

Impacto de la inteligencia artificial en el aprendizaje de vocabulario en español como lengua extranjera: un estudio cuasiexperimental con estudiantes sinohablantes

FAN WU

JING WU (CORRESPONDING AUTHOR)

Guangdong University of Foreign Studies, China

Received: 2025-05-13 / Accepted: 2025-10-06

DOI: <https://doi.org/10.30827/portalin.viXV.33566>

Porta Linguarum ISSN paper edition: 1697-7467, ISSN digital edition: 2695-8244

RESUMEN: Este estudio analiza el impacto de tres plataformas de aprendizaje asistido por inteligencia artificial (Duolingo, Eshelper y TalkAI) en la adquisición de vocabulario en español como lengua extranjera por parte de estudiantes sinohablantes de nivel A1-A2. El estudio adopta un diseño cuasiexperimental con un grupo de control y otro de intervención, complementado por análisis cuantitativo y cualitativo, y fundamentado en marcos teóricos sobre adquisición léxica, multimodalidad, contexto comunicativo y organización semántica. Los resultados evidencian mejoras estadísticamente significativas en la comprensión auditiva y en la retención léxica en el grupo experimental, atribuibles a la integración de estímulos multimodales y contextos comunicativos. No obstante, se detectan limitaciones en la adaptación cultural de los contenidos y en el tratamiento del léxico figurado. El estudio concluye con recomendaciones orientadas al diseño pedagógico de plataformas digitales, dirigidas a promover una enseñanza léxica más personalizada, contextualizada e interculturalmente relevante.

Palabras clave: adquisición léxica; inteligencia artificial; enseñanza de español como lengua extranjera; plataformas digitales; aprendientes sinohablantes

The impact of artificial intelligence on vocabulary acquisition in Spanish as a foreign language: a quasi-experimental study with Chinese learners

ABSTRACT: This study analyzes the impact of three artificial intelligence-assisted learning platforms (Duolingo, Eshelper, and TalkAI) on vocabulary acquisition in Spanish as a foreign language among Chinese-speaking learners at the A1–A2 level. The study adopts a quasi-experimental design with a control group and an intervention group, complemented by both quantitative and qualitative analyses, and grounded in theoretical frameworks on lexical acquisition, multimodality, communicative context, and semantic organization. The results show statistically significant improvements in listening comprehension and vocabulary retention in the experimental group, attributable to the integration of multimodal stimuli and communicative contexts. However, limitations were identified in the cultural adaptation of content and the treatment of figurative language. The study concludes with recommendations for the pedagogical design of digital platforms aimed at promoting more personalized, situated, and intercultural relevant vocabulary instruction.

Keywords: vocabulary acquisition; artificial intelligence; teaching Spanish as a foreign language; digital platforms; Chinese-speaking learners

1. INTRODUCCIÓN

La adquisición de vocabulario constituye un componente esencial en el desarrollo de la competencia comunicativa en una segunda lengua, ya que incide directamente en la fluidez, precisión y adecuación del discurso (Nation, 2001). Este proceso no ocurre de forma aislada, sino que interactúa con otras dimensiones lingüísticas —gramática, fonología y pragmática—, conformando un sistema interdependiente que sustenta la construcción progresiva del conocimiento lingüístico.

En este contexto, los recientes avances en inteligencia artificial (IA), especialmente en el procesamiento del lenguaje natural y el aprendizaje automático, han transformado los paradigmas tradicionales de la enseñanza de lenguas extranjeras, introduciendo entornos de aprendizaje más dinámicos, personalizados y centrados en la eficacia del proceso formativo. A través de algoritmos adaptativos y análisis de datos en tiempo real, estas plataformas permiten proporcionar *input* lingüístico variado, retroalimentación inmediata, seguimiento individualizado y ajustes pedagógicos continuos. Sin embargo, su aplicación también plantea desafíos que requieren un examen crítico, como la adecuación cultural de los contenidos, la protección de datos personales y la necesidad de una validación empírica rigurosa que evalúe su eficacia en contextos educativos específicos.

Paralelamente, el creciente interés por el español en China, impulsado por el fortalecimiento de los vínculos económicos, culturales y académicos con los países hispanohablantes, ha generado una demanda educativa sostenida. No obstante, los aprendientes sinohablantes en niveles iniciales (A1-A2, según el Marco Común Europeo de Referencia para las Lenguas, MCER) enfrentan múltiples obstáculos en la adquisición léxica del español. Entre las principales dificultades se encuentran la escasa exposición al idioma en contextos reales, las diferencias estructurales con la lengua materna y las limitaciones de enfoques tradicionales de enseñanza poco adaptados culturalmente.

En este escenario, las plataformas basadas en IA ofrecen un potencial prometedor para afrontar estos desafíos. Investigaciones recientes, como las de Palomares Marín (2022) y Zou y Wang (2021), han evidenciado beneficios del uso de la gamificación, la repetición espaciada y el seguimiento personalizado del progreso en entornos digitales de aprendizaje. No obstante, la aplicación de estas herramientas en la enseñanza de español como lengua extranjera (ELE) para sinohablantes, especialmente en niveles básicos, continúa siendo un terreno poco investigado. La limitada existencia de estudios empíricos sistemáticos restringe la comprensión de su impacto real y de los desafíos pedagógicos y culturales asociados.

Este estudio se propone analizar el impacto de tres plataformas basadas en IA (Duolingo, Eshelper y TalkAI) en la enseñanza léxica del español a estudiantes chinos de nivel A1-A2. En concreto, los objetivos específicos son: 1) evaluar su eficacia en la adquisición léxica; 2) examinar su capacidad de personalización y adaptación cultural; y 3) formular recomendaciones para el diseño pedagógico ajustado a las necesidades de este perfil de aprendientes.

A partir de estos objetivos, se formulan las siguientes preguntas de investigación: 1) ¿Qué efectos tiene el uso de estas plataformas en la adquisición léxica del alumnado chino de nivel A1-A2?; 2) ¿Qué desafíos plantea su implementación en términos de personalización y adecuación cultural en el contexto chino?; y 3) ¿Qué mejoras pueden incorporarse para optimizar su eficacia pedagógica?

Este trabajo aspira a contribuir a la innovación didáctica en la enseñanza de ELE mediada por tecnología. Desde una perspectiva conceptual, profundiza en la comprensión del impacto de la IA en el aprendizaje léxico en contextos socioculturales específicos; y desde un enfoque aplicado, ofrece orientaciones para el diseño de soluciones pedagógicas más eficaces, inclusivas y culturalmente pertinentes, especialmente en entornos con baja inmersión lingüística, caracterizados por un contacto limitado con la lengua meta tanto en el plano oral como en el cultural.

