Humanidades, Tecnología y Ciencia, del Instituto Politécnico Nacional
Certificado de Reserva de Derechos al Uso Exclusivo del Título: 04-2010-0326124-14000-203 ISSN: 2007-1957

DESARROLLO DE APLICACIÓN DE ESCENARIO INTERACTIVO CON REALIDAD VIRTUAL COMO APOYO AL APRENDIZAJE UNIVERSITARIO DEL IDIOMA INGLÉS

Maribel Leyva Gaxiola
Universidad Tecnológica de Corregidora
maribel.leyva@utcorregidora.edu.mx

Andrea Montserrat Cruz Rodríguez
Universidad Tecnológica de Corregidora
2310615016@utcorregidora.edu.mx

Abstract

The present study aims to develop a virtual reality educational environment designed to facilitate the teaching and practice the English language among university students. This environment promotes a dynamic, interactive, and effective learning experience by leveraging the immersive advantages of virtual reality to enhance pronunciation, listening comprehension, and oral expression. The project addresses the need to transform traditional language teaching methods through the integration of emerging technologies that foster meaningful learning based on experience and contextualized practice. A quantitative methodology with an experimental design was applied, using diagnostic and final assessments with 14 university students at basic English levels to measure their communicative performance before and after interacting with the virtual environment. Throughout the development process, an analysis of students' linguistic difficulties informed the design of three-dimensional scenarios focused on real-life communicative situations, implemented in Unity with C# scripts for user navigation and 3D models representing everyday contexts such as a restaurant, an airport, and a job interview. Each environment included AI-controlled avatars capable of engaging in basic conversations with users, creating an immersive experience that encouraged real- time English interaction. The experimental sessions revealed that, despite initial language limitations, students showed a strong willingness to participate, increased

confidence, and improved fluency when communicating with virtual characters. The results demonstrate a significant rise in motivation to learn English and an overall improvement in communicative competence. The virtual experience, by integrating constant practice and immediate feedback within an immersive environment, enabled students to learn the language in a more natural, participatory, and motivating way, highlighting the transformative potential of immersive technologies in university education.

Palabras clave: Virtual reality, Teaching the English language, Educational technologies.

El proyecto propone un entorno educativo en realidad virtual (VR) para mejorar la enseñanza del inglés en estudiantes universitarios (Freina & Ott, 2015). A través de escenarios inmersivos que simulan situaciones cotidianas, los usuarios desarrollan habilidades de pronunciación, comprensión auditiva y expresión oral en un ambiente seguro y motivador (Vandergrift & Goh, 2012). Esta iniciativa responde a la necesidad de modernizar las estrategias de enseñanza del idioma, promoviendo un aprendizaje más interactivo, participativo y significativo. En conjunto, busca ofrecer una alternativa innovadora al modelo tradicional, integrando tecnología avanzada y metodologías centradas en el estudiante para formar profesionistas competentes en un contexto globalizado.

1. Marco Teórico

1.1. Teorías del aprendizaje de idiomas

El aprendizaje de una lengua extranjera ha sido abordado desde diversas teorías que explican cómo las personas adquieren y desarrollan habilidades lingüísticas, influyendo en los métodos y tecnologías de enseñanza. El conductismo, propuesto por Skinner (1957), sostiene que el aprendizaje se produce mediante la repetición y el refuerzo de estímulos y respuestas. En contraste, el innatismo, encabezado por Chomsky (1965), plantea que los seres humanos poseen una capacidad biológica innata para aprender lenguas a través de un “dispositivo de adquisición del lenguaje” que facilita la comprensión natural de reglas y estructuras lingüísticas.

1.2. Habilidades lingüísticas esenciales

El Marco Común Europeo de Referencia para las Lenguas (MCER) clasifica las habilidades lingüísticas en dos grupos (Council of Europe, 2001): productivas (hablar y escribir) y receptivas (leer y escuchar), cada una con estrategias de enseñanza específicas. La comprensión auditiva implica interpretar mensajes orales activamente, mientras que la lectura permite acceder a información e inferir significados y estructuras (Vandergrift, 2012). La pronunciación, parte de la expresión oral, se apoya, hoy en día, en tecnologías con retroalimentación visual y auditiva. Finalmente, la escritura requiere planificación y revisión, siendo esencial para la comunicación académica y profesional.

