

**PROTOTIPO DE APLICACIÓN DE SOFTWARE BASADA EN  
VIDEOJUEGOS  
PARA LA MECANIZACIÓN Y REHABILITACIÓN DEL HABLA EN  
NIÑOS**

Daniel Andres Vasquez Murillo

Sebastian Castro Obando

Trabajo de Grado

**Director**

Andrés Adolfo Navarro Newball

Ingeniero

Pontificia Universidad Javeriana Cali

Facultad de Ingeniería y Ciencias

Ingeniería de Sistemas y Computación

Cali

2025

Santiago de Cali, 6 de Junio de 2025

Dr. Gerardo Mauricio Sarria, Director de Carrera  
Ingeniería de Sistemas y Computación Pontificia Universidad Javeriana.  
Cali.

Cordial saludo,

Nos permitimos presentar a su consideración el trabajo de grado denominado **“PROTOTIPO DE APLICACIÓN DE SOFTWARE BASADA EN VIDEOJUEGOS PARA LA MECANIZACIÓN Y REHABILITACIÓN DEL HABLA EN NIÑOS”**, con el fin de cumplir con los requisitos exigidos por la Universidad para llevar a cabo el proyecto de grado y posteriormente optar al título de Ingeniero de Sistemas y Computación.

Al firmar aquí, damos fe de que entendemos y conocemos las directrices para la presentación de Trabajos de Grado de la Facultad de Ingeniería y Ciencias aprobadas el 26 de noviembre de 2009, donde se establecen los plazos y normas para el desarrollo del trabajo de grado.

Atentamente,



---

Daniel Andres Vasquez Murillo  
Código: 8963154



---

Sebastian Castro Obando  
Código: 8965484

## TABLA DE CONTENIDOS

<b>Abstract</b>	<b>6</b>
<b>Resumen</b>	<b>7</b>
<b>1. Introducción</b>	<b>8</b>
1.1. Planteamiento del problema . . . . .	8
1.1.1. Formulación . . . . .	10
1.1.2. Sistematización . . . . .	10
1.2. Objetivos . . . . .	11
1.2.1. Objetivo general . . . . .	11
1.2.2. Objetivos específicos . . . . .	11
1.3. Justificación . . . . .	12
1.4. Alcance y Limitaciones . . . . .	13
<b>2. Marco Teórico y Trabajos Relacionados</b>	<b>15</b>
2.1. Marco de referencia . . . . .	15
2.1.1. Marco Teórico . . . . .	15
2.2. Trabajos relacionados . . . . .	20
<b>3. Metodología</b>	<b>24</b>
3.1. Enfoque de Prototipado Evolutivo . . . . .	24
3.1.1. ¿En qué consiste? . . . . .	24
3.1.2. Principios fundamentales . . . . .	24
3.1.3. Fases típicas . . . . .	24
3.1.4. Ventajas y desventajas . . . . .	25
3.1.5. Adecuación al presente proyecto . . . . .	25

3.2. Actividades planificadas . . . . .	27
3.2.1. Análisis de la terapia de mecanización de palabras . . . . .	27
3.2.2. Investigación sobre RV y videojuegos en terapia . . . . .	27
3.2.3. Implementación del prototipo . . . . .	28
3.2.4. Validación de la efectividad . . . . .	28
3.2.5. Fases del desarrollo . . . . .	29
3.2.6. Interacciones con Usuarios Durante el Desarrollo y Validación . . . . .	30
3.2.7. Alcance de la Documentación . . . . .	31
<b>4. Requerimientos del Sistema</b>	<b>32</b>
4.1. Verificación Interna de Requerimientos . . . . .	34
<b>5. Análisis y Diseño</b>	<b>36</b>
5.1. Modelo terapéutico de Daniel Ling y adaptación al juego . . . . .	36
5.2. Concepto inicial del juego: torre mágica y múltiples minijuegos . . . . .	38
5.3. Decisiones de diseño finales: enfoque en el minijuego de las pociones . . . . .	39
5.4. Estructura de niveles: planificación vs. producto final . . . . .	42
5.5. Interfaz de usuario, interacción RV y retroalimentación . . . . .	45
<b>6. Implementación</b>	<b>50</b>
6.1. Herramientas y tecnologías . . . . .	50
6.2. Desarrollo en Unity y Arquitectura General del Prototipo . . . . .	52
6.3. Componentes Clave Implementados: . . . . .	54
<b>7. Plan de Validación y Resultados</b>	<b>59</b>
7.1. Plan de Validación . . . . .	59
7.2. Resultados de las pruebas . . . . .	60
7.3. Evaluación de Rendimiento . . . . .	62
<b>8. Discusión</b>	<b>65</b>
8.1. Evaluación crítica del diseño y metodología del prototipo . . . . .	65

