

# Software para el aprendizaje de la lectoescritura basado en control gestual de manos y realidad aumentada

## Software for literacy learning based on hand gesture control and augmented reality

Raúl de Jesús Sánchez Martínez\*<sup>ID</sup>, María Antonieta Abud Figueroa<sup>ID</sup>, Ulises Juárez Martínez<sup>ID</sup>  
Hilarión Muñoz Contreras<sup>ID</sup> y Lisbeth Rodríguez Mazahua<sup>ID</sup>

División de Estudios de Posgrado e Investigación  
Tecnológico Nacional de México Campus Orizaba/Instituto Tecnológico de Orizaba  
Oriente 9, No. 852, Col. Emiliano Zapata, Orizaba, Veracruz, México, C.P. 94320.  
[\\*m15011248@orizaba.tecnm.mx](mailto:m15011248@orizaba.tecnm.mx)

### PALABRAS CLAVE:

Interacción Humano-Máquina, Realidad Aumentada, Control Gestual de Manos, Lectoescritura, Interfaces Humano-Máquina

### RESUMEN

Con los avances tecnológicos, distintos proyectos desarrollaron soluciones satisfactorias a problemáticas de aprendizaje principalmente relacionadas con el área de las matemáticas y la medicina utilizando realidad aumentada e interfaces humano-máquina, siendo un área de oportunidad la aplicación de estas tecnologías en el aprendizaje de la lectoescritura. Dado que según la última prueba PISA, aplicada en 2018, México se sitúa en el nivel 2 de comprensión lectora, por debajo del promedio de los países miembros de la Organización para la Cooperación y el Desarrollo Económico (OCDE), en respuesta a esta problemática, se desarrolló una herramienta que combina la realidad aumentada e interfaces humano-máquina aplicado a la lectoescritura. Esta aplicación despliega modelos tridimensionales renderizados a través de un dispositivo con cámara, permitiendo a los usuarios interactuar con estos modelos mediante el reconocimiento de movimientos de las manos. La aplicación fue probada exitosamente en estudiantes de jardín de niños divididos en dos grupos: rezagados y adelantados en el aprendizaje de las letras. Los resultados demostraron que esta herramienta es entretenida y efectiva, proporcionando un recurso significativo para profesores y maestros en la enseñanza de la lectoescritura.

### KEYWORDS:

Human-Machine Interaction, Augmented Reality, Hand Gesture Control, Literacy, Human-Machine Interfaces

### ABSTRACT

With the progression of technology, several projects have devised satisfactory solutions to learning problems mainly related to the area of mathematics and medicine using augmented reality and human-machine interfaces, being an area of opportunity the application of these technologies to literacy learning. Based on the latest PISA test applied in 2018 in Mexico, this country is rated at level 2 in reading comprehension, falling below the average of member countries in the Organization for Economic Cooperation and Development (OECD). To address this issue, a tool combining augmented reality and human-machine interfaces focused on literacy learning was developed. This application projects three-dimensional models rendered via a camera device, enabling users to engage with these models through hand movement recognition. The tool was effectively trialed on kindergarten students divided into two groups: those lagging behind and those advanced in letter learning. The outcomes indicated that this instrument is both engaging and efficient, offering a substantial asset for teachers in literacy learning.

• Recibido: 30 de junio 2023

• Aceptado: 31 de agosto 2023

• Publicado en línea: 30 de octubre de 2023



## 1. INTRODUCCIÓN

La prueba PISA (*Programme for International Student Assessment*, Programa Internacional para la Evaluación de Estudiantes) 2018 reveló que, entre los miembros de la OCDE, México ocupa una posición inferior en resultados promedio en cuanto a comprensión lectora se refiere [1]. Los estudios presentados en este documento sugieren que la habilidad para comprender textos complejos está estrechamente vinculada con la capacidad de abordar estructuras de pensamiento igual de complejas, por tanto, las habilidades de lectoescritura se convierten en un indicador confiable de la inteligencia de un individuo.

Las tecnologías de realidad aumentada (RA) han experimentado un crecimiento acelerado en la última década, principalmente impulsadas por su incorporación en el desarrollo de videojuegos. La posibilidad de superponer imágenes virtuales sobre la realidad física desencadena un sinfín de oportunidades, especialmente en el ámbito del aprendizaje, ya que permite la recreación de escenarios del mundo real en entornos virtuales libres de consecuencias. Este potencial ha propiciado su rápida adopción en sectores como la industria y la educación. También, existen desarrollos de interfaces de usuario que permiten una interacción directa con estos escenarios virtuales, facilitando su análisis. Estas interfaces, que generalmente operan mediante la captura de movimiento, posibilitan una comunicación efectiva entre los operadores humanos y entornos virtuales.

En este artículo se describe un sistema que utiliza realidad aumentada, además de una interfaz humano-máquina para manipular los modelos virtuales, en apoyo al aprendizaje de las habilidades más básicas de la lectoescritura en niños de preescolar.

El resto de este documento está presentado de la siguiente manera: la sección 2 describe los trabajos relacionados con el presente proyecto, mientras que el diseño y desarrollo del sistema propuesto son descritos en la

sección 3. La sección 4 describe los resultados obtenidos de la aplicación del software desarrollado a un grupo de estudiantes de jardín de niños. Por último, las conclusiones y el trabajo a futuro se discuten en la sección 5.

## 2. TRABAJOS RELACIONADOS

A continuación, se presentan los trabajos más importantes relacionados con este proyecto.

La geometría es un pilar fundamental de las matemáticas, sin embargo, muchos estudiantes consideran su aprendizaje como una tarea desafiante y poco agradable. En el primer trabajo analizado [2] se propuso un marco de trabajo para el aprendizaje de geometría utilizando software de realidad aumentada y reconocimiento de gestos manuales a través de interfaces humano-máquina. En su investigación, emplearon principalmente el dispositivo Leap Motion para la detección de gestos. Para el reconocimiento de marcadores de posición, así como la captura de video a través de una cámara y la renderización de modelos tridimensionales sobre el video capturado, se utilizaron las bibliotecas de ARToolkit.

