a la personalización del aprendizaje y a la autonomía del aprendiz.

A pesar del avance en el plano de la investigación lingüística, medir la riqueza y la variedad léxica de un modo totalmente objetivo es una labor que requiere tiempo, experiencia y la lectura atenta de la producción por parte del profesorado. En el entorno universitario, donde la distribución de grupos numerosos y niveles sumamente diferenciados suele ser la norma, esta tarea se convierte en doblemente complicada, tanto en la gestión y organización como en la labor pedagógica. En este sentido, la automatización del análisis puede ofrecer una vía prometedora, ya que permitiría transformar los descriptores cualitativos del Marco Común Europeo de Referencia para las lenguas (MCER) en objetivos cuantificables que reflejen el nivel de competencia escrita alcanzado por los estudiantes.

En este contexto, investigaciones recientes han comenzado a explorar el potencial de la IAG para la evaluación de la competencia escrita en español. Estudios pioneros han demostrado que modelos de lenguaje como GPT-3.5 pueden emplearse con éxito para la nivelación automática en entornos universitarios, validando la eficacia de técnicas de *zero-shot learning* frente al entrenamiento supervisado con corpus especializados (Cantero Romero; Vallecillo Rodríguez *et al.*, 2025). Asimismo, se ha destacado cómo la integración de estas tecnologías permite no solo medir la habilidad escrita, sino también personalizar el aprendizaje y reducir significativamente la carga del profesorado mediante la automatización de diagnósticos formativos (Cantero Romero; Martín-Valdivia *et al.*, 2026). Estas aportaciones subrayan la viabilidad de utilizar metadatos lingüísticos procesados por modelos de última generación para una clasificación precisa según el MCER.

Este estudio se sitúa en el marco de la innovación metodológica y el objetivo del mismo es observar hasta qué punto la inteligencia artificial, específicamente el modelo de lenguaje Gemini desarrollado por Google, puede interpretar y ofrecer retroalimentación sobre la competencia léxica de los estudiantes de ELE. Por ello, se define un método alternativo al de la evaluación de textos completos que se basa en un conjunto de métricas léxicas obtenidas utilizando la herramienta Textflow. Al proporcionar estas métricas al modelo, se puede comprobar si es capaz, no solo de asignar un nivel de competencia

escrita según el MCER, sino también de proporcionar recomendaciones correctas.

De esta manera, la combinación de Textflow y Gemini define un enfoque metodológico dual que pretende no solo identificar el nivel de competencia escrita, sino también generar una retroalimentación pedagógica correcta que mantenga coherencia con los descriptores del MCER.

El artículo se organiza del siguiente modo: en la siguiente sección se abordará el marco teórico sobre la importancia del léxico en la enseñanza de idiomas. A continuación, en el apartado metodología, se presenta el corpus, el procedimiento y las herramientas utilizadas en el presente estudio. En el cuarto apartado, se presenta el análisis de los resultados, en el quinto apartado y, finalmente, en el último apartado las conclusiones del estudio.

2 Marco teórico

El estudio del léxico constituye uno de los pilares esenciales en la adquisición de segundas lenguas. En palabras de Wilkins (1972, p. 11), “sin gramática se puede comunicar poco, pero sin vocabulario no se puede comunicar nada”. El conocimiento léxico permite no solo expresar significados con precisión, sino también comprender mensajes complejos en contextos reales. Son numerosas las investigaciones que han demostrado que la comprensión lectora y auditiva está directamente vinculada con la cantidad y calidad del vocabulario disponible del aprendiz. En este sentido, diversos estudios sobre cobertura léxica han señalado que para comprender un texto con un grado adecuado de autonomía es necesario reconocer entre el 95 % y el 98 % de las palabras que lo componen. Este umbral ha sido propuesto como referencia en investigaciones sobre comprensión textual y aprendizaje de vocabulario en segundas lenguas (Robles García, 2019; Degraeuwe y Goethals, 2024).

En el ámbito del español como Lengua Extranjera (ELE), la competencia léxica se define como el conocimiento del vocabulario y la capacidad de utilizarlo adecuadamente en distintos contextos comunicativos. El MCER la integra dentro de las competencias comunicativas y establece descriptores de dominio léxico asociados a cada nivel, desde A1 hasta C2. Por su parte, el Plan Curricular del Instituto Cervantes (2006) desarrolla esta perspectiva y clasifica el léxico en nociones generales, palabras de uso común, y nociones específicas, léxico temático vinculado a ámbitos concretos, estableciendo inventarios que se amplían y actualizan según las necesidades comunicativas del aprendiz.

A medida que el estudiante progresa en los niveles del MCER, la variedad y la sofisticación del léxico empleado aumenta. En niveles iniciales (A1–A2), predomina un vocabulario básico y de alta frecuencia; mientras que en los niveles intermedios y avanzados (B1–C1), se observa una mayor presencia de unidades léxicas intermedias o especializadas y un uso más preciso de colocaciones y expresiones idiomáticas. Este patrón de progresión léxica ha sido descrito en estudios sobre producción escrita de aprendientes de ELE, donde se observa un aumento progresivo de la variedad y la precisión del repertorio léxico a medida que aumenta el nivel de competencia (Vine Jara *et al.*, 2021).

La riqueza léxica, también denominada complejidad léxica, se ha consolidado como un constructo multidimensional compuesto por tres subcomponentes: la diversidad léxica, la sofisticación léxica y la densidad léxica. Diversos estudios han utilizado estos indicadores para analizar la calidad de la producción escrita tanto en hablantes nativos como en aprendientes de segundas lenguas. En el ámbito del español, Cuberos Vicente (2019) propone distintos indicadores léxicos para evaluar la calidad textual en producciones nativas y no nativas. En una línea similar, Vine Jara *et al.* (2021) analizan la riqueza léxica en textos de aprendientes de ELE y destacan su relación con la progresión del repertorio léxico en distintos niveles de competencia. Por su parte, Jiménez Pérez, López Martínez y Corbalán Berná (2024) exploran el uso de algoritmos de análisis textual para evaluar la complejidad y la comprensibilidad de textos en español.