2. DESAFÍOS DE LA INTELIGENCIA ARTIFICIAL EN LA ENSEÑANZA LÉXICA

La integración de la IA en la enseñanza de lenguas extranjeras ha despertado un creciente interés académico, impulsado por avances en procesamiento del lenguaje natural, el aprendizaje automático y el desarrollo de plataformas interactivas. Sin embargo, los estudios específicos sobre la enseñanza del vocabulario en ELE mediante IA continúan en una fase inicial, con especial escasez de estudios centrados en el alumnado sinohablante en contextos de ELE. Esta sección revisa cuatro áreas clave de investigación: 1) potenciación tecnológica de la IA; 2) innovación didáctica en modelos de enseñanza asistida por IA; 3) optimización de la enseñanza mediante el análisis de datos; y 4) desafíos y estrategias de respuesta.

2.1. Potenciación tecnológica de la IA en la enseñanza del vocabulario

El desarrollo de tecnologías basadas en IA, como ChatGPT, ha transformado los enfoques pedagógicos tradicionales del aprendizaje léxico, pues proporciona experiencias de aprendizaje enriquecidas con múltiples canales sensoriales que simulan situaciones comunicativas reales (Li, 2024; Zhang & Hong, 2023). Estas herramientas favorecen la práctica activa del vocabulario al integrar significados léxicos, usos pragmáticos y referencias culturales en contextos cotidianos, como compras, consultas médicas o interacciones administrativas (Hu, 2023).

Asimismo, plataformas de aprendizaje como Duolingo han incorporado tecnologías del procesamiento del lenguaje natural y algoritmos de repetición espaciada, lo que ha mostrado efectos positivos en la retención de vocabulario frecuente (Palomares Marín, 2022). La corrección automática en tiempo real ha demostrado ser un mecanismo compensatorio eficaz en entornos educativos con interacción oral limitada, contribuyendo así al desarrollo fonológico de los aprendientes (Liu et al., 2022).

De manera complementaria, el uso de tecnologías inmersivas, como la realidad aumentada y la realidad virtual, ha mostrado efectos prometedores en el aprendizaje de vocabulario contextualizado, al combinar *input* visual, interacción activa y estímulos situados (Lai & Chen, 2021; Zou et al., 2021). No obstante, persisten desafíos vinculados a la adaptabilidad dialectal y a la escasa representación de contenidos culturales específicos. Tal como indica Palomares Marín (2022), menos del 19 % de las aplicaciones actuales cubren las variedades lingüísticas del español en América Latina, lo que revela una brecha sustancial en la inclusión cultural de estas herramientas.

2.2. Innovación didáctica en modelos de enseñanza asistida por IA

En el marco de la innovación pedagógica, los modelos de enseñanza asistida por IA han promovido una transformación metodológica en la planificación y ejecución de la enseñanza léxica. Uno de los desarrollos más relevantes es la personalización del aprendizaje a través de sistemas adaptativos que ajustan dinámicamente los contenidos según el nivel léxico del alumnado. De acuerdo con estudios recientes, estos algoritmos, al analizar errores recurrentes, pueden generar tareas personalizadas que reducen la tasa de fallos y mejoran la retención del vocabulario frecuente hasta en un 28 % (Zou & Wang, 2021).

Paralelamente, ha cobrado relevancia el paradigma de colaboración entre el profesorado y la IA. Diversas investigaciones coinciden en que la IA debe entenderse como un recurso pedagógico complementario, y no como un sustituto del rol docente (Hu, 2023; Li, 2023). Aunque estas tecnologías pueden asumir tareas automatizables —como los dictados o la corrección sistemática—, el profesorado sigue siendo insustituible en el desarrollo de competencias cognitivas superiores y en la mediación intercultural (Kramsch, 1993; Risager, 2006). En consonancia con esta perspectiva, se ha propuesto el denominado “modelo de doble sujeto”, en el que los docentes diseñan secuencias didácticas digitalizadas que el alumnado internaliza mediante la interacción humano-máquina (Zhang & Hong, 2023; Li, 2024).

Otro ámbito emergente, estrechamente vinculado con lo anterior, es la formación docente mediada por IA. Un ejemplo lo constituyen los talleres de diseño didáctico que incorporan herramientas como ChatGPT, capaces de fortalecer la competencia tecnológica del profesorado y facilitar la integración crítica de recursos digitales en el aula (Li, 2024). No obstante, también se han señalado riesgos asociados a una dependencia excesiva de estas tecnologías. Un uso indiscriminado, por ejemplo, podría derivar en una memorización mecánica, desvinculada de una comprensión semántica profunda o contextualizada, especialmente en el caso del vocabulario con carga pragmática o metafórica (Monteagudo, 2024). Asimismo, diversas investigaciones subrayan la necesidad de una formación crítica y ética en el uso docente de la IA, en particular frente a los desafíos de su integración efectiva y responsable en los modelos educativos contemporáneos (Sanz Manzanedo, 2025).

2.3. Optimización de la enseñanza mediante el análisis de datos

El análisis de datos educativos ha abierto nuevas vías para optimizar la enseñanza del vocabulario en lenguas extranjeras. Una de las iniciativas más destacadas es la plataforma “Formative Assessment Data Platform for Foreign Language Teaching” (Kong & Qin, 2021), implementada en más de 30 cursos universitarios en China. Esta herramienta permite recopilar automáticamente datos generados durante el aprendizaje, tanto dentro como fuera del aula, lo que facilita un análisis longitudinal del desempeño y decisiones pedagógicas fundamentadas en evidencia. Entre sus funciones destacan el rastreo de trayectorias individuales, la identificación de errores frecuentes —como en los relacionados con colocaciones léxicas— y la activación automática de módulos de refuerzo personalizados cuando se detectan niveles elevados de dificultad léxica.

Más allá de los ajustes pedagógicos inmediatos, algunas investigaciones han aplicado el análisis de datos a dimensiones más complejas del aprendizaje, como la transcripción automática de interacciones orales para identificar carencias léxicas y patrones comunicati-

vos. Sin embargo, cuestiones como la privacidad de los datos y la transparencia algorítmica siguen siendo desafíos pendientes (Mohamed, 2024).

Otro campo emergente es la modelización del comportamiento de aprendizaje. Liu et al. (2022), mediante modelos de ecuaciones estructurales, demostraron que el diseño antropomórfico de agentes virtuales —como avatares o tutores virtuales con rasgos humanos— influye positivamente en la motivación del alumnado. En el contexto de la enseñanza léxica, Hu y Hu (2020) observaron que figuras culturalmente próximas pueden mejorar la atención y reducir la carga cognitiva. Asimismo, el estudio de Sanz Manzanedo (2025), centrado en docentes de lenguas en la región de Toscana, concluye que el diseño empático y el tono afable de los avatares virtuales favorecen su aceptación, especialmente entre mujeres docentes, lo que sugiere una motivación adicional derivada de la interacción con agentes digitalmente humanizados.

2.4. Desafíos y estrategias de respuesta

Pese a su gran potencial, la integración de la IA en la enseñanza léxica plantea desafíos significativos. Uno de los más críticos es la escasa adaptabilidad cultural de muchas plataformas, que tienden a replicar marcos culturales occidentales sin considerar las especificidades sociolingüísticas del alumnado sinohablante (Li, 2024). Esta desconexión puede repercutir desfavorablemente en la adquisición léxica y en el compromiso de los estudiantes con el aprendizaje.