En el segundo estudio examinado [3], los autores enfatizan la importancia de la percepción espacial para la comprensión de conceptos matemáticos complejos. En respuesta a esto, desarrollaron una aplicación de Android llamada cleARmaths, orientada a la enseñanza de geometría vectorial, basada en tecnología de realidad aumentada. Para su creación se empleó Unity, que tiene la capacidad para capturar y renderizar imágenes en tiempo real. Este recurso permitió desarrollar un código de comportamiento en clases que heredan de la clase Monobehaviour. En lo que respecta al seguimiento de los marcadores de imagen, se utilizó el SDK (*Software Development Kit*) Vuforia.

En el estudio [4], se creó una red neuronal convolucional diseñada para reconocer expresiones humanas y utilizarlas como instrucciones, aplicada en este caso al aprendizaje de anatomía. Los autores propusieron un proceso de entrenamiento para la aplicación que recurre a la computación en la nube, lo que facilita su interacción con diferentes dispositivos y agiliza el proceso de desarrollo. Para este proyecto, se utilizó la cámara de profundidad RealSense que, mediante su proyector láser infrarrojo, permite detectar la forma de los objetos escaneados y generar un modelo 3D utilizando un método de representación de superficies basado en puntos clave. Para visualizar los modelos generados por la red neuronal, se recurrió a la biblioteca de Python Tensorflow. Esta permitió la conexión con el dispositivo HoloLens, que posee la tecnología necesaria para rastrear movimientos de las manos a través de su cámara integrada. Además, su pantalla interactiva permite superponer la red neuronal sobre la imagen real de fondo.

El trabajo [5] introduce MagicHand, una herramienta que permite controlar dispositivos del Internet de las Cosas (IoT) mediante gestos manuales y usando realidad aumentada. Esta aplicación aprovecha la interacción inmersiva que la realidad aumentada ofrece para mejorar la manipulación visual de los modelos interactivos en dispositivos IoT. MagicHand, que cuenta con un prototipo de programa para facilitar la comunicación entre estos dispositivos y una interfaz IoT, permite interacciones a través de sistemas de sonido e iluminación. La interpretación de los movimientos manuales del usuario se logra mediante una red neuronal artificial. Adicionalmente, MagicHand cuenta con una cámara que puede leer códigos QR para crear objetos virtuales y reconocer figuras geométricas para su análisis. Aunque actualmente los dispositivos IoT se manejan principalmente a través de interfaces web y móviles, MagicHand se probó exitosamente en dos dispositivos móviles.

En el ámbito de la telemedicina, donde el funcionamiento adecuado de los dispositivos médicos y la interacción efectiva entre el médico y el paciente son esenciales, la realidad aumentada ha demostrado su potencial. En el estudio [6] se empleó la realidad aumentada para desarrollar una aplicación de telemedicina destinada a entrenar a usuarios sin experiencia en la realización de electrocardiogramas. En esta aplicación, se colocaron puntos de control en un maniquí que, conectados al dispositivo, facilitaban el reconocimiento espacial mediante una cámara. Además, se incorporaron cajas de texto e indicaciones audibles para guiar al usuario durante la secuencia de la prueba, así como comandos de voz para mejorar la operatividad del sistema. Tras probar la aplicación con 10 voluntarios sin experiencia previa, se encontró que la realidad aumentada promovió significativamente su aprendizaje. Esto se confirmó mediante una prueba posterior con pacientes reales, en la que la diferencia en la colocación de los electrodos del electrocardiograma fue de solo 4 milímetros entre la aplicación y la prueba real.

La importancia de la fonética en el aprendizaje del idioma inglés a menudo es subestimada por los estudiantes. La falta de comprensión de los sonidos lingüísticos puede generar errores de pronunciación que, a su vez, pueden afectar la comunicación oral y escrita. Por ello, en el trabajo [7] se presentó un modelo de enseñanza basado en realidad aumentada para facilitar el aprendizaje de la fonética en inglés. Para el desarrollo de este modelo, se utilizó la aplicación Aurasma, ahora renombrada como HP Reveal. Este recurso tecnológico, que combina objetos virtuales y videoclips en un entorno simultáneamente real y virtual, destaca la relevancia del estímulo físico en el aprendizaje de conceptos sencillos. A diferencia de otros proyectos que emplearon códigos QR, este modelo optó por el uso de marcadores de imágenes de uso libre obtenidos de Internet, editados en forma de

tarjetas relámpago. Al visualizar estas tarjetas en la pantalla de un dispositivo móvil, se muestra un modelo 3D de la imagen en la tarjeta, las palabras correspondientes y una grabación de la pronunciación correcta.

La integración de tecnología en el aprendizaje del inglés como segundo idioma, especialmente en lo que respecta a las habilidades de escritura, se explora con profundidad en el trabajo [8]. En este estudio, se diseñó una aplicación de realidad aumentada para la práctica de escritura ubicua consciente del contexto, con el objetivo de aumentar la motivación y reforzar la memoria a largo plazo. Se realizó una comparación entre los trabajos de escritura producidos y las percepciones de los estudiantes utilizando la aplicación frente a las obtenidas mediante el método clásico de enseñanza. Ambos enfoques arrojaron diferencias significativas en la percepción de los estudiantes, evidenciando que la aplicación desarrollada promovió un mejor recuerdo a largo plazo, incrementó la motivación y mejoró la percepción del proceso de escritura. El estudio concluyó que la herramienta creada debería integrarse en cursos de inglés a distancia para mejorar los resultados de aprendizaje. Además, se incluyó un procedimiento en cinco pasos para ayudar a los estudiantes de inglés como lengua extranjera a incorporar la práctica de la escritura ubicua en sus métodos de aprendizaje. Para el desarrollo de la aplicación se utilizó el software Unity.