La diversidad léxica se relaciona con la variedad de palabras empleadas en un texto; la sofisticación hace referencia a la frecuencia o rareza de los términos; y la densidad alude a la proporción de palabras con contenido semántico frente a las gramaticales. Estas dimensiones permiten describir con mayor precisión la amplitud y calidad del repertorio léxico utilizado por los aprendientes en su producción escrita. En estudios clásicos sobre producción escrita en segundas lenguas, Laufer y Nation (1995) señalan que la variedad del vocabulario constituye un indicador relevante del desarrollo lingüístico del

aprendiz. Investigaciones posteriores han ampliado este enfoque incorporando medidas de sofisticación léxica para analizar la evolución del repertorio léxico en textos de aprendientes, como muestran los trabajos de Kyle y Crossley (2016).

La presente investigación profundiza en la relación entre estas métricas y los niveles de competencia del MCER. Los estudios muestran que los aprendices de niveles más altos producen textos con mayor diversidad y densidad léxica, así como una selección más precisa y variada de categorías gramaticales (Vine Jara *et al.*, 2021). Este progreso léxico se acompaña, además, de una mayor capacidad para reorganizar el discurso y emplear estrategias de compensación, lo que refleja un desarrollo global de la competencia comunicativa.

En los últimos años, el análisis asistido por ordenador ha permitido medir de forma objetiva la complejidad lingüística a través de métricas automáticas. Herramientas como ANTCONC (2022) y Sistemas Inteligentes de Acceso a la Información (SINAI) (2025) o Text Inspector (2026) facilitan el cálculo de índices como la longitud media de palabra, la frecuencia léxica o la proporción de sustantivos, verbos y adjetivos. A partir de estos parámetros, se pueden establecer perfiles léxicos asociados a los niveles de competencia definidos por el MCER. Además, los modelos computacionales más recientes se benefician de recursos como el corpus *LexComSpaL2* (Degraeuwe y Goethals, 2024). Este conjunto de datos ha sido diseñado para entrenar clasificadores de dificultad léxica en aprendices de español L2, lo que demuestra el potencial de la inteligencia artificial para analizar y predecir la complejidad de los textos (Jiménez Pérez; López Martínez y Corbalán Berná, 2024).

La aplicación pedagógica de estas tecnologías es doble. Por un lado, permite generar materiales de aprendizaje ajustados al nivel real del estudiante; por otro, facilita procesos de retroalimentación automatizada, ofreciendo información precisa sobre los aspectos léxicos que requieren refuerzo. En el ámbito del análisis léxico aplicado a la enseñanza de lenguas, Cuberos Vicente (2019) propone distintos indicadores para evaluar la calidad textual y analizar la riqueza léxica en producciones escritas. Por su parte, Robles García (2019) subraya la importancia de generar oportunidades de aprendizaje léxico en el aula a partir de un análisis sistemático del vocabulario utilizado por los aprendientes. En una línea similar, Vine Jara *et al.* (2021) muestran cómo el estudio de la riqueza léxica en textos de aprendientes de ELE puede contribuir a describir el desarrollo del repertorio léxico y orientar la enseñanza del vocabulario. En este contexto, la combinación del análisis cuantitativo de métricas léxicas con una interpretación cualitativa —basada en la pertinencia, la variedad y la adecuación del léxico— constituye una vía prometedora para el desarrollo de herramientas educativas que contribuyan a mejorar la competencia escrita en ELE. Este enfoque orienta la propuesta metodológica del presente estudio, que explora el uso de métricas léxicas automatizadas y modelos de lenguaje para la generación de retroalimentación pedagógica.

3 Metodología

3.1 Objetivos

El objetivo principal de este estudio es analizar la capacidad de un modelo de lenguaje (Google, 2025) para estimar el nivel de competencia léxica de textos escritos por estudiantes de ELE a partir de métricas extraídas automáticamente con la herramienta *Textflow* (Sistemas Inteligentes de Acceso a la Información (SINAI), 2025).

De forma complementaria, se pretende:

1. Evaluar la correspondencia entre los niveles estimados por el modelo y los niveles reales asignados por docentes expertos.
2. Analizar la validez pedagógica de las recomendaciones generadas por el modelo como forma de retroalimentación automatizada.
3. Explorar los límites de la predicción léxica en los distintos niveles del MCER (A1–C1).

3.2 Corpus y participantes

El corpus analizado está compuesto por 316 textos producidos por estudiantes de ELE en el Centro de Estudios Avanzados en Lenguas Modernas de la Universidad de Jaén. Las producciones escritas fueron elaboradas por los estudiantes durante actividades realizadas en el aula de informática

bajo supervisión docente, lo que permitió garantizar que los textos no fueran generados mediante herramientas de inteligencia artificial.

Cada texto fue previamente nivelado por docentes del centro de lenguas con experiencia en la enseñanza de ELE y familiarizados con los descriptores del MCER. Los docentes responsables de esta evaluación cuentan con más de ocho años de experiencia en la enseñanza de ELE y formación especializada en el área, incluyendo estudios de posgrado en enseñanza de ELE. La asignación de niveles se realizó siguiendo los criterios habituales de evaluación de la producción escrita empleados en el centro y basados en los descriptores del MCER distribuyéndose de la siguiente manera: A1 (97 textos), A2 (85), B1 (71), B2 (54) y C1 (9).

Esta distribución refleja la composición real de los grupos en los cursos de ELE, con predominio de niveles iniciales. La menor presencia de producciones en los niveles más avanzados responde a la disponibilidad real de textos en el contexto educativo analizado, donde los cursos iniciales concentran un mayor número de estudiantes. Por este motivo, los resultados correspondientes a los niveles superiores deben interpretarse teniendo en cuenta esta limitación en la muestra.