Además, la evidencia empírica advierte que el uso acrítico de estas tecnologías puede fomentar un aprendizaje superficial y descontextualizado, centrado en definiciones rápidas pero carente de profundidad semántica (Zhang & Hong, 2023). En el plano ético, surgen preocupaciones en torno a la protección de datos, la transparencia algorítmica y el sesgo cultural en los contenidos generados. Como estrategias de respuesta, se han propuesto marcos éticos aplicables a entornos educativos (Li, 2023), así como la incorporación de formación ética sobre IA en los programas de capacitación docente (Monteagudo, 2024).

En síntesis, la literatura reciente pone de relieve los avances notables en personalización, análisis inteligente y entornos inmersivos aplicados al aprendizaje léxico. Sin embargo, persisten lagunas importantes en la adaptación cultural de los contenidos, la profundización semántica y la adecuación didáctica en contextos específicos.

3. FUNDAMENTOS TEÓRICOS

La adquisición de vocabulario en lenguas extranjeras ha sido abordada desde diversas perspectivas teóricas que destacan sus dimensiones cognitivas, comunicativas y pedagógicas. Este estudio se apoya en cuatro marcos conceptuales fundamentales —la teoría de la adquisición de vocabulario de Paul Nation, la teoría del contexto, la teoría de la multimodalidad y la teoría del campo semántico— que orientan tanto el diseño de los instrumentos como la interpretación de los resultados. Asimismo, se incorporan desarrollos recientes sobre la aplicación de la IA en la enseñanza léxica, con el fin de analizar hasta qué punto las plataformas digitales actuales integran estos principios teóricos y qué limitaciones presentan en contextos específicos de aprendizaje.

3.1. La teoría de la adquisición de vocabulario de Paul Nation

Paul Nation (2001) propone un modelo integral para el aprendizaje del vocabulario en L2, estructurado en cuatro dimensiones: *input* receptivo (lectura y audición), *output* productivo (habla y escritura), estudio formal (morfología, colocación y reglas gramaticales) y automatización (fluidez y uso espontáneo). A ello se suma una estructura tridimensional del conocimiento léxico: forma (pronunciación y ortografía), significado (básico, polisémico y asociativo) y uso (gramatical, pragmático y discursivo).

Este modelo ha sido ampliamente validado en la lingüística aplicada por su equilibrio entre aprendizaje explícito e implícito, así como por su versatilidad en contextos diversos de enseñanza. En el presente estudio, dicha teoría fundamenta el diseño de los instrumentos de evaluación, permitiendo un análisis multicomponente del léxico que trasciende la memorización y aborda de forma integrada la comprensión, la producción y el uso contextualizado.

3.2. La teoría del contexto

La teoría del contexto, formulada inicialmente por Malinowski (1923) y desarrollada posteriormente por Halliday (1985), sostiene que el significado léxico no puede entenderse aisladamente, sino que está condicionado por factores situacionales, socioculturales y lingüísticos.

Este planteamiento teórico —según el cual la comprensión y el uso del vocabulario dependen del contexto comunicativo— ha sido respaldado por aportaciones consolidadas (Nation, 2001), así como por estudios empíricos recientes que demuestran cómo los entornos de aprendizaje que simulan escenarios reales favorecen la retención léxica y su aplicación pragmática (Chen & Yen, 2013; Wang, 2016; Zhu et al., 2024). Entre las estrategias emergentes, destacan las plataformas digitales basadas en IA, que intentan integrar el contexto comunicativo en la instrucción léxica. TalkAI, por ejemplo, genera de forma automatizada diálogos cotidianos enmarcados en situaciones comunicativas reales.

No obstante, persisten limitaciones relevantes, especialmente en lo que respecta a la adaptación cultural de los contenidos y a la representación adecuada de las variaciones pragmáticas. Estas deficiencias se tornan más visibles en contextos cultural y lingüísticamente distantes, como el del alumnado sinohablante, donde la ausencia de referentes culturales compartidos puede dificultar la apropiación contextualizada del léxico y su empleo eficaz en situaciones de comunicación intercultural. Este fenómeno sugiere la necesidad de seguir perfeccionando los modelos de contextualización en las plataformas de IA, con especial atención a los matices sociopragmáticos de la lengua meta.

3.3. La teoría de la multimodalidad

Propuesta por Kress (2001), la teoría de la multimodalidad sostiene que el aprendizaje se potencia mediante la interacción simultánea de distintos modos semióticos —como el lenguaje verbal, la imagen, el sonido o la gestualidad— que intervienen en la construcción del significado. Investigaciones posteriores han indicado que la presentación de contenidos en formatos multimodales favorece la retención del vocabulario, reduce la carga cognitiva y estimula la motivación, facilitando así un aprendizaje más eficaz y sostenido en el tiempo (Mayer, 2005; Jewitt, 2009).

Este marco resulta especialmente pertinente, ya que las tres plataformas basadas en IA seleccionadas para este estudio integran diferentes recursos multimodales como imágenes, audio, animaciones y simulaciones conversacionales. Dichos elementos están diseñados para reforzar la asociación entre forma, significado y uso del léxico, en línea con los principios postulados por esta teoría. Por ejemplo, Duolingo utiliza estímulos auditivos y visuales combinados con señales correctivas inmediatas, mientras que TalkAI incorpora interacciones orales contextualizadas que estimulan la producción activa del vocabulario aprendido. La eficacia de estos recursos ha sido corroborada en estudios previos que demuestran cómo el uso de imágenes facilita la adquisición de vocabulario al fortalecer la conexión semántica (Boers et al., 2017) y cómo los entornos de simulación interactiva favorecen la retención a largo plazo del léxico aprendido (Franciosi et al., 2016).

Sin embargo, continúa siendo un reto importante la personalización efectiva de la experiencia multimodal, ya que muchas plataformas aún presentan secuencias homogéneas que no consideran adecuadamente las diferencias individuales en estilos de aprendizaje, ritmos cognitivos o preferencias sensoriales. En este sentido, uno de los objetivos del presente estudio es analizar en qué medida las aplicaciones evaluadas ofrecen adaptabilidad real en la presentación del *input* multimodal y si dicha adaptabilidad repercute en una mejora sustantiva en la adquisición del vocabulario, como han señalado estudios previos (Teng & Zhang, 2023).

3.4. La teoría del campo semántico

Formulada por Trier (1931), la teoría del campo semántico sostiene que el vocabulario no se adquiere como una suma de palabras aisladas, sino como parte de redes de significados interrelacionados, organizadas según afinidades conceptuales. Este enfoque facilita la memorización y la comprensión al permitir que los aprendientes establezcan asociaciones lógicas entre términos dentro de un mismo dominio léxico, contribuyendo así al desarrollo de la competencia contextual en L2 (Nation, 2001).