La realidad aumentada brinda la posibilidad de interactuar con objetos virtuales dentro de un contexto real, gracias a su habilidad para integrar ambos en un ambiente interactivo. Para proporcionar una interacción intuitiva y natural en un entorno de RA, las manos se han explorado como medio de interacción. En el estudio [9] se propone la creación de un entorno virtual donde los objetos pueden ser manipulados a través de gestos manuales en RA. Para lograr esto, se llevó a cabo una revisión de los trabajos existentes enfocados en la manipulación mediante gestos

manuales, seguido del uso de Leap Motion para capturar las expresiones manuales del usuario que permitirán la manipulación de los objetos en la aplicación de RA desarrollada. La fase de evaluación se realizó considerando la percepción del usuario, lo cual permitió determinar que la herramienta cumple con su propósito de brindar experiencias satisfactorias en la manipulación de modelos 3D mediante la gestión de gestos a través de Leap Motion. Para la implementación de la RA, se utilizó el SDK Vuforia.

En los casos analizados, la mayoría de las veces que se utilizó realidad aumentada, la plataforma de desarrollo seleccionada fue Unity, mientras que Leap Motion, en promedio, fue seleccionada como la interfaz humano-máquina para interacción con los objetos virtuales, en función de su costo accesible en comparación con otros dispositivos como Google Lens y HoloLens.

### 3. SOLUCIÓN PROPUESTA

En esta sección se describen las características del sistema propuesto, además de discutirse diferentes aspectos tecnológicos que sentaron las bases para su desarrollo.

#### 3.1. Software Development Kit de Realidad Aumentada

Un SDK es una colección de herramientas de desarrollo reunidas en un paquete de instalación único. Estas herramientas son específicas para una plataforma y sistema operativo determinados, a menudo toman la forma de Interfaces de Programación de Aplicaciones (*Application Programming Interfaces*, API), que son bibliotecas de funciones reutilizables para un lenguaje de programación específico. Normalmente, los SDK se incluyen en Entornos de Desarrollo Integrados (*Integrated Development Environment*, IDE) diseñados para plataformas específicas [10].

Vuforia Engine es la plataforma más ampliamente utilizada para el desarrollo de programas de RA, dada su gran compatibilidad con una multitud de dispositivos móviles. Permite a los desarrolladores integrar fácilmente funcionalidades de captura de cámara en sus dispositivos, para desplegar objetos de RA con los que los usuarios pueden interactuar de manera realista a través de la pantalla. Ofrece una extensa variedad de objetos, imágenes y entornos para la identificación de marcadores y el despliegue de modelos basados en ellos. Asimismo, permite el reconocimiento de formas de objetos utilizando modelos tridimensionales preexistentes en su propia biblioteca de imágenes y proyecta la RA basándose en estos reconocimientos. También cuenta con un *plugin* que facilita su integración con el motor de videojuegos Unity3D y Unreal Engine, permitiendo a los desarrolladores utilizar su interfaz para crear entornos de RA. Aunque Vuforia ofrece una versión gratuita, esta tiene limitaciones en cuanto al uso de la lectura de marcadores y áreas [11].

### 3.2. Interfaz de control

Las interfaces de control, también conocidas como interfaces humano-máquina, son sistemas que permiten la interacción entre el usuario y los objetos virtuales creados por un motor de gráficos. Estas interfaces pueden tomar varias formas, incluyendo controladores físicos, dispositivos de seguimiento de movimiento, reconocimiento de gestos, o incluso sistemas de control por voz.

El propósito de estas interfaces es permitir a los usuarios interactuar de forma intuitiva y eficiente con el entorno virtual, proporcionando una sensación de inmersión y control. Los avances en tecnología de seguimiento de movimiento y reconocimiento de gestos permiten interfaces más sofisticadas y naturales, lo que a su vez contribuye a la popularidad creciente de las aplicaciones de realidad aumentada y virtual.

Leap Motion Controller es un dispositivo avanzado de rastreo de manos y dedos que permite la interacción natural con contenido digital. Se basa en un modelo de seguimiento de puntos articulares del esqueleto, facilitando una representación detallada y precisa del movimiento de las manos. Este dispositivo es compatible con varios motores de desarrollo, incluyendo Unity3D y Unreal Engine, gracias a sus *plugins* disponibles. Aunque su SDK se centra en la programación en C#, también soporta lenguajes como C++, Java, JavaScript, Python y Objective-C en sus versiones más antiguas, aunque estos ya no reciben actualizaciones de soporte [12].

Ultra Leap, la empresa detrás de Leap Motion, ofrece varios dispositivos de seguimiento diseñados para distintas industrias. Sin embargo, para el uso de Leap Motion específicamente, solo se requiere una licencia personal, que es gratuita para uso no comercial. Esta accesibilidad hace de Leap Motion una opción popular para proyectos de realidad aumentada y virtual que buscan una interacción intuitiva y fluida con el entorno virtual [12].

### 3.3. Motor de videojuegos

Los motores de videojuegos son esencialmente conjuntos de rutinas de programación que proveen un marco de trabajo para el desarrollo de videojuegos y aplicaciones interactivas. Incluyen diversas funcionalidades y componentes integrados, tales como:

- 1 Motor de renderizado: Este es el componente que se encarga de dibujar y renderizar los gráficos 2D y 3D en la pantalla. Se encarga de todos los aspectos visuales del juego, incluyendo la iluminación, las sombras, los reflejos y otros efectos visuales [13].
- 2 Motor de física: Este motor simula las leyes de la física en el mundo del juego. Es el que se encarga de las colisiones, la gravedad, la fricción y otras interacciones físicas [13].

- 3 Animación: Este componente se encarga de dar vida a los personajes y objetos dentro del juego, permitiendo que se muevan y actúen de manera realista [13].
- 4 Inteligencia Artificial (IA): Este componente se encarga de controlar a los personajes no jugables (NPCs) en el juego, permitiendo que se comporten de manera creíble y respondan de manera inteligente a las acciones del jugador [13].
- 5 Escenarios gráficos: Los motores de videojuegos también incluyen herramientas para crear y diseñar los mundos y escenarios en los que se desarrolla el juego [13].
- 6 Codificación: Los motores de videojuegos a menudo incluyen su propio lenguaje de programación o permiten la integración con lenguajes de programación populares, permitiendo a los desarrolladores escribir el código que controla el juego [13].