La recopilación y tratamiento del corpus contó con la aprobación del comité de ética de la Universidad de Jaén. Todas las producciones escritas fueron anonimizadas antes de su procesamiento, garantizando la eliminación de cualquier dato personal identificable. El análisis automático se realizó mediante Textflow, una biblioteca desarrollada en Python que permite calcular diferentes métricas lingüísticas a partir de textos en formato plano. Esta herramienta se limita a procesar las secuencias textuales para extraer indicadores cuantitativos, sin almacenar ni tratar datos personales del alumnado.

3.3 Instrumentos

3.3.1 Herramienta Textflow y métricas utilizadas

El análisis de las producciones se realizó con Textflow (Sistemas Inteligentes de Acceso a la Información (SINAI), 2025), una herramienta desarrollada en Python que permite calcular métricas lingüísticas, estilométricas y de complejidad textual. De entre las múltiples medidas que ofrece, se seleccionaron trece métricas léxicas consideradas más relevantes para la descripción de la competencia escrita (Cuberos Vicente, 2019; Robles García, 2019; Vine Jara *et al.*, 2021; Degraeuwe y Goethals, 2024):

- *volumetrywords*: número total de palabras.
- *volumetryuniqueWords*: número de palabras únicas.
- *volumetryavgWordsLen*: longitud media de palabra.
- *stylometryTTR* y *stylometryRTTR*: índices de diversidad léxica.
- *stylometryFreqStopWords*: frecuencia de palabras funcionales.
- *RelFreqPOSNOUN*, *RelFreqPOSVERB*, *RelFreqPOSADV*, *RelFreqPOSADJ*: frecuencias relativas de sustantivos, verbos, adverbios y adjetivos.
- *complexitynComplexSentence*: número de oraciones complejas.
- *complexityavglenSentence*: longitud media de oración.
- *complexitynRareWords*: número de palabras poco frecuentes.

Estas métricas combinan aspectos cuantitativos (volumen y longitud), cualitativos (diversidad y sofisticación) y estructurales (complejidad). Los datos se exportaron a Excel para su posterior procesamiento por Gemini.

3.3.2 Gemini

Para la estimación del nivel del MCER a partir de las trece métricas extraídas con Textflow se empleó el modelo Gemini 2.0 Flash (Google, 2025), el cual fue seleccionado por su arquitectura de baja latencia y su alta capacidad de razonamiento. La eficacia de este modelo en el procesamiento del español ha sido validada recientemente en la tarea PROFE (Proficiency Assessment of Spanish Learners) en el marco de IberLEF 2025, donde los modelos de la familia Gemini demostraron una precisión sobresaliente (superior al 0.90) en la clasificación automática de niveles de competencia (Rodrigo Yuste *et al.*, 2025).

Gemini 2.0 Flash se integró mediante la API de Google Generative AI y fue configurado a través

de un prompt de rol especializado (Giray, 2023) que lo situaba como un experto en la enseñanza de ELE. La instrucción exacta utilizada para el procesamiento de los datos fue la siguiente:

“Eres un especialista en enseñanza del español como lengua extranjera (ELE), experto en nivelar textos siguiendo el MCER y PCIC. Analiza las métricas extraídas con Textflow de la producción escrita de un estudiante. Devuelve un JSON con:

- *“nivel_estimado”: nivel MCER (A1, A2, B1, B2 o C1)*
- *“justificacion”: breve frase que explique por qué*
- *“recomendaciones”: lista de 3 sugerencias pedagógicas individualizadas para mejorar el léxico*

Datos del texto: data”

Esta configuración técnica permitió que el modelo analizara exclusivamente los valores numéricos proporcionados en la variable data, sin acceder al contenido original de los textos. De este modo, la nivelación y la retroalimentación generada se basan estrictamente en los indicadores de complejidad, diversidad y densidad léxica procesados previamente por Textflow

3.4 Procedimiento

El flujo metodológico de la presente investigación se estructuró en dos fases consecutivas ejecutadas en el entorno de desarrollo Google Colab, utilizando la biblioteca google-generativeai para la gestión de peticiones a la API.

Fase 1: Extracción de métricas. Los 316 textos del corpus se procesaron con la herramienta Textflow para extraer las trece métricas léxicas seleccionadas. La salida de esta fase fue una matriz de datos donde cada producción escrita quedó reducida a una serie de valores numéricos.

Fase 2: Nivelación y generación de retroalimentación pedagógica con Gemini. Los valores numéricos de cada texto se integraron en el prompt especializado detallado en el apartado 3.3.2. Al trabajar exclusivamente con indicadores cuantitativos, se garantizó que el modelo no tuviera acceso al contenido textual original, asegurando la privacidad de los estudiantes. El modelo Gemini 2.0 Flash procesó cada vector numérico de forma individual para generar una salida estructurada en formato JSON, proporcionando la estimación del nivel (A1-C1), la justificación técnica y las propuestas de mejora léxica. Para asegurar la integridad del proceso frente a posibles latencias de la red, se implementó un sistema de gestión de reintentos automáticos y guardado automático cada 25 textos. Las métricas, el nivel estimado, la justificación y las recomendaciones generadas se almacenaron en un archivo de salida (*resultados_gemini.xlsx*) para la fase de análisis.

3.5 Limitaciones metodológicas

El diseño experimental, centrado exclusivamente en métricas léxicas, limita la capacidad del modelo para captar fenómenos discursivos (coherencia, cohesión o adecuación textual). Aun así, permite aislar y evaluar de forma precisa la influencia del componente léxico en la progresión de la competencia escrita, y explorar el potencial de los modelos de lenguaje como herramientas de evaluación y retroalimentación automatizada en contextos de enseñanza de ELE.