En el ámbito de la enseñanza asistida por IA, esta teoría ha sido incorporada en plataformas como Eshelper, que organiza el léxico por campos semánticos (alimentos, emociones o actividades cotidianas), con el fin de reforzar el aprendizaje asociativo. No obstante, aunque la evidencia reciente indica que este tipo de organización puede reforzar redes semánticas en aprendientes de ELE, los algoritmos actuales siguen dominados por relaciones semánticas básicas y lineales, mostrando claras limitaciones a la hora de procesar fenómenos más complejos como la polisemia, el lenguaje figurado o las expresiones idiomáticas (Teng & Zhang, 2023; Lai & Chen, 2021).

Aunque los avances en procesamiento del lenguaje natural y aprendizaje profundo han ampliado la capacidad de modelado léxico, sigue siendo un reto lograr una representación semántica que sea adaptable tanto a las necesidades individuales como a los marcos culturales del alumnado (van den Broek et al., 2018; Lee & Park, 2020). Este aspecto resulta especialmente relevante en contextos interculturales, como el de los estudiantes sinohablantes de español, cuya estructura cognitiva y referencias culturales difieren considerablemente del espacio hispanohablante.

3.5. Tendencias actuales en la enseñanza léxica con IA

En los últimos años, la integración de tecnologías basadas en IA ha favorecido el desarrollo de metodologías innovadoras, como el aprendizaje adaptativo, la corrección automatizada,

la evaluación inteligente y los entornos inmersivos, orientadas a personalizar el proceso de aprendizaje, mejorar la retención a largo plazo y aumentar la eficiencia del *input* léxico (Zou et al., 2021; Teng & Zhang, 2023; Mohamed, 2024). Plataformas como Duolingo, Eshelper y TalkAI ya integran algoritmos que ajustan dinámicamente los contenidos en función del rendimiento del alumnado, sus errores frecuentes y su progreso individual. Asimismo, el uso de tecnologías como la realidad virtual ha permitido simular situaciones comunicativas reales, lo que facilita la adquisición léxica en contextos pragmáticos y culturalmente relevantes.

No obstante, estos avances técnicos coexisten con importantes retos éticos, culturales y pedagógicos. Uno de los desafíos más notables es la escasa adaptabilidad cultural de muchas plataformas comerciales, que tienden a priorizar contextos europeos o anglófonos sin integrar adecuadamente las realidades socioculturales de los aprendientes sinohablantes (Li, 2024). Esta carencia afecta especialmente a la enseñanza de unidades léxicas con carga cultural, matices pragmáticos o significados figurados, dificultando su apropiación y uso comunicativo.

Aun considerando las limitaciones culturales, es imperativo también abordar las restricciones inherentes al diseño técnico y pedagógico de las plataformas. Desde el punto de vista técnico, el sesgo algorítmico puede obstaculizar una personalización efectiva, al generar recomendaciones que no reflejan con precisión el perfil real del alumnado (Mohamed, 2024). Además, en el ámbito pedagógico, persisten deficiencias en el diseño instruccional, particularmente en lo concerniente a la transferencia del vocabulario adquirido a contextos comunicativos auténticos, un aspecto determinante para el desarrollo funcional de la competencia léxica.

En términos generales, si bien la incorporación de la IA en la enseñanza léxica constituye una vía prometedora, su aprovechamiento real exige superar retos aún latentes, como la contextualización cultural, la integración semántica profunda y la personalización adaptativa del *input*. Las teorías aquí revisadas permiten asumir una mirada crítica hacia estas tecnologías, que guiará el análisis empírico del presente estudio.

4. METODOLOGÍA

En esta sección se expone la metodología del estudio, con especial atención al diseño experimental, los participantes, los recursos didácticos utilizados y el tratamiento de los datos.

4.1. Diseño de la investigación

Este estudio adopta un diseño cuasiexperimental con un grupo de intervención y un grupo de control, con el objetivo de examinar el efecto del uso sistemático de plataformas basadas en IA sobre la adquisición de vocabulario en estudiantes sinohablantes de español como lengua extranjera. La hipótesis principal sostiene que la incorporación de estas herramientas tecnológicas favorece significativamente el aprendizaje léxico en comparación con métodos tradicionales de enseñanza.

4.2. Participantes

La muestra estuvo compuesta por 52 estudiantes universitarios chinos (entre 18 y 19 años), matriculados en el segundo cuatrimestre del Grado en Español en una univer-

sidad pública del sur de China. Todos los participantes presentaban un nivel inicial A1-A2 según el MCER.

Para garantizar la homogeneidad del punto de partida, se administró una prueba diagnóstica de vocabulario y comprensión lectora basada en contenidos del nivel A2. A partir de los resultados obtenidos, los estudiantes fueron asignados aleatoriamente a dos grupos equivalentes —grupo experimental ($n = 26$) y grupo de control ($n = 26$)— mediante un procedimiento de aleatorización estratificada.

Ninguno de los participantes había residido en países hispanohablantes ni había mantenido contacto significativo con el español fuera del entorno académico, lo que permitió controlar la exposición externa al *input* léxico y asegurar condiciones iniciales comparables entre ambos grupos.

4.3. Instrumentos y materiales

Este apartado describe los recursos didácticos y tecnológicos utilizados durante la intervención, así como los instrumentos de evaluación aplicados para medir la adquisición léxica y la valoración del proceso por parte del alumnado.

4.3.1. Selección y justificación del léxico y de las plataformas

El léxico objeto de estudio fue extraído de los manuales Aula Internacional 1 y 2, ampliamente utilizados en la enseñanza de ELE. Se eligieron 120 unidades léxicas de alta frecuencia, adecuadas al nivel A2 del MCER y clasificadas en seis campos semánticos: alimentación, transporte, ocio, vivienda, salud y rutina. Estas 120 unidades conformaron el núcleo del contenido trabajado durante la intervención y sustentaron tanto las sesiones presenciales como las tareas digitales y las pruebas (escrita, auditiva y oral).

En cuanto a las herramientas tecnológicas, se seleccionaron tres aplicaciones de IA de uso extendido entre aprendientes sinohablantes: Duolingo, Eshelper y TalkAI. La selección se fundamentó en criterios de validez empírica, coherencia pedagógica y adecuación contextual. Estas plataformas integran estrategias esenciales para una adquisición léxica eficaz, como la repetición espaciada, la retroalimentación inmediata y la personalización del contenido, en consonancia con los principios formulados por Nation (2001).

1. Duolingo: plataforma internacional con un enfoque gamificado que ofrece microlecciones dinámicas, sistema de recompensas y ejercicios automatizados de pronunciación. Su diseño multimodal refuerza la asociación entre sonido, imagen y forma escrita, lo que favorece la retención léxica mediante actividades lúdicas estructuradas.
2. Eshelper: herramienta desarrollada para sinohablantes, que está centrada en la corrección morfosintáctica y semántica a partir de prácticas contextualizadas.
3. TalkAI: aplicación emergente orientada a la producción oral a través de conversaciones automatizadas, que facilitan la activación del léxico en situaciones comunicativas simuladas.