8.2. Percepción de los usuarios: respuesta de niños y terapeutas al enfoque gamificado	67
<b>9. Conclusiones</b>	<b>71</b>
<b>10. Trabajo Futuro</b>	<b>74</b>
<b>11. Anexos</b>	<b>80</b>
<b>Anexo 2: Video de referencia para configuración de Unity RV</b>	<b>82</b>

## LISTA DE FIGURAS

1.1. Imagen del panel de estadísticas dentro del juego . . . . .	13
5.1. Estados de aprendizaje de consonantes . . . . .	36
5.2. Orden natural de adquisición de sonidos . . . . .	37
5.3. Esquema de niveles del prototipo . . . . .	38
5.4. Entorno virtual del juego con temática de torre de mago . . . . .	40
5.5. Elementos principales para la interacción (items) . . . . .	42
5.6. Diagrama de la arquitectura del sistema . . . . .	44
5.7. Diagrama de flujo del minijuego . . . . .	45
5.8. Interfaz gráfica prototipo . . . . .	46
5.9. Paneles de la interfaz gráfica de usuario con menús de inicio y registro . . . . .	47
5.10. Vista del entorno del caldero con un ingrediente resaltado listo para ser agarrado y la sílaba propuesta . . . . .	49
6.1. Pila tecnológica (hardware → contenido). . . . .	50
6.2. Gafas de realidad virtual Meta Quest 3 utilizadas . . . . .	52
6.3. Modelo 3D de la torre de mago utilizada en el entorno de juego . . . . .	57
7.1. Imagen de rendimiento del videojuego . . . . .	63
7.2. Imagen de la correcta interacción con los objetos . . . . .	63
7.3. Imagen de la correcta eliminación de los objetos . . . . .	64
7.4. Imagen del correcto funcionamiento de la progresión en los niveles . . . . .	64

## ABSTRACT

This project presents "Fonetic Tower", a Virtual Reality (VR) and video game-based software prototype for speech rehabilitation (word mechanization) in children aged 6-9. Its aim is to improve children's motivation and participation through an immersive, playful environment, overcoming limitations of traditional methods and prior tools. Development involved a literature review, interviews with therapists from the Instituto para Niños Ciegos y Sordos del Valle del Cauca (INCSVC), and evolutionary prototyping in Unity for Meta Quest 3. "Fonetic Tower", is a fantasy-themed game (wizard's tower) featuring a potion preparation minigame with seven levels of pronunciation exercises based on Daniel Ling's model. It includes gamification and visual feedback, with therapist-supervised pronunciation validation at this stage. Pilot tests with four children (7-11 years) and therapists at INCSVC showed high child engagement, motivation, and immersion, and positive therapist feedback on its potential and data logging. Although technical goals were met, limitations like manual voice validation and restricted content pave the way for future enhancements, including automatic speech recognition and activity expansion.