En resumen, los motores de videojuegos son herramientas potentes que permiten a los desarrolladores centrarse en la creación de su juego, sin tener que preocuparse por programar todas estas funcionalidades desde cero [13].

Unity3D es un motor de desarrollo de videojuegos muy popular y potente que proporciona una gama de funcionalidades para la creación de experiencias interactivas y multimedia. Se lanzó por primera vez en 2005, y se ha convertido en una opción preferida para muchos desarrolladores debido a su capacidad de crear rápidamente prototipos y su facilidad para exportar proyectos a varias plataformas [14].

Unity3D es conocido por su editor gráfico intuitivo y su potente motor de física. El editor de Unity permite a los desarrolladores trabajar de manera más visual, creando y manipulando objetos directamente en el espacio 3D. Uno de los puntos fuertes de Unity es su capacidad para exportar juegos a una amplia gama de plataformas. Con Unity, es posible desarrollar un juego y luego publicarlo en Windows, Mac, iOS, Android,

consolas de videojuegos, dispositivos de realidad virtual (VR) y más. En cuanto a la programación, aunque Unity originalmente admitía C y C++, hoy en día está orientado hacia el uso de C#, que es un lenguaje de programación versátil y fácil de aprender [14].

#### 3.4. Arquitectura de la aplicación propuesta

La arquitectura del programa propuesto se fundamentó en el análisis de las tecnologías y técnicas adoptadas en los estudios y proyectos previos discutidos en la segunda sección de este artículo. Este diseño, que es una arquitectura en capas, se enfoca en la incorporación de una interfaz humano-máquina para la interpretación en tiempo real de los gestos de la mano del usuario. Se conceptualiza un sistema compuesto por tres módulos esenciales: el Traductor de Gestos Manuales, el Motor de videojuego y la Biblioteca SDK.

El módulo de Traductor de Gestos Manuales se encarga de la traducción inmediata de los movimientos manuales del usuario. Cada vez que el usuario mueve sus manos, este módulo traduce ese movimiento en información que puede ser leída por el Motor de Videojuego. Este último es responsable de traducir la información proporcionada por el Traductor de Gestos Manuales, reproduciendo los movimientos del usuario a la pantalla en tiempo real. También es responsable de la obtención de los modelos tridimensionales, que son proporcionados por la Biblioteca SDK. De igual manera, a través de la programación que se puede hacer por medio de las rutinas prediseñadas en el Motor de Videojuego, se puede modificar el comportamiento de los modelos, permitiendo su manipulación. La Biblioteca SDK, por su parte, dispone de bibliotecas para el seguimiento e interpretación de marcadores y permite el acceso a una nube que almacena los modelos tridimensionales a desplegar. Cuando la cámara enfoca un modelo tridimensional, el Motor de Videojuego consulta la Biblioteca SDK que recupera el

modelo necesario, accediendo a la nube y mostrándose al usuario.

Debido a su flexibilidad y capacidad de integración con diversas interfaces de control, se utilizó Unity 3D como el motor de juego. Su amplio soporte para diferentes plataformas y la disponibilidad de *plugins* que facilitan la interacción con varias interfaces humano-máquina lo hacen especialmente adecuado para este diseño. Además, Unity 3D permite la implementación de comportamientos a través de la programación en C#.

En cuanto a la interfaz de control, Leap Motion emergió como la opción sólida. Se favoreció su implementación debido a su compatibilidad con Unity 3D y su asequibilidad en comparación con otras opciones identificadas en el estado del arte.

Por último, se utilizó Vuforia Engine como SDK, debido a su extenso uso en proyectos relacionados y su compatibilidad con una variedad de motores de desarrollo, incluido Unity3D. Además, existe un *plugin* específico para la integración de Unity3D con Vuforia.

En la Figura 1, se visualiza la arquitectura del programa propuesto.

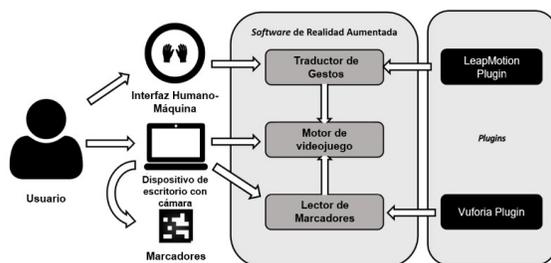


Figura 1. Arquitectura de la aplicación.

La Figura 2 presenta un diagrama ilustrativo de la disposición física del *hardware* requerido para una implementación adecuada de la arquitectura del programa. Esta configuración incluye una computadora (ya sea de escritorio o portátil), una cámara Web, marcadores para su lectura y la cámara Leap Motion Controller.

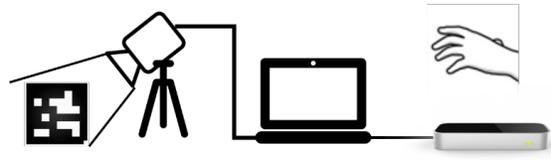


Figura 2. Disposición física del hardware.

#### 4. DESARROLLO DEL PROGRAMA

Para el desarrollo del programa se utilizó la metodología SCRUM, que debido a sus características, permitió una planificación y gestión efectiva del proyecto. Las reuniones constantes con el cliente facilitaron el desarrollo del sistema, al asegurar el cumplimiento de los requisitos del cliente y permitir el arreglo inmediato de bugs inesperados. La aplicación realizada tiene como objetivo ser una herramienta en apoyo al aprendizaje de las habilidades de la lectoescritura, el programa implementa técnicas de gamificación para facilitar dicho proceso.