4 Resultados

El análisis de los textos se organizó en tres fases complementarias, con el objetivo de ofrecer una visión integrada de la competencia léxica y de su evaluación automatizada. En primer lugar, se llevó a cabo un análisis descriptivo de las métricas extraídas con Textflow, orientado a identificar patrones de progresión léxica en el corpus según los niveles del MCER. En segundo lugar, se realizó un análisis cuantitativo del rendimiento del modelo Gemini en la clasificación automática de los textos, con el fin de evaluar el grado de correspondencia entre los niveles estimados y los niveles reales. Por último, se desarrolló un análisis cualitativo de la retroalimentación generada por el modelo, centrado en valorar su pertinencia pedagógica.

4.1 Análisis descriptivo del corpus (métricas de Textflow)

Se realizó un análisis descriptivo de las métricas extraídas por Textflow con el fin de caracterizar el corpus y confirmar que la progresión léxica se ajusta a lo esperado según los descriptores del MCER

y del PCIC (2006). El corpus total presentó una longitud media de 162 palabras, con una diversidad léxica moderada ($\text{stylometryTTR} = 0,49$; $\text{stylometryRTTR} = 0,82$). La distribución gramatical mostró una alta concentración en sustantivos (0,31) y verbos (0,27).

Como se observa en la Tabla 1, las métricas clave muestran un incremento progresivo y constante en los niveles iniciales (A1-B1), que tiende a estabilizarse en los niveles superiores (B2-C1). Este patrón refleja que la complejidad léxica aumenta con el nivel de competencia, pero las diferencias cuantitativas se vuelven menos significativas una vez que el estudiante ha adquirido un vocabulario funcional intermedio.

Tabla 1. Progresión de métricas léxicas clave según el nivel del MCER.

Nivel real	Nº de muestras	Media de palabras	Media TTR (riqueza léxica)	Frecuencia relativa de sustantivos	Frecuencia relativa de adjetivos
A1	97	125	0,42	0,28	0,09
A2	85	155	0,47	0,30	0,11
B1	71	185	0,52	0,32	0,14
B2	54	175	0,54	0,31	0,13
C1	9	190	0,55	0,31	0,14
Media general	316 (total)	162	0,49	0,31	0,12

Fuente: Elaboración propia.

El índice TTR (Type-Token Ratio) y la Frecuencia Relativa de Adjetivos fueron los indicadores más informativos. En conjunto, estos resultados muestran una progresión léxica diferenciada según el nivel de competencia. En los niveles iniciales (A1–A2) se observa un incremento gradual en la diversidad y en la sofisticación del léxico, que se intensifica en el nivel B1. A partir de este nivel, los valores medios tienden a estabilizarse en B2 y C1.

Desde esta perspectiva, los resultados ponen de manifiesto que la progresión de la competencia léxica no responde necesariamente a un crecimiento continuo e in crescendo a lo largo de todos los niveles del MCER. Por el contrario, las métricas analizadas muestran fases de expansión más marcadas en los niveles iniciales y una posterior estabilización en los niveles superiores. Este comportamiento permite delimitar el alcance metodológico de los indicadores léxicos automáticos, que resultan especialmente adecuados para caracterizar las primeras etapas del aprendizaje, pero menos sensibles a los matices propios de la competencia avanzada.

4.2 Análisis cuantitativo del rendimiento de clasificación

Para evaluar la precisión del modelo en la estimación del nivel MCER, se compararon los niveles estimados por Gemini con los niveles reales asignados por los docentes. Se calcularon las métricas estándar de evaluación de modelos de clasificación: precisión, cobertura (recall) y F1-score, utilizando la biblioteca scikit-learn de Python. La precisión mide el porcentaje de aciertos sobre las predicciones positivas, el recall calcula el porcentaje de textos correctamente clasificados dentro de cada nivel real, y la medida F1-score es la media armónica entre precisión y recall, como medida de equilibrio general.

En la Tabla 2 se presentan los resultados de clasificación, donde se puede observar que el modelo tiene un desempeño significativamente más elevado en los niveles iniciales.

En los niveles A1 y A2 se obtienen valores de F1-score de 0.51 y 0.49, respectivamente, lo que indica que las métricas léxicas constituyen indicadores eficaces para la identificación de textos de niveles básicos, donde el repertorio léxico aún se encuentra en expansión y presenta diferencias cuantificables claras.

En los niveles intermedios y avanzados, el rendimiento disminuye de forma notable. En el nivel B1, el modelo muestra dificultades para clasificar correctamente los textos, mientras que en los niveles B2

Tabla 2. Progresión de métricas léxicas clave (valores medios) por nivel MCER.

Nivel	Precisión	Recall	F1-score	N.º de textos
A1	0.84	0.37	0.51	97
A2	0.33	0.92	0.49	85
B1	0.26	0.14	0.18	71
B2	0.00	0.00	0.00	54
C1	0.00	0.00	0.00	9

Fuente: Elaboración propia.

y C1 no se registran aciertos. Este patrón confirma la tendencia observada anteriormente, según la cual la capacidad discriminativa de las métricas léxicas se reduce en los niveles superiores.

En conjunto, los resultados delimitan el alcance empírico de la clasificación automática basada en métricas léxicas, confirmando su utilidad en los estadios iniciales del aprendizaje y evidenciando la necesidad de incorporar otros indicadores lingüísticos para abordar la evaluación de niveles más elevados del MCER.

4.3 Análisis cualitativo del potencial pedagógico

Por último, se llevó a cabo un análisis cualitativo de las justificaciones y recomendaciones generadas por Gemini con el fin de valorar su utilidad desde una perspectiva pedagógica. Este análisis no se centró en la exactitud de la predicción del nivel, sino en la pertinencia y calidad de la retroalimentación léxica ofrecida al alumnado.