**Keywords:** Virtual Reality, Therapeutic Video Games, Speech Rehabilitation, Word Mechanization, Child Therapy.

## RESUMEN

Este trabajo presenta "Fonetic Tower", un prototipo de software en Realidad Virtual (RV) y videojuegos para la rehabilitación del habla (mecanización de palabras) en niños de 6 a 9 años. Su objetivo es mejorar la motivación y participación infantil mediante un entorno inmersivo y lúdico, superando limitaciones de métodos tradicionales y herramientas previas. Se desarrolló mediante revisión bibliográfica, entrevistas con terapeutas del Instituto para Niños Ciegos y Sordos del Valle del Cauca (INCSVC) y prototipado evolutivo en Unity para Meta Quest 3. "Fonetic Tower", es un juego de fantasía (torre de mago) con un minijuego de preparación de pociones y siete niveles de ejercicios de pronunciación basados en el modelo de Daniel Ling. Incluye gamificación y retroalimentación visual, con validación de pronunciación supervisada por terapeutas en esta etapa. Pruebas piloto con cuatro niños (7-11 años) y terapeutas en INCSVC mostraron alto compromiso, motivación e inmersión infantil, y una percepción positiva de los terapeutas, quienes valoraron su potencial y el registro de datos. Aunque los objetivos técnicos se lograron, limitaciones como la validación manual de voz y el contenido restringido sentaron las bases para futuras mejoras, como el reconocimiento de voz automático y la ampliación de actividades.

**Palabras Clave:** Realidad Virtual, Videojuegos Terapéuticos, Rehabilitación del Habla, Mecanización de Palabras, Terapia Infantil.

## 1. INTRODUCCIÓN

La rehabilitación del habla en niños a menudo enfrenta el desafío de mantener el interés y la participación de los pacientes más jóvenes, ya que las técnicas tradicionales pueden resultar monótonas. Frente a esto, enfoques lúdicos como los videojuegos y la realidad virtual (RV) emergen como herramientas prometedoras para transformar la terapia. La gamificación de los ejercicios terapéuticos tiene el potencial de convertir la práctica repetitiva en actividades dinámicas y atractivas, lo que podría incrementar la adherencia al tratamiento.

Este proyecto de grado se enfocó en el desarrollo de "Fonetic Tower", un prototipo de videojuego terapéutico en RV. Esta herramienta busca proporcionar un entorno inmersivo e interactivo diseñado para la mecanización de palabras en niños de 6 a 9 años, complementando la labor del terapeuta. El prototipo se materializa en un escenario de fantasía, una torre de mago, donde los niños realizan ejercicios de pronunciación a través de un minijuego de preparación de pociones. Este minijuego se estructura en siete niveles de dificultad progresiva, basados en el modelo de Daniel Ling, para asegurar una progresión gradual.

Un aspecto clave en esta fase inicial es que la validación de la pronunciación es supervisada directamente por un terapeuta, quien confirma los intentos del niño dentro del juego. Desarrollado en Unity para la plataforma Meta Quest 3, "Fonetic Tower", utiliza controles de movimiento para una interacción natural con objetos virtuales. Esta iniciativa buscó innovar en la terapia del habla infantil mediante la integración de tecnología de RV y principios de diseño centrados en el usuario, con el objetivo de hacer las sesiones terapéuticas más efectivas y amenas, superando limitaciones identificadas en herramientas previas.

### 1.1. Planteamiento del problema

El desarrollo del habla en los niños es crucial durante su crecimiento y en cómo interactúan con el mundo. No obstante, los niños con dificultades en el habla enfrentan barreras que

pueden afectar su capacidad de comunicarse de manera efectiva. Una correcta capacidad comunicativa es fundamental para el desarrollo social, cognitivo y emocional. En este contexto, la tecnología actual ofrece herramientas valiosas para apoyar este proceso. Por ejemplo, Maite Frutos et al. [1] presentan un juego educativo computarizado diseñado específicamente para niños y adolescentes con espectro autista. Este sistema, que combina aplicaciones de gestión y aprendizaje, tiene como objetivo reforzar el lenguaje mediante actividades interactivas que fomentan la pronunciación correcta y mejoran la motivación y la concentración. Así, herramientas como esta contribuyen a que los usuarios construyan relaciones, comprendan mejor su entorno y participen activamente en la sociedad.