Se realizaron 12 *sprints* para la realización del proyecto, abarcando once el desarrollo del programa como tal, siendo el doceavo la realización de pruebas y obtención de resultados. Los *sprints* se realizaron de la siguiente manera:

- *Sprint 1*: Análisis de los requerimientos. Para la obtención de las funcionalidades requeridas se utilizó un libro de pedagogía del idioma español llamado "Aprender a Leer: Programa Kantor para niños en Preescolar" escrito por Julio Varela. Dicho libro contiene conceptos teóricos sobre el aprendizaje de las letras en niños de preescolar, además de diferentes ejercicios que se utilizaron como base para la funcionalidad de los diferentes niveles.
- *Sprint 2*: Diseño de la arquitectura. Dicha arquitectura se ha explicado a detalle en la sección D. Arquitectura de la aplicación propuesta del apartado III de este artículo.
- *Sprint 3*: Diseño de los marcadores y modelos tridimensionales. Gracias al libro

escrito por Julio Varela, se desarrollaron marcadores basados en el sistema de tarjetas relámpago, un método de enseñanza de las letras bastante usado en las aulas, además se hizo una plantilla cuadriculada para indicar al usuario la disposición física de los marcadores. Por otro lado, se modelaron distintas figuras tridimensionales para su uso en el juego. Un ejemplo de marcador con su respectivo modelo tridimensional se observa en la Figura 3.



Figura 3. Marcador y modelo tridimensional de la letra O.

- *Sprint 4:* Diseño de los niveles y *wireframes* del sistema. Se realizaron *wireframes* que facilitaron el diseño de los menús, de igual manera se bosquejó el funcionamiento de cada uno de los niveles. En la Figura 4 se visualiza el *wireframe* del nivel 1.

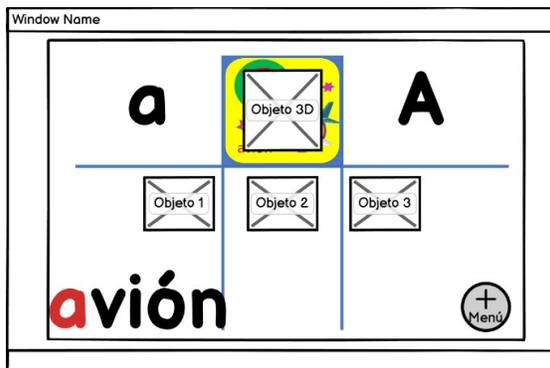


Figura 4. Wireframe nivel 1.

- *Sprint 5:* Grabación de los archivos multimedia. Para cada uno de los niveles se grabaron instrucciones específicas para orientar al usuario en el uso de la aplicación durante la ejecución de los niveles.
- *Sprint 6:* Desarrollo de los menús. Se diseñaron una serie de escenas con función

de menús: El menú principal, sección de créditos, video de ayuda y el menú de niveles. En la Figura 5 se observa el menú de niveles.



Figura 5. Menú de niveles.

- *Sprint 7:* Diseño del nivel 1. Este nivel contiene el funcionamiento base de los otros tres niveles, al iniciarse, se reproduce un audio de bienvenida que indica al usuario que debe colocar frente a la cámara el marcador que va a utilizar. Una vez se ha hecho esto aparece un modelo tridimensional encima del marcador cuyo nombre empieza con la letra representada en dicho marcador, además de una plantilla en pantalla que muestra el nombre del objeto resaltando la letra, también la muestra en mayúscula y minúscula. Esto se logra gracias al *plugin* de Vuforia Engine, que permite la lectura de los marcadores. Luego se reproduce otro audio que indica al usuario que debe tomar el objeto, sostenerlo y repetir lo que escuche, por ejemplo, en el caso de que se esté repasando la letra A, el objeto que aparece es un avión, entonces, se reproducirá el audio en el formato “a de avión”. Esta interacción con el objeto tridimensional se realiza a través de Leap Motion Controller, que a través de su cámara captura el movimiento de las manos y despliega en pantalla un esqueleto virtual de las articulaciones, esto gracias a su *plugin* de integración con Unity3D. Una vez que el usuario ha repetido con su boca la instrucción señalada, se suelta el objeto agarrado, a partir de aquí, el comportamiento del programa difiere

según el nivel. En este caso, siendo el nivel uno, en cuanto el usuario suelta el objeto, aparecen tres objetos en pantalla, cuyos nombres también comienzan con la letra que se está repasando, también aparece un marcador en cero en la esquina inferior derecha. Suena un audio que instruye al usuario a sujetar cada uno de estos objetos que, de igual forma al modelo principal, cuando son sostenidos, se reproduce el audio con sus nombres respectivos en el formato “a de avión”. El audio también recalca que dichos objetos deben arrastrarse hacia el modelo principal para ganar puntos. Una vez que el usuario ha ganado tres puntos, aparece el mensaje de felicitaciones por concretar el nivel y regresa al usuario al menú de niveles. Un ejemplo del funcionamiento de este nivel se observa en la Figura 6.



Figura 6. Funcionamiento del nivel uno.

- *Sprint 8:* Realización del video de ayuda. Dicho video se utilizará para explicar al usuario de forma básica los elementos necesarios para el correcto funcionamiento del programa: Los marcadores, Leap Motion Controller, la plantilla de marcadores, una cámara web.
- *Sprint 9:* Diseño del nivel 2. El funcionamiento de este nivel es derivado del nivel uno, con la excepción de que, en lugar de tres objetos aparecen cinco: los tres que aparecieron en el nivel anterior, además de dos objetos extra cuyo nombre no inicia con la letra que se está repasando. Cuando uno de estos objetos erróneos es sujetado por el usuario, de igual manera se reproduce el audio con el nombre del

objeto en el formato “a de avión”. Si se arrastra un objeto erróneo al modelo principal, todo el conjunto de objetos se reordena de forma aleatoria, incluyendo el que sujetaba el usuario. Un ejemplo del funcionamiento de este nivel se observa en la Figura 7.



Figura 7. Funcionamiento del nivel dos.

- *Sprint 10:* Diseño del nivel 3. El comportamiento es exactamente el mismo que el del nivel dos, con la diferencia de que al tocar el modelo que sale sobre el marcador, el audio que se reproduce es simplemente el nombre del objeto. También los otros modelos tridimensionales son reemplazados con palabras, también tridimensionales. Un ejemplo del funcionamiento de este nivel se observa en la Figura 8.