Para sistematizar la evaluación de las salidas del modelo, se diseñó una rúbrica de cuatro niveles (1–4), adaptada a los descriptores del MCER y del PCIC, que permitió analizar las recomendaciones en tres dimensiones: pertinencia léxica, progresión léxica y coherencia con los descriptores oficiales. La aplicación de esta rúbrica fue realizada por los mismos docentes de ELE que participaron en la nivelación inicial del corpus (véase apartado 3.2). Antes de la evaluación, los evaluadores revisaron conjuntamente los criterios de valoración con el fin de asegurar una interpretación homogénea de las categorías, resolviendo posteriormente las posibles discrepancias mediante discusión hasta alcanzar consenso. Los criterios y niveles de valoración empleados en esta evaluación se presentan en la Tabla 3.

Con el fin de ilustrar el tipo de retroalimentación generada por el modelo, se presentan en la Tabla 4 algunos ejemplos representativos de las recomendaciones producidas por Gemini a partir de las métricas léxicas analizadas. Estos ejemplos tienen un carácter ilustrativo y no sustituyen el análisis sistemático realizado mediante la rúbrica descrita anteriormente.

A partir de la rúbrica presentada en la Tabla 3, se evaluaron las justificaciones y recomendaciones generadas por el modelo en los distintos niveles de competencia. Los resultados del análisis cualitativo, organizados por dimensión y agrupados en niveles iniciales (A1–A2) y niveles intermedios–avanzados (B1–C1), se recogen en la Tabla 5.

En los niveles iniciales (A1–A2), el análisis muestra que el modelo identifica con precisión rasgos léxicos característicos de estos estadios, como el uso predominante de vocabulario de alta frecuencia, la limitación de los campos semánticos y la ausencia de variación semántica significativa. Este comportamiento resulta coherente con las métricas de frecuencia léxica proporcionadas al modelo, que informan sobre la recurrencia de determinadas palabras en los textos. En este sentido, aunque Gemini no tuvo acceso a las producciones completas del alumnado, las métricas de frecuencia léxica proporcionadas al modelo incluyen información sobre la recurrencia de determinadas palabras, lo que permite detectar patrones de uso y formular recomendaciones acordes con los ámbitos léxicos más representados en los datos de entrada. Asimismo, las recomendaciones resultan concretas y realistas, orientadas a una ampliación progresiva del repertorio léxico coherente con la progresión esperada hacia el nivel siguiente.

En los niveles intermedios y avanzados (B1–C1), las justificaciones del modelo tienden a formularse

Tabla 3. Rúbrica de evaluación cualitativa (1–4).

Valor	Pertinencia léxica	Progresión léxica	Coherencia MCER/PCIC
1 – Muy baja	Las observaciones no identifican rasgos léxicos propios del nivel o son incorrectas.	Las recomendaciones no tienen relación con la progresión del nivel o no ayudan a avanzar.	No hay correspondencia con los descriptores; conceptos clave ausentes o erróneos.
2 – Baja	Las observaciones son demasiado generales o incompletas; identifican algún rasgo pertinente, pero sin precisión.	Recomendaciones genéricas, sin secuenciación clara o sin relación directa con necesidades del nivel.	Se mencionan aspectos del nivel, pero de forma superficial o incompleta.
3 – Adecuada	El modelo identifica varios rasgos léxicos propios del nivel con cierta precisión, aunque con falta de detalle.	Recomendaciones funcionales que apoyan el avance típico del nivel, aunque sin gran especificidad.	Coincidencia general con el MCER/PCIC, aunque faltan matices o fineza analítica.
4 – Alta	Las observaciones describen con precisión rasgos léxicos característicos del nivel (amplitud, campo semántico, complejidad) y los conectan con el desempeño real del texto.	Recomendaciones precisas, graduadas y directamente vinculadas a campos léxicos y funciones propias del nivel.	Las observaciones y recomendaciones reflejan con precisión los descriptores léxicos del nivel.

Fuente: Elaboración propia

de manera más general, lo que dificulta la identificación de rasgos léxicos diferenciales propios de cada nivel. No obstante, las recomendaciones mantienen una orientación pedagógica válida, centrada en la ampliación del vocabulario, la precisión semántica y la consolidación del léxico académico, aunque con un menor grado de especificidad que en los niveles iniciales.

En conjunto, los resultados del análisis cualitativo ponen de relieve que, aunque la capacidad del modelo para discriminar niveles disminuye a medida que aumenta la complejidad de la competencia escrita, la retroalimentación generada conserva un valor formativo relevante, especialmente como apoyo al aprendizaje autónomo y como complemento a la intervención docente en contextos de enseñanza de ELE.

5 Discusión

Los resultados obtenidos confirman la hipótesis inicial de que las métricas léxicas proporcionadas por *Textflow* son indicadores válidos para discriminar los niveles iniciales de competencia escrita en el aprendizaje de ELE, pero insuficientes para caracterizar adecuadamente los niveles intermedios y avanzados. Esta tendencia coincide con lo señalado por De la Torre García (2016), quien observa que la progresión lingüística en ELE implica un aumento de la variedad verbal y de la complejidad estructural que trasciende la dimensión léxica.

En los niveles A1 y A2, Gemini logró asociar correctamente los patrones de diversidad y frecuencia léxica con descriptores del MCER, lo que respalda la correlación entre riqueza léxica y competencia comunicativa descrita por Robles García (2019). La relación entre la cantidad y variedad de vocabulario y el desarrollo de la competencia escrita es especialmente visible en los estadios iniciales, donde la progresión depende de la adquisición y consolidación de un repertorio básico de unidades léxicas. En estos niveles, las métricas como *stylometryTTR* o la frecuencia relativa de sustantivos y adjetivos se comportan como predictores fiables del nivel de dominio.

Resulta especialmente significativo que estas asociaciones se produzcan sin acceso del modelo a los textos completos del alumnado, pero sí a métricas que incluyen información sobre la frecuencia de determinadas palabras. Aunque Gemini no procesa el contenido discursivo íntegro de las producciones,

Tabla 4. Ejemplos de recomendaciones generadas.