De manera complementaria, las terapias de rehabilitación del habla recurren a un conjunto de técnicas diseñadas para mejorar la capacidad comunicativa de los niños. Entre estas se encuentra la mecanización de palabras, un proceso que implica la repetición estructurada y controlada de palabras y frases con el objetivo de mejorar la pronunciación, fluidez y articulación del habla. La integración de estas técnicas con enfoques tecnológicos innovadores, como los videojuegos y la realidad virtual, puede potenciar los beneficios terapéuticos. Además, la intervención temprana, apoyada por profesionales especializados, ha demostrado ser efectiva para corregir o mitigar los problemas de habla y lenguaje, contribuyendo significativamente a mejorar la calidad de vida del niño [2].

En este contexto, los videojuegos se han consolidado como una herramienta innovadora y eficaz para complementar las terapias de rehabilitación. Gracias a su capacidad para crear entornos controlados y adaptativos, permiten a los terapeutas diseñar ejercicios personalizados que responden a las necesidades específicas de cada paciente. Esto resulta particularmente valioso en procesos de rehabilitación prolongados, donde la motivación y el compromiso del paciente son cruciales. Elementos como recompensas, retroalimentación inmediata y niveles de dificultad ajustables transforman los ejercicios tradicionales en actividades dinámicas y atractivas. Para los niños, estos estímulos lúdicos no solo hacen más amena la terapia, sino que también potencian su progreso en la mejora de habilidades comunicativas y cognitivas, contribuyendo a una recuperación más efectiva y sostenida [3].

El Instituto para niños ciegos y sordos del Valle del Cauca (INCSVC) ha implementado

un prototipo de aplicación de software basada en videojuegos para apoyar la rehabilitación. Estos prototipos han mantenido el interés de los niños durante las sesiones terapéuticas. Sin embargo, debido a una falta de niveles de dificultad adecuados, un enfoque visual pobre y mala usabilidad, los prototipos resultan poco atractivos. Las técnicas clásicas de rehabilitación del habla, basadas en ejercicios de repetición y formularios, pueden ser monótonas y aburridas para los niños. En particular, el proyecto TalkingToTeo [4], diseñado para apoyar el desarrollo del lenguaje, presenta varias falencias. La interfaz del software no es intuitiva ni entretenida, dificultando su uso independiente y reduciendo su eficacia. Además, el contenido no es variado y el sistema del terapeuta está separado del videojuego.

### **1.1.1. Formulación**

Se propuso un prototipo de aplicación de software basado en videojuegos que, mediante el uso de realidad virtual (RV), ofrezca un entorno inmersivo para facilitar la mecanización de palabras y otros aspectos clave de la terapia del habla en niños.

¿Cómo implementar un prototipo de aplicación de software accesible e interactivo basado en realidad virtual y videojuegos para la mecanización de palabras en la rehabilitación del habla en niños de 6 a 9 años?

### **1.1.2. Sistematización**

- ¿Cómo funciona la mecanización de palabras en niños?
- ¿Cómo se puede favorecer el desarrollo de la mecanización de palabras en niños a través de una aplicación de software basada en videojuegos y realidad virtual?
- ¿Cómo se puede implementar el prototipo de una aplicación de software basada en videojuegos y realidad virtual para favorecer la mecanización de palabras?
- ¿Cómo validar la aplicación implementada?

## **1.2. Objetivos**

### **1.2.1. Objetivo general**

Desarrollar un prototipo de aplicación de software accesible e interactiva basado en videojuegos y realidad virtual para la mecanización de palabras, orientado a la rehabilitación del habla en niños de 6 a 9 años, que apoye a los profesionales de la salud en sus terapias.