4.3.2. Pruebas de evaluación del vocabulario

Los instrumentos de evaluación fueron diseñados para medir la adquisición léxica desde una perspectiva multidimensional, conforme al modelo de Nation (1990; 2001), considerando tres dimensiones fundamentales: la dicotomía entre amplitud y profundidad léxica (Nation & Waring, 1997), la diferenciación entre capacidad receptiva y productiva, así como la triple dimensión de forma, significado y uso.

1. Prueba escrita: versión adaptada de la *Vocabulary Levels Test* (Schmitt & Clapham, 2001), enfocada en vocabulario del nivel A2 del MCER, con ítems de opción múltiple contextualizados.
2. Prueba auditiva: tareas de completamiento de textos a partir de audios, que evalúan reconocimiento auditivo y deducción semántica.
3. Prueba oral: descripción de imágenes, evaluando la precisión, fluidez y pertinencia léxica.

Estos instrumentos fueron elaborados por tres expertos en didáctica del español y validados mediante dos rondas del método Delphi para lograr consenso experto sobre la adecuación de los ítems al nivel A2 del MCER. La confiabilidad interna fue confirmada mediante coeficientes alfa de Cronbach superiores a 0.82 en todas las pruebas.

4.3.3. Cuestionario de percepción

Se aplicó un cuestionario estructurado para medir la percepción de los estudiantes en relación con la usabilidad, motivación y eficacia de las plataformas de IA. Este instrumento fue adaptado de modelos validados en estudios previos sobre tecnología educativa (Vesselinov & Grego, 2016) y se sometió a pruebas de validez de constructo mediante análisis factorial exploratorio, obteniendo KMO = .709 y Bartlett $p < .001$. El alfa de Cronbach global fue de 0.847, indicando alta consistencia interna.

4.4. Procedimiento de recopilación de datos

La intervención se desarrolló a lo largo de doce semanas. El grupo experimental trabajó con las plataformas de IA durante 30 minutos diarios, cinco días a la semana, realizando actividades personalizadas que se adaptaban a su progreso y ofrecían retroalimentación correctiva automática en tiempo real.

Paralelamente, el grupo de control continuó con métodos tradicionales de enseñanza, centrados en la memorización de listas de vocabulario, ejercicios escritos y explicaciones impartidas por el docente. Ambos grupos recibieron el mismo número de horas de instrucción formal y trabajaron con contenidos léxicos equivalentes.

La instrucción léxica se estructuró en torno a tres sesiones presenciales semanales de 45 minutos. El contenido abordado fue el mismo en ambos casos y se correspondía con los ítems léxicos seleccionados previamente (véase el apartado 4.3.1), lo que permitió garantizar la equidad en la exposición temática y la coherencia didáctica.

4.5. Análisis de datos

El análisis de los datos se llevó a cabo mediante un enfoque mixto, que combinó técnicas estadísticas inferenciales y herramientas cualitativas para ofrecer una visión más completa de los resultados.

4.5.1. Evaluación de supuestos estadísticos

Para el análisis de datos se utilizó el programa SPSS (v. 25.0). Se comprobaron los supuestos de normalidad mediante las pruebas de Kolmogorov-Smirnov (K-S) y Shapiro-Wilk (S-W), y homocedasticidad con la prueba de Levene. En función de los resultados obtenidos, se recurrió a pruebas *t* de Student o, en caso de no normalidad, a la prueba de Mann-Whitney.

4.5.2. Parámetros de reporte

Se empleó un nivel de significación bilateral de $\alpha = 0.05$. La magnitud de los efectos se interpretó mediante:

1. *d* de Cohen, efecto pequeño (0.2), mediano (0.5) y grande (0.8).
2. *r* de Rosenthal, correlación débil (0.1), moderada (0.3) y fuerte (0.5).

4.5.3. Análisis cualitativo y estadístico descriptivo del cuestionario

El cuestionario fue elaborado a partir de modelos previamente validados y revisado mediante un proceso de validación de contenido y constructo. La fiabilidad interna se confirmó con un coeficiente alfa de Cronbach, superior a 0.80, considerado adecuado para instrumentos de corte actitudinal.

Además del análisis cualitativo basado en codificación temática emergente, se llevó a cabo un análisis estadístico descriptivo para identificar patrones generales en las respuestas. Este análisis incluyó:

1. el cálculo de la media, la desviación estándar y la distribución de frecuencias de cada ítem, con el fin de detectar tendencias centrales y variabilidad.
2. un análisis porcentual de las categorías de respuesta (“muy de acuerdo”, “de acuerdo”, “en desacuerdo” y “muy en desacuerdo”) para visualizar gráficamente la distribución de actitudes.

Este enfoque mixto —cuantitativo y cualitativo— permitió triangular la información, enriquecer la interpretación de los resultados y contextualizar la experiencia de aprendizaje mediada por IA.

4.6. Consideraciones éticas e integridad de datos

Se examinó la base de datos para verificar su integridad, sin encontrar valores perdidos ni atípicos extremos, lo que refuerza la validez interna y la fiabilidad estadística del estudio.

Asimismo, no se registraron abandonos durante el periodo de intervención, lo cual garantizó la estabilidad del diseño cuasiexperimental.

La participación del alumnado fue voluntaria y anónima, en conformidad con los principios éticos de la investigación educativa. La implementación del protocolo fue supervisada mediante plataformas con trazabilidad de uso en el caso del grupo experimental, y a través de observaciones presenciales sistemáticas en el grupo de control. Este control metodológico permitió asegurar una exposición homogénea al input léxico y condiciones de intervención comparables en ambos grupos.

5. RESULTADOS

Esta sección presenta los resultados obtenidos mediante el análisis cuantitativo y cualitativo de los datos recopilados durante la intervención.

5.1. Resultados cuantitativos

A continuación, se exponen los resultados cuantitativos obtenidos mediante las pruebas de evaluación aplicadas a ambos grupos.

5.1.1. Pruebas de normalidad

Las pruebas de normalidad K-S y S-W indicaron que las puntuaciones obtenidas en la prueba escrita y en la comprensión auditiva seguían una distribución normal ($p > .05$). En cambio, los resultados de la prueba de expresión oral del grupo de control no cumplían con este supuesto ($p < .05$), lo que justificó el uso de pruebas no paramétricas para dicha variable.

5.1.2. Comparación de medias: prueba t para muestras independientes

El test de homogeneidad de varianzas de Levene no arrojó diferencias estadísticamente significativas ($p > .05$), lo que permitió aplicar la prueba t de Student. En la prueba escrita, no se observaron diferencias estadísticamente relevantes entre los grupos ($t(50) = -.834$, $p = .409$, $d = .231$), lo que indica un efecto bajo o inexistente de la intervención en el rendimiento escrito. Tal como indica la Tabla 1, los valores medios entre ambos grupos se mantienen dentro de un rango similar, lo que refuerza la conclusión de que no hubo un impacto sustancial en esta habilidad.