Nivel	Ejemplo de recomendación
A1	■ Trabajar la concordancia de género y número entre sustantivos y adjetivos (ej. “buenos día” → “buen día” o “buenos días”). ■ Practicar el uso correcto de las preposiciones “a”, “en”, “de” para indicar tiempo y lugar. ■ Ampliar el vocabulario relacionado con invitaciones y planes, incluyendo alternativas más formales e informales.
A2	■ Practicar el uso correcto de las preposiciones “en”, “a”, “de” para indicar lugar y posesión. ■ Repasar la conjugación de verbos regulares e irregulares en presente de indicativo, prestando especial atención a las formas de “ser”, “estar” y “gustar”. ■ Ampliar el vocabulario relacionado con lugares, actividades de tiempo libre y expresiones de afecto.
B1	■ Ampliar el vocabulario para expresar opiniones de manera más precisa y matizada. ■ Trabajar en la corrección de errores gramaticales comunes, como la concordancia de género y número, y el uso correcto de los tiempos verbales. ■ Practicar la escritura de textos más estructurados y cohesionados, utilizando conectores discursivos para mejorar la claridad y la fluidez.
B2	■ Ampliar el vocabulario para expresar ideas de manera más precisa y matizada. ■ Trabajar en la corrección de errores gramaticales sutiles, como la concordancia y el uso de preposiciones, mediante ejercicios de revisión y edición. ■ Practicar la redacción de textos argumentativos más extensos para mejorar la organización de ideas y la cohesión textual.

Fuente: Elaboración propia.

la disponibilidad de datos de frecuencia léxica le permite identificar patrones de uso recurrente y formular observaciones coherentes con el repertorio léxico empleado. Esto refuerza la utilidad de un enfoque de evaluación apoyado en metadatos lingüísticos que, aún sin incorporar el análisis textual completo, ofrece información pedagógicamente relevante.

Sin embargo, a partir del nivel B1, los resultados muestran una caída significativa en la precisión del modelo. Este hallazgo coincide con los planteamientos de Cuberos Vicente (2019) y Consolini y Kyle (2024), quienes sostienen que la complejidad léxica por sí sola no basta para explicar el desarrollo global de la competencia escrita, ya que intervienen otros factores —como la organización textual, la cohesión y la adecuación discursiva— que no se reflejan en las métricas numéricas. En este sentido, la menor capacidad predictiva del modelo en los niveles superiores sugiere que el análisis basado exclusivamente en métricas léxicas resulta necesariamente parcial. Podría ser necesario añadir a las medidas de análisis del discurso y de la sintaxis, la media de longitud, las oraciones subordinadas y la variedad de conectores utilizados. Esta idea se refiere a lo que propone Hu (2021) cuando manifiesta que a medida que se incorporan factores de complejidad léxico-sintáctica, la predicción de los niveles del MCER mejora.

Por otro lado, conviene señalar que las métricas léxicas analizadas en este estudio no constituyen una medida exhaustiva del nivel global de producción escrita según el MCER. El análisis se centra específicamente en la dimensión léxica de la competencia escrita, que representa solo uno de los componentes del desempeño comunicativo descrito por dicho marco. En consecuencia, las recomendaciones generadas por el modelo se interpretan como orientaciones pedagógicas generales basadas en la lectura de estas métricas, y no como una identificación directa de errores concretos en las producciones escritas.

Tabla 5. Tabla de evaluación de las salidas.

Dimensión	Nivel A1–A2	Nivel B1–C1	Valoración
1. Pertinencia léxica	El modelo identifica repertorios limitados, uso predominante de vocabulario de alta frecuencia y ausencia de variación semántica. Ej.: Identificación precisa de un “vocabulario limitado a temas personales relacionados con el estado de salud y los deseos” (A1).	Las justificaciones se vuelven vagas y generales. Ej.: Observaciones como “vocabulario adecuado” o “imprecisiones léxicas” que no discriminan rasgos diferenciales entre B1 y B2.	Alta (3–4) en iniciales; baja–media (1–2) en avanzados.
2. Progresión léxica	Las recomendaciones son concretas y realistas, sugiriendo una ampliación controlada del repertorio. Ej.: Sugerir “ampliar el vocabulario sobre salud y bienestar” (A1) o “sobre lugares y profesiones” (A2).	Las recomendaciones pierden especificidad. Ej.: Proponer solo “ampliar el vocabulario temático” sin concretar estrategias de precisión léxica, variación estilística o colocaciones.	Clara (3–4) en iniciales; limitada (1–2) en avanzados.
3. Coherencia MCER/PCIC	Se evidencia un ajuste claro a los descriptores del PCIC. Ej.: Las observaciones reflejan el repertorio limitado y el control parcial de la morfología esperado para esos niveles.	Hay una ausencia de referencia a rasgos léxicos propios de niveles intermedios-altos en los descriptores. Ej.: Omisión de conceptos como el uso de léxico abstracto, matices semánticos o colocaciones.	Alta (3–4) en iniciales; baja (1–2) en avanzados.

Fuente: Elaboración propia.

Desde el punto de vista pedagógico, el modelo Gemini presenta un potencial muy importante para la retroalimentación formativa. A pesar de que tiene menor exactitud para los niveles más altos, la retroalimentación generada, en forma de comentario o sugerencias específicas, es un aspecto que puede ser relevante en la enseñanza del léxico. Este tipo de retroalimentación guarda relación con la visión del Plan Curricular del Instituto Cervantes (2006) que enfatiza la autonomía del aprendiente y las estrategias de ampliación léxica. Las recomendaciones que pueden ser generadas por el modelo Gemini, como por ejemplo practicar la morfología verbal, reforzar el vocabulario temático o recomendaciones sobre el uso de recursos digitales de memorización, pueden ser utilizadas como estrategia en la secuencia didáctica adaptada al perfil del alumnado.