### **1.2.2. Objetivos específicos**

- Analizar la terapia de mecanización de palabras para comprender los mecanismos y etapas de aprendizaje del habla, identificando su eficacia para la rehabilitación a través de actividades interactivas.
- Investigar cómo la realidad virtual (RV) y los videojuegos pueden utilizarse para diseñar un prototipo de aplicación de software que pueda apoyar la mecanización de palabras en niños, evaluando las ventajas y limitaciones de estas tecnologías en el contexto terapéutico.
- Desarrollar e implementar un prototipo de aplicación de software basado en videojuegos y realidad virtual para facilitar la mecanización de palabras en niños mediante actividades, ejercicios y dinámicas interactivas. Que fomenten el interés, la participación activa y la retroalimentación constante en un entorno seguro.
- Evaluar la efectividad del prototipo a través de retroalimentación de profesionales al igual que buscando la opinión de los niños, para de este modo ir recopilando datos cuantitativos y cualitativos para ajustar y mejorar la funcionalidad del sistema, su precisión, usabilidad y su impacto en la rehabilitación del habla.

### 1.3. Justificación

El desarrollo de un prototipo de aplicación de software basado en realidad virtual y videojuegos para la rehabilitación del habla en niños es altamente útil, ya que aborda una necesidad crítica en la intervención terapéutica infantil. La realidad virtual (RV) ofrece un entorno seguro y controlado que facilita la mecanización de palabras a través de la práctica repetitiva en un ambiente lúdico, lo que aumenta la adherencia al tratamiento. Estudios han demostrado que la RV es eficaz en la mejora de habilidades cognitivas y lingüísticas en niños con dificultades de aprendizaje, como la dislexia y otras condiciones neuropsicológicas [5]. Esto sugiere que un enfoque similar podría ser beneficioso para la rehabilitación del habla, proporcionando un entorno atractivo y efectivo para los niños.

El impacto de la implementación de este sistema podría ser significativo, no sólo en términos de resultados terapéuticos, sino también en la forma en que los tratamientos de rehabilitación del habla son percibidos y adoptados. Los niños, a menudo reacios a participar en sesiones tradicionales, pueden estar más dispuestos a participar en terapias basadas en videojuegos y RV debido al componente lúdico. Además, este sistema podría minimizar el impacto negativo asociado con la falta de interés o compromiso, proporcionando una alternativa atractiva y motivante. Investigaciones previas han mostrado que la gamificación y el uso de RV pueden mejorar significativamente la motivación y la participación en actividades terapéuticas. Este resultado positivo nos lo da Stephanie Crowe, et al. en donde se diseñó un modelo de colaboración entre pacientes con lesiones cerebrales y terapeutas basado en realidad virtual, que buscaba modificar la consulta tradicional y hacerla más lúdica. El modelo no solo logró resultados positivos, sino que además, se logró demostrar que la informalidad e interacción entre paciente y terapeuta se volvía más amigable y manejable a través de la realidad virtual [6].

En cuanto a la viabilidad, el desarrollo de este sistema requiere una combinación de recursos tecnológicos y humanos. Es necesario contar con software de realidad virtual, hardware compatible, y el acceso a datos relevantes sobre patrones de habla y necesidades terapéuticas. Sin embargo, el costo y la disponibilidad de estos recursos están disminuyendo con el tiempo,

lo que hace que este tipo de intervención sea menos accesible. Además, el respaldo académico y clínico en conjunto al instituto de niños ciegos y sordos del Valle del Cauca (INCSVC), como el mostrado en estudios recientes, subraya la viabilidad técnica y científica del uso de RV en rehabilitación. Esto indica que, con la planificación y los recursos adecuados, este proyecto puede ser implementado exitosamente y con un impacto positivo duradero en los tratamientos de rehabilitación del habla para niños.

## 1.4. Alcance y Limitaciones

El proyecto se enfocó en la creación de un prototipo de aplicación de software que facilitó la mecanización de palabras en un entorno de realidad virtual y videojuegos. Este sistema les permitió a los niños reforzar habilidades del habla en un ambiente controlado y lúdico, fomentando la práctica activa y motivadora mediante dinámicas interactivas. El prototipo integró reportes con estadísticas de uso individual, lo que posibilitó un análisis más eficiente y dinámico por parte de los profesionales de la salud (ver Figura 7.4).