En contraste, en la comprensión auditiva se encontraron diferencias estadísticamente significativas a favor del grupo experimental ($t(50) = -3.517$, $p = .001$, $d = .975$), lo que confirma un efecto positivo del uso de plataformas de IA en la comprensión auditiva del vocabulario. Como muestra la Tabla 1, la media del grupo experimental fue sensiblemente más alta, lo que sugiere una mejora clara en la adquisición receptiva del léxico.

Tabla 1. Media ($\pm$ DE), valores t y p de las pruebas escrita y auditiva por grupo

Prueba	Experimental (n = 26)	Control (n = 26)	t	p
Prueba escrita	92.50 $\pm$ 2.70	91.73 $\pm$ 3.85	-.834	.409
Prueba auditiva	87.08 $\pm$ 4.66	82.31 $\pm$ 5.11	-3.517	.001**

* $p < .05$ ** $p < .01$

5.1.3. Comparación no paramétrica: prueba U de Mann-Whitney

Dado que la variable de expresión oral del grupo de control no cumplía el supuesto de normalidad, se aplicó la prueba no paramétrica U de Mann-Whitney. Los resultados mostraron una diferencia estadísticamente significativa entre los grupos ($U = 178$, $Z = -2.945$, $p = .003$), con un tamaño del efecto moderado-alto ($r = .408$), según los criterios de Cohen (1988). Estos resultados indican que la intervención con plataformas de IA tuvo un efecto positivo significativo en el desarrollo de la expresión oral. La Tabla 2 muestra medianas consistentemente superiores en el grupo experimental, lo que respalda la eficacia de la práctica oral digitalizada.

Tabla 2. Mediana (P_{25} - P_{75}), valores U , Z y p de la prueba oral por grupo

Prueba	Experimental (n = 26)	Control (n = 26)	U	Z	p
Prueba oral	88.5 (84.0; 92.0)	83.0 (80.0; 86.3)	178	-2.945	.003**

* $p < .05$; ** $p < .01$

5.2. Resultados cualitativos: percepción estudiantil sobre el uso de plataformas de IA

En esta sección se presentan los principales hallazgos derivados del análisis cualitativo de las respuestas, las cuales reflejan la percepción general del alumnado respecto al uso de plataformas de IA en el proceso de aprendizaje.

5.2.1. Resultados generales de la encuesta (n = 52)

Los resultados del cuestionario revelan una actitud mayoritariamente positiva hacia el uso de herramientas basadas en IA. El 80,8 % manifestó estar “de acuerdo” o “muy de acuerdo” con que dichas plataformas mejoraban la eficiencia en el aprendizaje de vocabulario, y el 84,6 % percibió que las plataformas favorecían la retención de nuevo vocabulario. No obstante, se identificaron críticas relevantes: el 77 % consideró que los escenarios de vocabulario presentados no reflejaban sus contextos culturales cotidianos en China. Además, el 88,4 % señaló que las explicaciones socioculturales asociadas a los términos en español eran poco claras o insuficientes, y el 100 % manifestó que las plataformas no abordaban adecuadamente las diferencias gramaticales entre el español y el chino.

5.2.2. *Análisis de las respuestas abiertas*

En las respuestas cualitativas, los estudiantes propusieron diversas mejoras centradas en la adaptación cultural y lingüística. Entre las principales sugerencias se destacan:

1. La introducción de contenidos léxicos y situaciones comunicativas representativas de la cultura china, con el fin de fomentar una mayor conexión identitaria y contextual.
2. La incorporación de explicaciones culturales más claras y accesibles sobre los significados de los términos propios de los países hispanohablantes.
3. La implementación de mecanismos de corrección fonética personalizada, ajustada al perfil fonológico del usuario.

En suma, los resultados muestran mejoras notables en el rendimiento léxico y una actitud mayoritariamente favorable hacia el uso de plataformas de IA. No obstante, los comentarios cualitativos revelan áreas críticas de mejora, especialmente en lo relativo a la contextualización cultural y la adecuación lingüística. En consecuencia, estos hallazgos serán abordados en la siguiente sección.

6. DISCUSIÓN

Los resultados de esta investigación evidencian que el uso de plataformas basadas en IA puede favorecer el aprendizaje de vocabulario en ELE, particularmente en las dimensiones receptivas y orales del léxico. No obstante, su impacto en la precisión escrita se muestra moderado. A continuación, se discuten estos hallazgos a la luz de investigaciones previas, se examinan sus implicaciones pedagógicas y se abordan las limitaciones metodológicas del estudio.

6.1. Interpretación de los resultados

Tal como muestran los datos cuantitativos, el grupo experimental presentó mejoras estadísticamente significativas en la retención léxica auditiva y en el reconocimiento de vocabulario en contexto oral (véanse Tabla 1 y Tabla 2). Estos avances sugieren que la exposición multimodal —mediante imágenes, audio y ejemplos contextualizados— contribuye a facilitar la codificación del léxico y, en consecuencia, su posterior recuperación. Este resultado coincide con estudios previos que subrayan la eficacia de los entornos digitales enriquecidos para aprendientes de nivel inicial (Xing & Zhang, 2025).

En cambio, no se observaron progresos cuantificables en la producción escrita. Este hallazgo indica que, si bien la IA puede fortalecer competencias receptivas, no garantiza por sí sola la consolidación de las habilidades vinculadas a esta modalidad. Una posible explicación es la escasa inclusión de tareas abiertas o de redacción guiada en las plataformas utilizadas, lo que concuerda con investigaciones que destacan la importancia del diseño instruccional para el desarrollo de esta destreza (Kong & Qin, 2021). Además, muchas aplicaciones priorizan ejercicios cerrados y repetitivos que no estimulan procesos de composición más complejos, limitando así el dominio de estructuras gramaticales y ortográficas más avanzadas (Palomares Marín, 2022).

Desde un enfoque cualitativo, los datos recopilados complementan estos hallazgos. La mayoría del alumnado manifestó una valoración positiva del uso de IA, destacando su accesibilidad, claridad y eficacia. No obstante, también se señalaron deficiencias en la adaptación cultural de los contenidos, especialmente entre los estudiantes chinos, quienes percibieron una desconexión entre los materiales y su contexto sociocultural. La carencia de referentes culturales próximos en los materiales puede incidir negativamente en la motivación y restringir el aprendizaje léxico en situaciones comunicativas reales. Desde esta perspectiva, los resultados subrayan la necesidad de desarrollar entornos digitales que integren no solo *input* lingüístico, sino también contenidos culturalmente pertinentes y anclados en las realidades socioculturales del alumnado.

6.2. Implicaciones pedagógicas

A partir de los resultados obtenidos, se derivan implicaciones pedagógicas relevantes para la enseñanza de lenguas extranjeras en contextos caracterizados por una exposición limitada al idioma meta.

Una primera consideración apunta a la necesidad de potenciar las habilidades orales y auditivas mediante el uso de tecnologías basadas en inteligencia artificial. Estas herramientas permiten mitigar la escasez de *input* auténtico gracias a la oferta de estímulos auditivos variados y situaciones de interacción simulada. Tales entornos facilitan tanto la comprensión auditiva como la producción oral en contextos donde el contacto natural con la lengua meta es limitado. La literatura empírica respalda esta línea, señalando que los entornos virtuales pueden mejorar significativamente la retención del vocabulario receptivo a largo plazo (Franciosi et al., 2016).