Estos resultados refuerzan la idea de que el análisis de métricas cuantitativas puede convivir con enfoques cualitativos. Como sostienen Vine Jara *et al.* (2021) y Jiménez Pérez, López Martínez y Corbalán Berná (2024), la combinación de ambos enfoques permite una visión más completa del desarrollo de la competencia léxica. En este sentido, la información objetiva debe entenderse como un punto de partida que requiere, a su vez, de una interpretación pedagógica fundamentada. Desde esta perspectiva, el modelo Gemini no sustituye en ningún momento la valoración del docente, sino que funciona como un asistente que proporciona, en un formato eficaz y coherente con los descriptores del MCER, información diagnóstica útil para orientar la retroalimentación, ajustar las actividades de refuerzo léxico y apoyar la toma de decisiones didácticas.

Por último, la consideración de Gemini como herramienta de apoyo pone de relieve el potencial de la inteligencia artificial en la didáctica de ELE. Los resultados coinciden con Degraeuwe y Goethals (2024) en señalar los beneficios de los modelos de IA en la creación de recursos adaptativos. No obstante, como advierte Hu (2021), la fiabilidad de estos sistemas depende tanto de la solidez de las métricas empleadas como de su interpretación lingüística contextualizada.

En este sentido, una línea prometedora para futuras investigaciones es la exploración de modelos de lenguaje específicamente ajustados al ámbito educativo, como LearnLM. A diferencia de los modelos generalistas, como Gemini 2.0 Flash, estos sistemas se diseñan bajo principios pedagógicos

que priorizan el andamiaje del aprendizaje. Su integración podría contribuir a evaluar si una mayor sensibilidad didáctica permite abordar con mayor precisión la falta de granularidad observada en los niveles B2 y C1, especialmente en relación con la complejidad discursiva que las métricas léxicas por sí solas no capturan.

Además, en futuras investigaciones, sería pertinente analizar empíricamente el impacto pedagógico de las recomendaciones generadas por el modelo en el aprendizaje de los estudiantes, con el fin de evaluar su efectividad didáctica en contextos reales de enseñanza de ELE.

6 Conclusiones

Este estudio ha puesto de manifiesto que el análisis de métricas léxicas con *Textflow* y su interpretación automatizada con Gemini permite visualizar tendencias significativas dentro de la evolución de la competencia escrita en el aprendizaje de ELE, sobre todo en los niveles iniciales del MCER. Las variables de diversidad léxica y frecuencia léxica fueron los mejores predictores del nivel de los textos. No obstante, su capacidad discriminativa disminuye en los niveles intermedios y avanzados, lo que refuerza la idea de que la complejidad lingüística no puede explicarse únicamente a partir de la dimensión léxica.

A pesar de las limitaciones, la integración del modelo Gemini aporta un valor pedagógico. El hecho de que el modelo pueda ofrecer recomendaciones personalizadas, dependiendo de los resultados numéricos, significa un avance hacia una retroalimentación formativa más ágil, objetiva y adaptada al perfil del alumno. En este sentido, la inteligencia artificial se convierte en un recurso que puede complementar la práctica docente, útil en el diagnóstico de necesidades léxicas, en el fomento de la autonomía en el aprendizaje y en el seguimiento continuo del progreso de los aprendices.

Desde una perspectiva de metodología de trabajo, el presente trabajo evidencia la posibilidad de cruzar instrumentos de análisis lingüístico automático (*Textflow*) junto a modelos de lenguaje generativo (Gemini) en una analítica cuantitativa-cualitativa, que no sólo es capaz de clasificar producciones escritas, sino que permite establecer información pedagógica que se puede trasladar al aula, orientada a la mejora del repertorio léxico y poder ampliar las estrategias de aprendizaje.

Entre las líneas de continuidad futuras, se acoge la idea de incluir también métricas sintácticas y métricas de la dimensión discursiva que, junto a las medidas léxicas, podrán contribuir a medir las producciones más complejas y completas. Igualmente, se espera ampliar la recogida de textos y el tamaño de la muestra con textos de otros contextos educativos y con distintas variedades de aprendientes. También se muestra interesante explorar el valor que pueden tener los docentes, y los aprendientes en sus interacciones (retroalimentación automática), para poder valorar la utilidad de estos ejercicios en el aprendizaje y en la enseñanza de ELE.

En resumen, los resultados de este estudio caminan hacia una situación muy congruente entre lingüística aplicada e inteligencia artificial: las métricas reflejan un acercamiento al análisis objetivo y replicable, mientras que los modelos de lenguaje denotan un acercamiento al análisis con un potente valor pedagógico. Esto permite dar viabilidad a la combinación de ambas para abordar el análisis, la evaluación y la enseñanza del léxico del aula de ELE.

Referencias

ANTCONC. *AntConc*. [S. l.]: [s. n.], 2022. Disponible en: <https://www.laurenceanthony.net/software/antconc/>. Acceso en: enero 2026.

CANTERO ROMERO, María Victoria; MARTÍN-VALDIVIA, María Teresa; ORTIZ-COLÓN, Ana María y JIMÉNEZ-ZAFRA, Salud María. Medición de la habilidad escrita en español como lengua extranjera con inteligencia artificial generativa. *RIED – Revista Iberoamericana de Educación a Distancia*, Madrid, volumen 29, número 1, páginas 353-379, 2026. DOI: 10.5944/ried.45486.

CANTERO ROMERO, María Victoria; VALLECILLO RODRÍGUEZ, María Estrella; ORTIZ-COLÓN, Ana María y JIMÉNEZ-ZAFRA, Salud María. GPT-3.5 for ELE leveling in a university environment: a comparative study of zero-shot learning and fine-tuning. *Procesamiento del Lenguaje Natural*, Alicante, número 75, páginas 53-63, 2025.

CONSOLINI, Carla H. y KYLE, Kristopher. The relationship between L2 Spanish proficiency and features of written lexical and lexicogrammatical use. *Applied Linguistics*, Oxford, volumen 45, número 6, páginas 992-1009, 2024. DOI: 10.1093/applin/amae032.