Figura 8. Funcionamiento del nivel tres.

- *Sprint 11:* Diseño del nivel 4. Su programación es diferente a la de los anteriores niveles en cierta medida. En cuanto se suelta el modelo principal, aparece una palabra la cual contiene la letra que se está repasando, no necesariamente al principio. El usuario tiene que tomar cada

una de las letras correctas dentro de dicha palabra y arrastrarlas al modelo principal. En caso de que se arrastre una letra errónea, ésta reaparece en el lugar que le corresponde dentro de la palabra. Una vez que en dicha palabra se han arrastrado al modelo principal las letras correctas, las letras restantes desaparecerán, se ganará un punto y aparecerá la siguiente palabra. Se necesitan cuatro puntos para completar el nivel. Cabe mencionar que las palabras siempre aparecerán en un orden aleatorio. Un ejemplo del funcionamiento de este nivel se observa en la Figura 9.



Figura 9. Funcionamiento del nivel cuatro.

- *Sprint 12*: Aplicación de un caso de estudio. Las pruebas y el análisis de resultados se describen con mayor detalle en el apartado siguiente.

## 5. RESULTADOS OBTENIDOS

En este apartado se describe el caso de estudio que se llevó a cabo para evaluar la aplicación propuesta. Se describe la institución donde se llevaron a cabo las pruebas, las características de las mismas y una breve explicación de los resultados obtenidos.

### 5.1. Caso de estudio

Debido a que el programa está enfocado al desarrollo de las habilidades tempranas de la lectoescritura (el reconocimiento de las letras y relacionarlas con su respectivo sonido, así como con objetos reconocibles para el usuario), se determinó que el grado educativo

ideal para realizar las pruebas al programa era la etapa preescolar. Para la implementación del proyecto desarrollado se acudió al Jardín de Niños Cavadonga, ubicado en la ciudad de Orizaba, Veracruz, México. Las pruebas se realizaron a cargo de la directora de la institución, quien permitió realizar las pruebas con ayuda de los niños de los tres grados escolares.

Debido a que la institución es una escuela privada y que se requería en todo momento la presencia de una maestra, la cantidad de niños que ayudaron a realizar la prueba fue mínima, también, por causa del avance académico presentado en cada uno de los grados escolares, se determinó que los niños de primero auxiliarían en las pruebas de los niveles uno y dos, los niños de segundo en los niveles dos y tres, mientras que los niños de tercero con los niveles tres y cuatro. En total diez niños participaron:

- Primer grado: dos niños, uno adelantado en el aprendizaje de las habilidades de lectoescritura, el otro rezagado.
- Segundo grado: cuatro niños, dos adelantados, dos rezagados.
- Tercer grado: cuatro niños, dos adelantados, dos rezagados.

Se realizaron cuatro rondas de pruebas, en las que se repasó con los estudiantes las letras que les costaban más trabajo aprender. Previo a cada prueba al niño se realizó una ronda de preguntas:

- 1 ¿Cómo se llama esto? (Se le mostró un objeto de los que se observan en los marcadores, se solicitó una vocalización específica del nombre).
- 2 ¿Con que letra empieza? (Se solicitó vocalización de la letra).
- 3 ¿Cómo se ve esa letra? (Se le mostraron diversos marcadores de las letras, incluida aquella que contiene el objeto que se le mostró).

Posterior a la realización de la prueba, se le repitieron las preguntas previas al niño, agregando otras preguntas de control:

- 4 ¿Te gustó el juego?
- 5 ¿Se te hizo difícil en algún momento?
- 6 Menciona qué te gusto/disgusto del juego.
- 7 ¿Volverías a jugarlo?

La finalidad de realizar preguntas de control al iniciar y finalizar la prueba es determinar en qué grado el programa está ayudando a largo plazo al estudiante a aprender las letras repasadas por medio de asociación: El objeto cuyo nombre inicia con cierto sonido, es relacionado por el estudiante con el símbolo que representa la letra, se buscaba determinar si esta relación perdura en la mente del niño a largo plazo, gracias a la herramienta desarrollada en este proyecto. En la Figura 10 se visualiza a uno de los estudiantes durante la aplicación de las pruebas.



Figura 10. Estudiante durante aplicación de prueba.

## 5.2. Interpretación de los resultados

Para la correcta interpretación de los resultados, se recurrió a la maestra que supervisaba las pruebas, pues debido al grado académico de los estudiantes y modelo educativo de la escuela, era imposible realizarles pruebas a los niños que representaran la obtención de datos duros. Mas bien, gracias a las preguntas de control y lo observado durante la realización de las pruebas se obtuvieron los siguientes resultados:

- Nueve de los diez estudiantes demostraron un interés pleno y sostenido en el tiempo por el uso del programa.
- La herramienta es especialmente útil en niños con problemas de déficit de atención diagnosticada.
- Se determinó que tres de los cinco niños que padecen rezago en el aprendizaje de las letras, deben esta incapacidad a un problema del lenguaje, más específicamente a problemas con la vocalización en general, aunque son capaces de identificar las letras si es otra persona la que pronuncia correctamente la letra o palabra.
- La RA fue especialmente llamativa para los niños, que se emocionaban bastante con los objetos que aparecían en pantalla.
- El uso de una interfaz humano-máquina como fue Leap Motion Controller demostró ser verdaderamente importante, ya que lo que más les gustó a absolutamente todos los niños fue agarrar los objetos y las palabras a través de la pantalla, el poder arrastrarlas y supuestamente tocarlas.
- El uso de puntos, sonidos de felicitación y voces sirvió de motivación en especial a los niños rezagados.
- El único estudiante que mostró renuencia y exasperación en la última ronda de prueba es un niño que además de problemas en el lenguaje sufre déficit de atención, lo cual hizo complicada la realización de la prueba sin apoyo de un docente indicándole qué hacer.
- La herramienta, de hecho, demostró tener la capacidad de hacer que niños con rezago académico en el aprendizaje de las letras pudieran discriminar con mayor facilidad palabras y letras las unas de las otras.
- El programa también auxilió en el aprendizaje de nuevas palabras.
- Los niños demostraron una pronunciada avidez por seguir jugando, nuevamente, debido a que les encantó tomar los objetos y las palabras con las manos para jugar con ellos.