Además, resulta fundamental concebir estas tecnologías como herramientas complementarias dentro de modelos pedagógicos híbridos, más que como sustitutos del docente. Su integración adecuada permite enriquecer los enfoques tradicionales a través de actividades adaptativas, retroalimentación inmediata y estrategias de gamificación. Estas características no solo intensifican la experiencia de aprendizaje, sino que también favorecen una interacción más significativa con el léxico cuando se articulan con un diseño instruccional estructurado (Teng & Zhang, 2023).

Por otro lado, la personalización del aprendizaje se perfila como un potencial pedagógico clave en entornos mediados por IA. Gracias al monitoreo automatizado del desempeño, estas plataformas permiten ajustar los contenidos y niveles de dificultad en función de las necesidades individuales del alumnado. Este enfoque favorece el aprendizaje autónomo y una atención más eficaz a la diversidad presente en el aula. Además, estudios recientes han demostrado que los sistemas con atributos antropomórficos tienden a generar una mayor implicación emocional, lo cual repercute positivamente en la adquisición de vocabulario (Liu et al., 2022).

Finalmente, no puede obviarse el papel que desempeña la motivación y la adecuación cultural en la eficacia del aprendizaje asistido por tecnología. Para maximizar el impacto de las plataformas de IA, es esencial implementar estrategias de localización cultural que adapten los contenidos léxicos, los contextos comunicativos y los referentes visuales a las realidades socioculturales del alumnado sinohablante. Como han señalado Kramsch (1993) y

Risager (2006), el proceso de adquisición lingüística está íntimamente vinculado al entorno cultural en que se desarrolla. En este sentido, la integración de materiales culturalmente pertinentes no solo incrementa la motivación, sino que también optimiza la eficacia de la enseñanza. Una aplicación pedagógica crítica de la IA requiere, por tanto, articular la innovación tecnológica con la sensibilidad intercultural y una planificación didáctica coherente.

6.3. Limitaciones del estudio

Pese a que este estudio ofrece evidencia empírica relevante sobre el impacto de las plataformas de IA en el aprendizaje léxico de ELE, conviene reconocer ciertas limitaciones que condicionan la interpretación y generalización de los resultados.

En primer lugar, la muestra estuvo compuesta exclusivamente por estudiantes universitarios sinohablantes, lo que restringe la aplicabilidad de los hallazgos a otros perfiles de aprendientes, niveles educativos o contextos culturales. Esta homogeneidad limita la validez externa del estudio y podría haber influido en la receptividad observada hacia las herramientas tecnológicas empleadas.

En segundo lugar, la investigación se centró únicamente en la dimensión léxica del aprendizaje, sin integrar otros componentes fundamentales como la gramática, la pragmática o la interacción discursiva. Su incorporación posibilitaría una evaluación más holística del impacto de la IA en la competencia comunicativa y enriquecería los modelos pedagógicos híbridos.

Por último, aunque las preguntas abiertas ofrecieron una perspectiva complementaria, su alcance resultó limitado. El empleo de técnicas cualitativas más robustas —como entrevistas semiestructuradas o grupos focales— habría facilitado una comprensión más matizada de las percepciones estudiantiles. La elección metodológica, aunque pragmática, redujo el potencial para alcanzar una saturación analítica significativa.

6.4. Direcciones para futuras investigaciones

A partir de las limitaciones señaladas, se proponen tres líneas de investigación prioritarias para avanzar en la comprensión del papel de la IA en el aprendizaje de lenguas extranjeras.

En primer lugar, es necesario profundizar en la dimensión sociocultural de los entornos digitales. La contextualización cultural de contenidos y referentes visuales podría influir de manera determinante en la apropiación del léxico y en la motivación, especialmente en contextos interculturales. Investigar esta variable permitiría desarrollar materiales más inclusivos y culturalmente pertinentes.

Por otro lado, conviene explorar el potencial de personalización adaptativa que ofrecen los sistemas basados en IA. El ajuste dinámico de las rutas de aprendizaje según el rendimiento académico y las características individuales del alumnado plantea nuevos horizontes para atender a la diversidad. De igual manera, incorporar una evaluación más holística, que contemple gramática, expresión escrita e interacción discursiva, facilitaría una valoración más integral del impacto tecnológico.

Finalmente, se requiere emprender estudios longitudinales que analicen los efectos sostenidos del uso de plataformas inteligentes sobre la retención léxica, la transferencia al

uso espontáneo del vocabulario y la progresión de la competencia lingüística. Este enfoque temporal ofrecería una base más sólida para evaluar la eficacia de las intervenciones tecnopedagógicas a medio y largo plazo.

7. CONCLUSIONES

Este estudio se inscribe en la creciente línea de investigación sobre el potencial transformador de la IA en el aprendizaje de lenguas extranjeras, aportando evidencia empírica sobre su impacto en el desarrollo léxico de estudiantes sinohablantes en niveles iniciales de ELE. Más allá de constatar la eficacia de determinadas funcionalidades tecnológicas —en especial en las habilidades receptivas y orales—, los resultados invitan a una reflexión crítica sobre las condiciones didácticas, culturales y epistemológicas necesarias para una integración significativa de estas herramientas.

Una de las principales contribuciones del trabajo radica en su enfoque metodológico mixto, sustentado en la triangulación entre datos cuantitativos y cualitativos en un entorno educativo auténtico. Esta aproximación ha permitido validar avances objetivos en el rendimiento léxico y, a su vez, visibilizar la dimensión afectiva del aprendizaje, lo que refuerza la necesidad de integrar la eficacia instrumental y la apropiación subjetiva en las propuestas tecnopedagógicas, en consonancia con el enfoque sociocognitivo del aprendizaje léxico.

Asimismo, la investigación revela desafíos estructurales como la limitada contextualización cultural de los materiales digitales. Esta carencia no es un simple vacío de contenido, sino una desconexión estructural entre la innovación tecnológica y los marcos interpretativos del aprendizaje intercultural. La inclusión de referentes socioculturales significativos constituye una condición esencial para garantizar procesos de aprendizaje inclusivos, motivadores y sostenibles.

Desde esta perspectiva, la IA no debe concebirse como una solución universal, sino como un mediador didáctico que puede potenciar o limitar la calidad del aprendizaje, en función del paradigma pedagógico en que se articule. Su verdadero valor reside en complementar la labor docente, siempre que se integre con rigor didáctico, sensibilidad intercultural y responsabilidad epistemológica que reconozca el papel activo del diseño tecnológico en la construcción del conocimiento lingüístico y cultural.

En conclusión, este trabajo contribuye no solo con datos empíricos aplicables al diseño de recursos digitales para ELE, sino también con una mirada reflexiva sobre las condiciones para una implementación pedagógicamente coherente y culturalmente consciente de la tecnología. Tales resultados permiten delinear una agenda de investigación orientada a una integración ética y enraizada en la diversidad sociocultural, que potencie el uso significativo de la IA en el aprendizaje de lenguas.