1.3. Concepto y características de la realidad virtual

La VR es una tecnología que crea entornos digitales tridimensionales inmersivos mediante dispositivos como visores y sensores de movimiento, siendo cada vez más accesible y utilizada en la educación. Su principal característica es la sensación de presencia, que integra los sentidos para generar experiencias realistas y participativas (Bacca, et al., 2014). Esta inmersión favorece el compromiso emocional y cognitivo del estudiante, permitiendo una construcción activa del conocimiento a través de la exploración, la toma de decisiones y la resolución de problemas dentro del entorno virtual.

1.4. Beneficios pedagógicos de la realidad virtual

La VR se ha consolidado como una herramienta educativa innovadora que promueve un aprendizaje más inmersivo, activo y personalizado, especialmente útil en la enseñanza del inglés como segunda lengua. Su principal ventaja pedagógica radica en la inmersión, que genera una sensación de presencia y concentración total, favoreciendo el compromiso cognitivo y emocional del estudiante (Bacca, et al., 2014).

Diversos estudios destacan la importancia de la retroalimentación inmediata, pues los entornos virtuales pueden corregir errores de pronunciación y expresión oral en tiempo real, mejorando la precisión y la confianza comunicativa (Chapelle & Sauro, 2017).

La VR también potencia el aprendizaje experiencial, coherente con los principios constructivistas de Piaget (1970) y Vygotsky (1978), al permitir que los estudiantes aprendan mediante la acción, la interacción y la reflexión dentro de contextos simulados. Asimismo, ofrece personalización y adaptación a distintos niveles y estilos de aprendizaje, lo que incrementa la motivación intrínseca y la autonomía.

Finalmente, al integrar estímulos visuales, auditivos y kinestésicos, la VR fortalece la retención y la memoria, tal como explica el modelo de codificación dual de Paivio (1991). En conjunto, estas ventajas convierten a la realidad virtual en una herramienta pedagógica integral que mejora las competencias lingüísticas y fomenta un aprendizaje más significativo, motivador y duradero.

1.5. Meta Quest 3

El Meta Quest 3, desarrollado por Meta Platforms, es un dispositivo de realidad virtual autónomo que combina potencia, portabilidad y precisión, posicionándose como una herramienta ideal para la educación inmersiva (Meta Platforms Inc., 2023). Gracias a su procesador Snapdragon XR2 Gen 2, ofrece un alto rendimiento gráfico y baja latencia, evitando mareos y garantizando una experiencia fluida. Su seguimiento de manos y movimiento permite una interacción natural con objetos virtuales, favoreciendo el aprendizaje práctico del inglés mediante gestos y manipulación de elementos. Al ser autónomo, no requiere conexión a un ordenador, lo que facilita su uso en espacios educativos sin infraestructura compleja. Pedagógicamente, permite crear escenarios tridimensionales y conversacionales con inteligencia artificial y audio espacial, mejorando la comprensión y pronunciación del idioma. Además, su modo de realidad mixta integra el entorno físico con el virtual, ampliando las posibilidades educativas.

El dispositivo es compatible con Unity y otras herramientas de desarrollo, lo que posibilita

diseñar contenidos adaptados al contexto local (*Unity Technologies*, 2023). También incluye sistemas de monitoreo de uso y rendimiento, útiles para la evaluación formativa. Su diseño ergonómico promueve el confort en sesiones prolongadas, aunque su costo elevado y la escasez de contenidos educativos en español son limitaciones actuales. Por ello, la creación de aplicaciones propias resulta esencial para aprovechar su potencial en la enseñanza universitaria del inglés.

1.6. Unity

Unity es una plataforma de desarrollo multiplataforma ampliamente utilizada para crear aplicaciones de realidad virtual, aumentada y mixta. Su motor gráfico avanzado y su lenguaje de programación C# permiten diseñar entornos tridimensionales inmersivos, interactivos y con gran realismo, lo que la convierte en una herramienta ideal para proyectos educativos como el destinado a la enseñanza del inglés en contextos universitarios (*Unity Technologies*, 2023).

La compatibilidad con múltiples dispositivos, especialmente con Meta Quest 3, y su soporte para APIs de VR como OpenXR y el SDK de Meta, facilitan la integración de funciones como el seguimiento ocular, control por gestos y sincronización labial (*Meta Platforms Inc.,* 2023). Además, su *Asset Store* provee modelos, sonidos y *scripts* que agilizan el desarrollo sin necesidad de crear cada elemento desde cero.

Desde la perspectiva pedagógica, *Unity* permite incorporar estrategias interactivas y gamificadas, simulando situaciones reales que promueven el aprendizaje activo, la práctica comunicativa y el pensamiento crítico. También posibilita integrar inteligencia artificial y aprendizaje automático, permitiendo crear avatares conversacionales adaptativos —como los implementados mediante Convai— que enriquecen la experiencia del estudiante.