CUBEROS VICENTE, María del Rocío. *Indicadores léxicos de calidad textual en español nativo y no nativo*. 2019. Tesis (Doctorado en Didáctica de las Ciencias, las Lenguas, las Artes y las Humanidades) – Universidad de Granada, Granada.

DE LA TORRE GARCÍA, Nuria. La complejidad lingüística en los niveles de competencia del MCER: el caso de la variedad verbal en la expresión escrita en ELE. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, Madrid, número 20, 2016.

DEGRAEUWE, Jasper y GOETHALS, Patrick. LexComSpaL2: a lexical complexity corpus for Spanish as a foreign language. En: PROCEEDINGS of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024). Torino: ELRA; ICCL, 2024. páginas 10432-10447. International Conference on Language Resources and Evaluation (LREC-COLING), 2024.

GIRAY, Louie. Prompt engineering with ChatGPT: a guide for academic writers. *Annals of Biomedical Engineering*, New York, volumen 51, número 12, páginas 2629-2633, 2023.

GOOGLE. *Gemini API: Gemini models*. [S. l.]: [s. n.], 2025. Disponible en: <https://ai.google.dev/gemini-api/docs/models>. Acceso en: oct. 2025.

HU, Xiaolin. Predicting CEFR levels in L2 oral speech, based on lexical and syntactic complexity. *Asia Pacific Journal of Corpus Research*, Incheon, volumen 2, número 1, páginas 35-45, 2021. DOI: 10.22925/apjcr.2021.2.1.35.

INSTITUTO CERVANTES. *Plan curricular del Instituto Cervantes: niveles de referencia para el español*. Madrid: [s. n.], 2006. Disponible en: https://cvc.cervantes.es/ensenanza/biblioteca_ele/plan_curricular/default.htm. Acceso en: oct. 2025.

JIMÉNEZ PÉREZ, Elena del Pilar; LÓPEZ MARTÍNEZ, Olivia y CORBALÁN BERNÁ, Javier. LECOLE algorithm: readability, readability, and comprehensibility in Spanish. *XLinguae*, Nitra, volumen 17, número 1, páginas 12-25, 2024. DOI: 10.18355/XL.2024.17.01.12.

KYLE, Kristopher y CROSSLEY, Scott. The relationship between lexical sophistication and independent and source-based writing. *Journal of Second Language Writing*, Amsterdam, volumen 34, páginas 12-24, 2016. DOI: 10.1016/j.jslw.2016.10.003.

LAUFER, Batia y NATION, Paul. Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics*, Oxford, volumen 16, número 3, páginas 307-322, 1995. DOI: 10.1093/applin/16.3.307.

ROBLES GARCÍA, Pablo. Competencia léxica y oportunidades de aprendizaje léxico en el aula de lenguas: una aproximación teórica. *Lengua y Habla*, Mérida, número 23, páginas 157-182, 2019.

RODRIGO YUSTE, Álvaro; MORENO ÁLVAREZ, Sergio; PÉREZ GARCÍA-PLAZA, Alberto; PEÑAS PADILLA, Anselmo; AGERRI, Rodrigo; FRUNS-JIMÉNEZ, Javier y SORIA-PASTOR, Inés. Overview of PROFE at IberLEF 2025: language proficiency evaluation. *Procesamiento del Lenguaje Natural*, Alicante, número 75, páginas 487-497, 2025.

SISTEMAS INTELIGENTES DE ACCESO A LA INFORMACIÓN (SINAI). *Textflow*. [S. l.]: [s. n.], 2025. Disponible en: <https://github.com/sinai-uja/textflow>. Acceso en: oct. 2025.

TEXT INSPECTOR. *Text Inspector*. [S. l.]: [s. n.], 2026. Disponible en: <https://textinspector.com>. Acceso en: enero 2026.

VINE JARA, Ana; FUENTES RIFFO, Karina; NEIRA MARTÍNEZ, Angie y POZA MOLINA, Carlos. Riqueza léxica en textos escritos por aprendientes de ELE: una contribución a las nociones específicas del Plan Curricular del Instituto Cervantes. *Nueva Revista del Pacífico*, Valparaíso, número 75, páginas 215-240, 2021. DOI: 10.4067/S0719-51762021000200201.

Contribuciones de los autores

María-Victoria Cantero-Romero: Conceptualización, Curación de datos, Investigación, Metodología, Visualización, Recursos, Administración de proyectos, Redacción – borrador original; **Ana María Ortiz Colón:** Validación, Supervisión, Redacción – Revisión y edición; **Salud María Jiménez Zafra:** Validación, Software, Análisis formal, Adquisición de financiamiento, Supervisión, Redacción – Revisión y edición.

Agradecimientos

Agradecimientos al Centro de Estudios Avanzados en Lenguas Modernas de la Universidad de Jaén. Este trabajo ha sido parcialmente financiado por el Ministerio para la Transformación Digital y de la Función Pública y por el Plan de Recuperación, Transformación y Resiliencia —financiado por el programa NextGenerationEU de la UE— en el marco del proyecto Desarrollo de modelos ALIA. También ha contado con el apoyo del Proyecto CONSENSO (PID2021-122263OB-C21), financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea NextGenerationEU/PRTR, del proyecto ROMANET (CERV-2024-CHAR-LITI -101215052), financiado por la Unión Europea en el marco del programa Ciudadanos, Igualdad, Derechos y Valores, de los proyectos HEART-NLP-UJA (PID2024-156263OB-C21) y VERITAS-H (AIA2025-163322-C64), financiados por MICIU/AEI/10.13039/501100011033 y por FEDER/UE, y del proyecto GALENO-IA (DGP_PIDI_2024_00852), financiado por la Unión Europea en el marco del Programa de la Estrategia de Especialización Inteligente para la Sostenibilidad de Andalucía (S4Andalucía) 2021-2027.

Disponibilidad de datos

Los datos de investigación solo están disponibles previa solicitud.