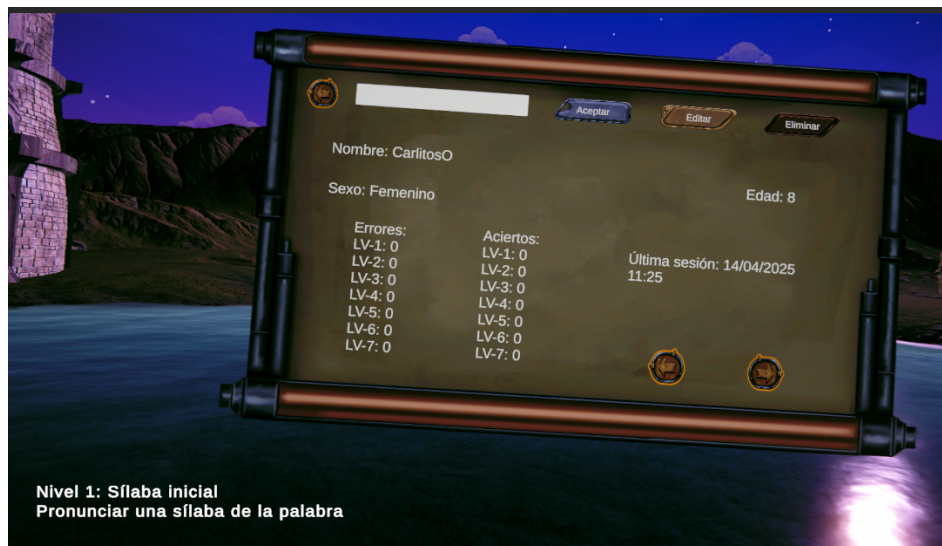


Figura 1.1: Imagen del panel de estadísticas dentro del juego

Para su validación, el prototipo se probó en un entorno de laboratorio con profesionales del Instituto para Niños Ciegos y Sordos del Valle del Cauca (INCSVC) y con niños, contando

con la autorización previa de un comité de ética en salud para la participación de los menores. Estas pruebas buscaron evaluar su efectividad y usabilidad en un contexto real. El sistema fue diseñado específicamente para la plataforma Meta Quest 3, con un enfoque inicial en el sistema operativo Windows y la posibilidad de compatibilidad con Android. El desarrollo se centró en implementar buenas prácticas de ingeniería de software a fin de asegurar un diseño robusto, eficiente y seguro.

En esta etapa de desarrollo, el reconocimiento del habla no se automatizó mediante tecnología; la validación de la correcta pronunciación de los ejercicios fue supervisada directamente por un profesional a cargo, quien validó manualmente las producciones de los niños y evitó así la dependencia de sistemas de reconocimiento de voz aún en desarrollo.

El proyecto presentó algunas limitaciones significativas: el sistema requirió la supervisión constante de profesionales de la salud (fonoaudiólogos) durante las sesiones para validar las actividades realizadas, ya que estos contaban con el criterio adecuado para verificar el progreso de los infantes. Por último, las capacidades del hardware Meta Quest 3 determinaron la calidad y el alcance de la experiencia inmersiva proporcionada por el prototipo.

## **2. MARCO TEÓRICO Y TRABAJOS RELACIONADOS**

### **2.1. Marco de referencia**

En esta sección se abordarán y explicarán los conceptos fundamentales necesarios para comprender y desarrollar el prototipo de aplicación de software para la rehabilitación del habla. Se explorarán los procesos del desarrollo y rehabilitación del habla, con un enfoque particular en la población infantil, así como la mecanización de palabras como estrategia central en la terapia. Además, se examinarán los sistemas de reconocimiento de voz y su aplicación en la rehabilitación, junto con la integración de videojuegos en procesos terapéuticos. Un aspecto clave será el análisis del papel de la Realidad Virtual (RV) en la creación de entornos inmersivos para la terapia, permitiendo así comprender el importante rol que juega la RV en el diseño del entorno y las mecánicas del prototipo propuesto. El objetivo es proporcionar una comprensión integral de las tecnologías y metodologías involucradas, sentando las bases para el diseño y desarrollo informado del prototipo de aplicación que busca crear un entorno terapéutico más atractivo y efectivo.