## 6. CONCLUSIONES Y TRABAJO A FUTURO

En la actualidad, México exhibe un rendimiento académico insatisfactorio en lo que respecta a la comprensión lectora, un indicador fundamental del desarrollo intelectual de los estudiantes. Dado que una capacidad de comprensión lectora limitada conlleva una menor aptitud para formular y desarrollar ideas complejas, un fortalecimiento sistemático de las habilidades de lectoescritura tendría un impacto positivo en la formación académica integral de los estudiantes. En el análisis presentado en este estudio, se demuestra que diversas investigaciones y proyectos han generado soluciones satisfactorias a las problemáticas en el aprendizaje enfrentadas por los estudiantes en diversas áreas de estudio, tales como las matemáticas, la ética y la medicina. Estas investigaciones utilizaron de forma consciente y eficaz la RA, así como diferentes interfaces humano-máquina que permitieron la correcta resolución de las problemáticas presentadas en sus estudios correspondientes.

La herramienta descrita en este trabajo se presentó como una aplicación que ayuda en el desarrollo de las habilidades más básicas de la lectoescritura: la identificación de sonidos con el símbolo de su respectiva letra, así como la asociación de ambos con un objeto específico. La utilización de RA y Leap Motion Controller como interfaz de control, probó ser un éxito en lo que se refiere al proceso para captar la plena atención de los estudiantes de preescolar, puesto que la capacidad de los estudiantes por tocar los objetos, palabras y letras generó en ellos afección por seguir usando el programa y por extensión, seguir aprendiendo. También, la implementación de técnicas de gamificación demostró que, en conjunto, el sistema desarrollado es eficiente en apoyo al aprendizaje de las vocales y consonantes que se repasaron con los estudiantes; siendo de hecho, muy útil con niños diagnosticados con déficit de atención. Por ende, se considera

que esta herramienta, única en su tipo, con respecto a su enfoque en el aprendizaje del idioma español, y a la fecha de la última revisión del estado del arte, es eficiente, útil, e innovadora para la enseñanza de las habilidades más básicas de la lectoescritura.

Como trabajo a futuro, se tiene la intención de ampliar el marco de utilidad de cada uno de los niveles, expandiendo el uso a todas las letras del abecedario, además de proponer el desarrollo de niveles enfocados al desarrollo de la percepción espacial de los niños, como también al diseño de niveles enfocados en la enseñanza de las sílabas. Siempre considerando que las letras r, c, s, q, p, g, j y ñ, se enseñan de forma diferente a las demás que componen el abecedario. Es importante llevar a cabo una investigación exhaustiva sobre interfaces humano-máquina que sean capaces de detectar manos más pequeñas. Dado que se ha identificado que el programa actual presenta dificultades para detectar las manos de los niños más pequeños, se recomienda explorar opciones tecnológicas que permitan una detección precisa y adecuada en este grupo de usuarios. Esto garantizará que todos los niños, independientemente de su tamaño, puedan interactuar de manera efectiva con el programa. Se buscará ampliar el grupo de participantes en las pruebas del programa. Involucrar a más niños con diferentes niveles de habilidad y antecedentes académicos permitirá obtener una perspectiva más amplia y representativa de los posibles usuarios. Esto proporcionará datos más sólidos sobre la efectividad del programa en diferentes contextos y permitirá realizar análisis comparativos más rigurosos. Finalmente se evaluará la posibilidad de implementar el programa a largo plazo en entornos educativos. Para ello, se sugiere realizar un seguimiento de los niños participantes en el estudio para medir su progreso a lo largo del tiempo y determinar si el programa tiene un impacto sostenido en el aprendizaje de la lectoescritura. Esto permitirá evaluar la viabilidad de utilizar el programa como una herramienta

complementaria en la enseñanza regular y explorar la posibilidad de colaborar con escuelas o instituciones educativas para su implementación a mayor escala.

### Agradecimientos

Este trabajo de investigación fue financiado por el Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) y la Secretaría de Educación Pública (SEP) de México. También se agradece al Tecnológico Nacional de México (TecNM) por el apoyo brindado para la realización de esta investigación.

### REFERENCIAS

- [1] Salinas, D., De Moraes, C., Schwabe M. Programa para la Evaluación Internacional de Alumnos (PISA) PISA 2018 - Resultados - Nota País México. 2018, 1-12. Recuperado el 20 de enero de 2023 de: [https://www.oecd.org/pisa/publications/PISA2018\\_CN\\_MEX\\_Spanish.pdf](https://www.oecd.org/pisa/publications/PISA2018_CN_MEX_Spanish.pdf)
- [2] Le, H. Q., Kim J. An augmented reality application with hand gestures for learning 3D geometry, 2017 IEEE Int. Conf. on Big Data and Smart Computing (BigComp). 2017, 34-41, doi: [10.1109/BIGCOMP.2017.7881712](https://doi.org/10.1109/BIGCOMP.2017.7881712).
- [3] Schutera, S., Schnierle, M., Wu, M., Pertz, T., Seybold, J., Bauer, P., Teutscher, D., Raedle, M., Heß-Mohr, N., Röck, S., Krause, M. J. On the potential of augmented reality for mathematics teaching with the application cleARmaths, *Education Sciences*. 2021, 11(8), 368, doi: [10.3390/educsci11080368](https://doi.org/10.3390/educsci11080368).
- [4] Karambakhsh, A., Kamel, A., Sheng B., Li, P., Yang, P., Feng, D. Deep gesture interaction for augmented anatomy learning, *Int. Journal of Information Management*. 2019, 45, 328-336, doi: [10.1016/j.ijinfomgt.2018.03.004](https://doi.org/10.1016/j.ijinfomgt.2018.03.004).
- [5] Sun, Y., Armengol-Urpi, A., Reddy Kantareddy, S. N., Siegel, J., Sarma, S. MagicHand: A Deep Learning Approach towards Manipulating IoT Devices in Augmented Reality Environment, *IEEE Conf. on Virtual Reality and 3D User Interfaces*. 2019, 1-9.
- [6] Bifulco P., Narducci, F., Vertucci, R., Ambrosi, P., Cesarelli, M., Romano, M. Telemedicine supported by Augmented Reality: An interactive guide for untrained people in performing an ECG test, *BioMedical Engineering OnLine*. 2014, 13(1), 1-16, doi: [10.1186/1475-925X-13-153](https://doi.org/10.1186/1475-925X-13-153).
- [7] Nugraha, I., Suminar, A. R., Octaviana, D. W., Hidayat, M. T., Ismail, A. The application of augmented reality in learning English phonetics, *Journal of Physics: Conference Series*. 2019, 1402(7), 077024, doi: [10.1088/1742-6596/1402/7/077024](https://doi.org/10.1088/1742-6596/1402/7/077024).