Asimismo, Unity ofrece herramientas para evaluar el desempeño del usuario mediante el registro de datos y métricas que pueden ser analizadas por los docentes para ajustar estrategias de enseñanza. Aunque su curva de aprendizaje puede ser pronunciada y la optimización para dispositivos autónomos exige conocimientos técnicos avanzados, su versatilidad, comunidad activa y constante actualización consolidan a *Unity* como una plataforma esencial para la innovación educativa en entornos de realidad virtual.

1.7. Convai

Convai es una plataforma de inteligencia artificial conversacional diseñada para crear avatares interactivos y realistas en entornos de realidad virtual y aumentada (Convai Inc, 2024). Su tecnología, basada en procesamiento del lenguaje natural (PLN) y aprendizaje automático, permite que los personajes virtuales mantengan diálogos naturales y contextuales con los usuarios, lo que la convierte en una herramienta ideal para la enseñanza del inglés en entornos inmersivos.

La plataforma se integra fácilmente con Unity, posibilitando la sincronización labial, el seguimiento de la mirada y la animación facial, elementos clave para mejorar la pronunciación, entonación y realismo comunicativo (Unity Technologies, 2023). Además, ofrece funciones de personalización de diálogos y retroalimentación automática, lo que permite crear escenarios educativos adaptativos, como entrevistas o conversaciones cotidianas.

Convai también recopila datos de desempeño del usuario (errores, tiempo de interacción,

progreso), útiles para la evaluación formativa y la personalización del aprendizaje (Convai Inc, 2024). Entre sus beneficios destacan la práctica oral autónoma, la reducción de la ansiedad comunicativa y la retroalimentación inmediata, aunque enfrenta desafíos como la comprensión de acentos o expresiones idiomáticas y la necesidad de equilibrar su uso con la interacción humana real (Chapelle & Sauro, 2017).

En conjunto, Convai representa una herramienta innovadora que potencia el aprendizaje de lenguas mediante interacciones inmersivas, adaptativas y emocionalmente realistas, integrando inteligencia artificial y realidad virtual en la educación (Convai Inc, 2024).

1.8. Evaluación comparativa

La evaluación comparativa en el aprendizaje de idiomas mediante realidad virtual (VR) es una herramienta esencial para analizar la eficacia de las tecnologías inmersivas frente a los métodos tradicionales (Freina & Ott, 2015).

Este enfoque permite medir tanto resultados académicos (como mejoras en pronunciación, comprensión auditiva y fluidez oral) como factores cualitativos (motivación, confianza y percepción del estudiante) (Vandergrift & Goh, 2012).

Investigaciones como las de Freina & Ott (2015) y Bacca et al. (2014) han demostrado que la VR favorece una mayor participación, reduce la ansiedad al hablar y genera experiencias de aprendizaje más auténticas, simulando interacciones comunicativas reales. Además, la evaluación comparativa considera variables contextuales (nivel previo, experiencia tecnológica y apoyo docente) y la duración de la exposición, ya que estos factores pueden influir significativamente en los resultados.

Este tipo de evaluación también permite identificar los elementos del entorno virtual que más contribuyen al aprendizaje —como los avatares conversacionales, la interactividad o la gamificación—, promoviendo un diseño educativo basado en evidencias. En el caso del proyecto universitario descrito, se propone medir el progreso de los estudiantes antes y después del uso del entorno VR y compararlo con un grupo control, utilizando tanto pruebas de desempeño lingüístico como cuestionarios de percepción.

2. Metodología

2.1. Tipo de investigación

La investigación fue de tipo descriptiva y aplicada. Descriptiva ya que se analizaron y detallaron las características de la problemática educativa relacionada con la enseñanza del inglés, y aplicada porque se propuso y desarrolló una solución funcional y tecnológica.

2.2. Diseño metodológico y fases del proyecto

El desarrollo se dividió en seis fases principales:

Fase 1: Planeación

1. Definir objetivos

Se establecieron los objetivos generales y específicos del proyecto, con base en la necesidad de mejorar la enseñanza del idioma inglés en estudiantes universitarios mediante un entorno inmersivo.

2. Definir requerimientos del usuario

Se identificaron las necesidades del usuario final mediante entrevistas breves a estudiantes

universitarios, quienes señalaron la dificultad de practicar inglés conversacional de forma dinámica.