8. REFERENCIAS

- Boers, F., Warren, P., He, L., & Deconinck, J. (2017). Does adding pictures to glosses enhance vocabulary uptake from reading? *System*, 66, 113–129. <https://doi.org/10.1016/j.system.2017.03.017>
- Chen, I. J. & Yen, J. C. (2013). Hypertext annotation: Effects of presentation formats and learner proficiency on reading comprehension and vocabulary learning in foreign languages. *Computers and Education*, 63, 416–423. <https://doi.org/10.1016/j.compedu.2013.01.005>

- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Franciosi, S. J., Yagi, J., Tomoshige, Y., & Ye, S. (2016). The effect of a simple simulation game on long-term vocabulary retention. *CALICO Journal*, 33(3), 355–379. <https://doi.org/10.1558/cj.v33i2.26063>
- Halliday, M. A. K. (1985). *An introduction to functional grammar*. Edward Arnold.
- Hu, J. H. & Hu, J. S. (2020). Theoretical and paradigm evolution of foreign language education informatization in the past 70 years. *Technology Enhanced Foreign Language Education*, 1(1), 17–23. <https://doi.org/10.20139/j.issn.1001-5795.2020.01.001>
- Hu, Z. L. (2023). ChatGPT talk about foreign language education. *Foreign Languages in China*, 20(3), 1, 12–15. <https://doi.org/10.13564/j.cnki.issn.1672-9382.2023.03.003>
- Jewitt, C. (Ed.). (2009). *The Routledge handbook of multimodal analysis*. Routledge.
- Kong, L. & Qin, H. W. (2021). Digitalization of formative assessment data in foreign language teaching: Design and practice. *Technology Enhanced Foreign Language Education*, 2021(2), 57–64. <https://doi.org/10.20139/j.issn.1001-5795.2021.02.007>
- Kramsch, C. (1993). *Context and culture in language teaching*. Oxford University Press.
- Kress, G. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Edward Arnold.
- Lai, K. W. K. & Chen, H. J. H. (2021). A comparative study on the effects of a VR and PC visual novel game on vocabulary learning. *Computer Assisted Language Learning*, 36(3), 312–345. <https://doi.org/10.1080/09588221.2021.1928226>
- Lee, S. M. & Park, M. (2020). Reconceptualization of the context in language learning with a location-based AR app. *Computer Assisted Language Learning*, 33(8): 936–959. <https://doi.org/10.1080/09588221.2019.1602545>
- Li, W. W. (2023). AI-enabled reform in foreign language education: conceptual innovation and action logic. *China Higher Education*, (9), 49–52.
- Li, Z. W. (2024). ChatGPT for empowering foreign language teaching: Scenarios and strategies. *Journal of Beijing International Studies University*, 46(1), 109–118. <https://doi.org/10.12002/j.bisu.501>
- Liu, Q., Yang, H. J., & Zhang, X. Q. (2022). Language learning supported by Artificial Intelligence—How does anthropomorphic technology affect learning. *Technology Enhanced Foreign Language Education*, 43(10), 102–109. <https://doi.org/10.13811/j.cnki.eer.2022.10.014>
- Malinowski, B. (1923). The problem of meaning in primitive languages. In C. K. Ogden & I. A. Richards (Eds.), *The meaning of meaning* (pp. 296–336). Harcourt, Brace & World. <https://doi.org/10.4324/9781315010847-25>
- Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. Cambridge University Press.
- Mohamed, A. M. (2024). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a foreign language (EFL) teaching: Perceptions of EFL faculty members. *Educational Information Technologies*, 29, 3195–3217. <https://doi.org/10.1007/s10639-023-11917-z>
- Monteagudo Canales, A. (2024). Posibilidades y limitaciones de ChatGPT para el aprendizaje del español como LE/L2. *El Español por el Mundo*, 6(1), 35–45. <https://doi.org/10.59612/epm.i1.141>
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Newbury House.

- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nation, I. S. P. & Waring, R. (1997). Vocabulary size, text coverage, and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 6–19). Cambridge University Press.
- Palomares Marín, M. del M. (2022). *El español como lengua extranjera en aplicaciones adaptativas y multimedia: el caso de Duolingo* (Tesis doctoral). Universidad de Murcia. <http://hdl.handle.net/10201/115763>
- Risager, K. (2006). *Language and culture: Global flows and local complexity*. Multilingual Matters.
- Rosenthal, R. (1991). *Meta-Analytic Procedures for Social Research*. Sage.
- Sanz Manzanedo, M. (2025). La IA en la enseñanza de idiomas: chatbots y formación del profesorado. *European Public & Social Innovation Review*, 10(3). <https://doi.org/10.31637/epsir-2025-513>
- Schmitt, D. & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55–88. <https://doi.org/10.1177/026553220101800103>
- Teng, M. F. & Zhang, D. (2023). The associations between working memory and the effects of multimedia input on L2 vocabulary learning. *International Review of Applied Linguistics in Language Teaching*, 61(3), 1021–1049. <https://doi.org/10.1515/iral-2021-0130>
- Trier, J. (1931). *Der deutsche Wortschatz im Sinnbezirk des Verstandes: Die Geschichte eines sprachlichen Feldes* (Vol. 1). Winter.
- van den Broek, G. S. E., Takashima, A., Segers, E., & Verhoeven, L. (2018) Contextual richness and word learning: Context enhances comprehension but retrieval enhances retention. *Language Learning*, 68(2): 546–585. <https://doi.org/10.1111/lang.12285>
- Vesselinov, R. & Grego, J. (2016). *Duolingo effectiveness study: Final report*. Duolingo.
- Wang, Y. H. (2016). Promoting contextual vocabulary learning through an adaptive computer-assisted EFL reading system. *Journal of Computer Assisted Learning*, 32(4), 291–303. <https://doi.org/10.1111/jcal.12132>
- Xing, B. & Zhang, H. (2025). A study of the effect of multimodal input on vocabulary acquisition: Evidence from online Chinese language learners. *Language Teaching Research*. <https://doi.org/10.1177/13621688241313017>
- Zhang, Z. Y. & Hong, H. Q. (2023). ChatGPT-supported foreign language teaching: Empowerment, challenges, and strategies. *Foreign Language World*, (2), 38–44.
- Zhu, T., Zhang, Y. & Irwin, D. (2024). Second and foreign language vocabulary learning through digital reading: A meta-analysis. *Education and Information Technologies*, 29, 4531–4563. <https://doi.org/10.1007/s10639-023-11969-1>
- Zou, B. & Wang, M. J. (2021). Artificial Intelligence Technology and English Language Teaching and Learning: Present and Future. *Foreign Languages and Literature*, 37(3), 124–130.
- Zou, D., Huang, Y., & Xie, H. (2021) Digital game-based vocabulary learning: Where are we and where are we going? *Computer Assisted Language Learning*, 34(5–6): 751–777. <https://doi.org/10.1080/09588221.2019.1640745>