- [8] Lin, V., Liu, G. Z., Chen, N. S. The effects of an augmented-reality ubiquitous writing application: a comparative pilot project for enhancing EFL writing instruction, *Computer Assisted Language Learning*. 2020, 47, 989-1030, doi: [10.1080/09588221.2020.1770291](https://doi.org/10.1080/09588221.2020.1770291).
- [9] SyafiqahSafiee, N., Wanis Ismail, A., AR Home Deco: Virtual Object Manipulation Technique Using Hand Gesture in Augmented Reality, *UTM Comp. Proc. On Innovations in Computing Technology and Applications*. 2018, 979, 1-6, 2018.
- [10] Microsoft, Visual Studio Code. Recuperado el 27 de febrero de 2022 de <https://visualstudio.microsoft.com/es>.
- [11] Shamsee, H., Klebenov, N., Fayed, D., CCNA Data Center DCICT 640-916: Official Cert Guide. Cisco. 2015.
- [12] UltraLeap, Leap Motion Controller. Recuperado el 27 de febrero de 2023 de: <https://www.ultraleap.com/product/leap-motion-controller/#overview>
- [13] Atherton J., Wang G. Chunity: Integrated Audiovisual Programming in Unity. *Proc. of the Int. Conf. on New Interfaces for Musical Expression*. 2018, 102-107, doi: [10.5281/zenodo.1302695](https://doi.org/10.5281/zenodo.1302695).
- [14] Haas, J.K. A History of the Unity Game Engine An Interactive. Worcester: Worcester Polytechnic Institute. 2014. Recuperado el 1 de marzo de 2023 de: <https://digitalcommons.wpi.edu/iqp-all/3207>.

### ACERCA DE LOS AUTORES



#### Raúl de Jesús Sánchez Martínez.

Ingeniero en Sistemas Computacionales egresado del Instituto Tecnológico de Orizaba en el año 2021, actualmente en proceso de titulación para la obtención de su grado de Maestro en Sistemas Computacionales en la misma institución. Ha publicado anteriormente un artículo con las bases teóricas para el desarrollo del programa expuesto en el documento presentado en este Congreso. Dicha investigación se presentó en la 9ª Jornada de Ciencia y Tecnología Aplicada, con el nombre: Arquitectura de Software para utilizar realidad aumentada con control gestual de manos en apoyo al aprendizaje de la lectoescritura.



**María Antonieta Abud Figueroa**, nació en la ciudad de Orizaba, Ver. Es ingeniero en electrónica por la UAM-Iztapalapa, México DF en el año 1984, y maestra en ciencias en sistemas de información

por el ITESM-Morelos, en la ciudad de Cuernavaca, Mor. en el año 1991. Desde el año 1995 es profesora-investigadora en el área de posgrado del Instituto Tecnológico de Orizaba, en la ciudad de Orizaba, Ver. México. Su línea de investigación es la Ingeniería de Software. La M.C. Abud es miembro de IEEE.



**Ulises Juárez Martínez.**

Realizó su investigación doctoral sobre la programación orientada a aspectos con capacidades de corte en variables locales, obteniendo una plataforma para la verificación de aseveraciones a tiempo de ejecución. Actualmente es profesor investigador de la División de Estudios de Posgrado e Investigación del Tecnológico Nacional de México Instituto Tecnológico de Orizaba. Ha dirigido proyectos de investigación, tesis de nivel maestría y doctorado aplicando la orientación a aspectos en ingeniería de software, líneas de productos de software y en la implementación de lenguajes naturalísticos de propósito general. Recientemente ha iniciado investigaciones sobre la parte naturalística aplicada en el contexto del Metaverso.



**Lisbeth Rodríguez Mazahua.**

Licenciada en Informática (2000-2004) y Maestra en Ciencias Computacionales egresada del Instituto Tecnológico de Orizaba (2005-2007). Estudió el Doctorado en Ciencias en Computación en el Centro de Investigación y

de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN) de 2008 a 2012. Fue profesora de tiempo completo en la Universidad Autónoma del Estado de México de febrero de 2012 a enero de 2014. Realizó una estancia posdoctoral en el Instituto Tecnológico de Orizaba de febrero de 2014 a enero de 2016. Desde marzo de 2016 trabaja como profesora-investigadora en el Instituto Tecnológico de Orizaba. Es miembro del Sistema Nacional de Investigadores desde el 2016. Sus áreas de interés son Bases de Datos, Minería de Datos y Big Data.



**Hilarión Muñoz Contreras.**

Profesor investigador de la División de Estudios de Posgrado e Investigación en el Instituto Tecnológico de Orizaba. Dr. en Ingeniería industrial, colaborador en las áreas de Sistemas Computacionales e Ingeniería Industrial. Sus áreas investigación son: inteligencia artificial, calidad, recursos humanos y ética. Ha colaborado con diferentes instituciones de educación superior, y ha asesorado a empresas para la solución de problemas en mejora continua.