Fase 2: Investigación Técnica

3. Investigar herramientas y APIs de reconocimiento de voz

Se realizó una revisión comparativa de APIs que permiten reconocimiento de voz y comprensión del lenguaje natural. Finalmente, se seleccionó la API de Convai, ya que permite integrar fácilmente personajes virtuales con capacidades conversacionales realistas, incluyendo reconocimiento de intención, contexto y generación de respuesta natural, lo cual se alineaba perfectamente con el propósito educativo del proyecto.

Fase 3: Diseño

Fase 4: Desarrollo

5. Integrar la API

Se conectó la aplicación desarrollada en Unity con la API de Convai. Para esto se configuró el personaje virtual dentro de Convai Studio, entrenando sus respuestas y personalidad. El personaje pudo reaccionar con animaciones, gestos y sincronización labial según la respuesta generada por la IA.

6. Programación

Se utilizó un sistema de *raycasting* para detectar la zona objetivo, y mediante un script en C#, se actualizó la posición del XR Rig al punto seleccionado. Esta técnica fue preferida por encima del desplazamiento continuo, con el fin de evitar mareos y mantener la experiencia accesible a todos los usuarios. Además, se realizó la integración con scripts que permiten teletransportar al usuario por los diferentes escenarios.

Fase 5: Pruebas e Iteración

7. Pruebas internas de la APK

Se generaron *builds* preliminares del proyecto para dispositivos Meta Quest 3 y se realizaron pruebas internas para asegurar la funcionalidad del reconocimiento de voz, tiempo de respuesta de la API, y comportamiento general de la aplicación. Se identificaron errores como fallos desincronización labial o desconexiones temporales con la API.

8. Corrección de errores

Se realizaron ajustes en el código para mejorar la estabilidad, como optimizar las animaciones para evitar caídas de FPS. También, se modificaron los *prompts* del personaje en Convai Studio para mejorar su comprensión en distintos contextos.

Fase 6: Evaluación y Resultados

9. Pruebas con usuarios

Se aplicaron pruebas con estudiantes universitarios para observar su interacción con el personaje. Se hicieron anotaciones para analizar reacciones, fluidez de conversación y nivel de participación. Se utilizó una lista de control que evaluaba la naturalidad de la conversación, claridad de interfaz y utilidad percibida en la práctica del idioma.

10. Recolección de resultados

Se recopilaron datos cualitativos a partir de entrevistas posteriores a la experiencia, así como observaciones directas. Los usuarios señalaron una mayor motivación para practicar el idioma debido a la interactividad del personaje, aunque también expresaron sugerencias para mejorar las respuestas contextuales del avatar.

11. Análisis de resultados

Se realizó un análisis descriptivo de la información obtenida y se identificaron patrones en la participación de los estudiantes. Esto permitió validar la hipótesis de que la interacción con personajes conversacionales en entornos virtuales mejora el interés por el aprendizaje de un segundo idioma.

3. Resultados

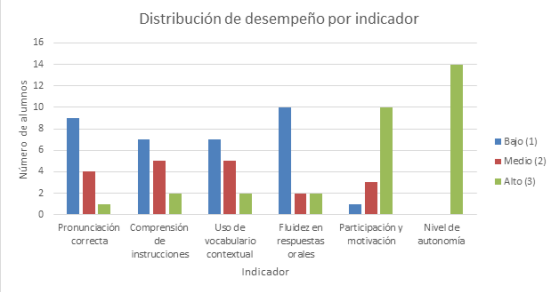
Durante la fase de evaluación del proyecto, se llevó a cabo una prueba piloto con una muestra de catorce estudiantes universitarios, cuyo objetivo fue analizar las reacciones iniciales y el desempeño comunicativo de los participantes al interactuar en un entorno de realidad virtual (VR) diseñado para la práctica del idioma inglés. El escenario, desarrollado en Unity e integrado con la API de Convai, recreaba situaciones cotidianas que incentivaban el uso espontáneo del idioma en un contexto seguro, inmersivo y libre de ansiedad social. Cada estudiante contó con un intervalo de cinco minutos para mantener una conversación con un avatar conversacional inteligente, capaz de procesar el lenguaje natural y responder de forma contextualizada mediante voz sintética.

Los resultados obtenidos fueron significativos en términos de motivación, participación y disposición hacia la práctica oral. A pesar de que la mayoría de los estudiantes cuenta con un nivel inicial bajo en inglés, especialmente en pronunciación y fluidez verbal, la experiencia inmersiva promovió una actitud activa frente al aprendizaje. La interacción con el personaje virtual generó un entorno donde los participantes se sintieron más cómodos para comunicarse, lo que sugiere que la realidad virtual, combinada con inteligencia artificial conversacional, puede reducir las barreras emocionales y la ansiedad comunicativa que suelen limitar el desempeño oral en contextos presenciales tradicionales. Este hallazgo coincide con estudios previos que destacan el potencial motivacional y afectivo de los entornos inmersivos para la enseñanza de lenguas extranjeras. La evaluación del desempeño se realizó mediante una rúbrica analítica que contempló seis indicadores principales: pronunciación correcta, comprensión de instrucciones, uso contextual del vocabulario, fluidez en la expresión oral, motivación y participación, y nivel de autonomía. Los resultados mostraron bajas puntuaciones en los indicadores lingüísticos técnicos (pronunciación y fluidez), pero niveles medios y altos en participación, motivación y autonomía. Estos datos sugieren que el entorno virtual contribuyó significativamente al compromiso y la autoconfianza del estudiante, factores esenciales para un aprendizaje comunicativo efectivo. Asimismo, la naturaleza interactiva del escenario permitió observar una mayor disposición al ensayo y error, favoreciendo una práctica más natural del idioma y el desarrollo gradual de la competencia comunicativa.

No obstante, es importante reconocer las limitaciones metodológicas de esta primera fase. La aplicación se restringió a un único escenario de interacción y a un tiempo de exposición limitado, lo que impidió evaluar progresos longitudinales en el aprendizaje del idioma. Además, la metodología utilizada se basó en observaciones cualitativas, sin incluir pruebas estandarizadas que permitieran medir de manera objetiva los avances en las habilidades lingüísticas. Para futuras etapas del proyecto, se recomienda ampliar el número de escenarios, aumentar la duración de las sesiones y complementar la evaluación con instrumentos cuantitativos, como exámenes de competencia o pruebas de pronunciación asistidas por *software*,

con el fin de obtener resultados más robustos y comparables.

Tabla 3 Resultados de pruebas. Elaboración propia



4. Conclusiones

La investigación demostró que el uso de entornos inmersivos de realidad virtual (VR) en la enseñanza del inglés genera mejoras significativas en la fluidez oral, la comprensión auditiva y la motivación de los estudiantes universitarios. La interacción con escenarios realistas y avatares conversacionales favoreció la producción espontánea del lenguaje, redujo la ansiedad comunicativa y fortaleció la confianza al hablar, validando la eficacia pedagógica de la VR frente a los métodos tradicionales.

Asimismo, la exposición a diálogos naturales y contextos auténticos potenció la comprensión auditiva y el vocabulario receptivo, mientras que la interactividad y el carácter lúdico del entorno incrementaron el interés y el compromiso con el aprendizaje. En conjunto, estos resultados confirman que la realidad virtual favorece tanto el rendimiento lingüístico como la dimensión emocional del proceso educativo.

Se recomienda que las instituciones educativas integren tecnologías inmersivas en sus programas de idiomas, acompañadas de formación docente y evaluaciones longitudinales que midan el progreso real de los estudiantes. Finalmente, se propone ampliar futuras investigaciones hacia el desarrollo de avatares con inteligencia artificial y la inclusión de poblaciones con necesidades educativas especiales, consolidando la VR como una herramienta transformadora en la enseñanza de lenguas extranjeras.

5. Referencias

Bacca, J., Baldiris, S., Fabregat, R., Graf, S., & Kinshuk. (2014). *Augmented reality trends in education: A systematic review of research and applications. Educational Technology & Society, 17*(4), 133–149.

Chapelle, C. A., & Sauro, S. (2017). *The handbook of technology and second language teaching and learning.* Wiley-Blackwell.

Chomsky, N. (1965). *Aspects of the theory of syntax.* MIT Press.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment.* Cambridge University Press.

Convai Inc. (2024). *Convai AI platform: Technical documentation and integration guide.* Convai Developer Portal.

Freina, L., & Ott, M. (2015). *A literature review on immersive virtual reality in education: State of the art and perspectives. eLearning & Software for Education, 1*, 133–141.

Meta Platforms Inc. (2023). *Meta Quest 3: Product specifications and developer overview.* Meta Developer Documentation.

Paivio, A. (1991). *Dual coding theory: Retrospect and current status. Canadian Journal of Psychology, 45*(3), 255–287.

Piaget, J. (1970). *Psychology and epistemology: Towards a theory of knowledge.* Penguin Books.

Skinner, B. F. (1957). *Verbal behavior.* Appleton-Century-Crofts.

Unity Technologies. (2023). *Unity manual: XR development with OpenXR and Meta SDK.* Unity Technologies Documentation.

Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action.* Routledge.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.