#### **2.1.1. Marco Teórico**

##### **Desarrollo del habla en niños**

El desarrollo del habla en niños es un proceso complejo que involucra la adquisición de habilidades lingüísticas y la capacidad de producir sonidos del habla de manera precisa. Este proceso se divide en etapas que van desde el balbuceo en los primeros meses de vida hasta el dominio de estructuras gramaticales complejas en la edad escolar. Los bebés comienzan a producir sonidos de vocales alrededor de los 3 a 8 meses de edad, seguidos por la combinación de consonantes y vocales en un patrón conocido como balbuceo entre los 5 y 10 meses, lo que constituye un hito temprano en el desarrollo del habla [7]. Factores como la genética, el

entorno y la estimulación juegan un papel crucial en este desarrollo. Los trastornos del habla, como la dislalia o el tartamudeo, pueden surgir durante este proceso y requieren intervención temprana para prevenir dificultades a largo plazo en la comunicación y el desarrollo social del niño.

### **Rehabilitación del habla**

Según la clínica profesional de Cleveland [8], La rehabilitación del habla comprende un conjunto de técnicas y estrategias diseñadas para mejorar la comunicación oral en individuos con trastornos del habla. La mecanización de palabras, un componente clave de esta rehabilitación, implica la práctica repetitiva y estructurada de sonidos y palabras para mejorar la articulación y fluidez. Las técnicas tradicionales incluyen ejercicios de repetición, juegos de palabras y terapia de espejo. La intervención temprana es crucial, ya que aprovecha la plasticidad neuronal de los niños pequeños, permitiendo una mejor adaptación y aprendizaje. La eficacia de la rehabilitación depende de la consistencia de la práctica y la motivación del paciente, aspectos que las nuevas tecnologías buscan mejorar.

### **Mecanización de palabras**

Este método implica la repetición sistemática y estructurada de palabras o frases específicas con el objetivo de automatizar la correcta articulación y pronunciación. El proceso se basa en la teoría del aprendizaje motor, que sostiene que la práctica repetitiva puede conducir a la automatización de habilidades motoras, incluyendo los movimientos necesarios para el habla. En la terapia, se comienza con sonidos aislados, progresando hacia sílabas, palabras y finalmente frases completas. Mejora la precisión articulatoria, la fluidez y la velocidad del habla. Esta progresión gradual desde elementos fonéticos simples hacia estructuras más complejas es un principio fundamental que guía la estructura de niveles implementada en el prototipo "Fonetic Tower", donde los ejercicios iniciales se centran en fonemas y sílabas básicas, avanzando progresivamente hacia palabras y combinaciones que demandan mayor habilidad articulatoria, tal como se detalla en las secciones 5.1 y 5.4 de este documento.

## **Realidad Virtual (RV) en terapia**

La Realidad Virtual (RV) es una tecnología avanzada que genera entornos tridimensionales simulados por computadora, ofreciendo experiencias inmersivas y controladas para la práctica terapéutica. Esta tecnología permite a los usuarios explorar e interactuar con mundos virtuales a través de diversos periféricos, como gafas de RV, mandos y auriculares, creando una experiencia multisensorial. En el contexto específico de la terapia del habla, la RV presenta ventajas significativas: **Inmersión**, proporcionando un entorno controlado y envolvente que minimiza las distracciones externas, facilitando la concentración del paciente; **Interactividad**, que permite la manipulación directa de objetos virtuales, lo cual puede aumentar el compromiso y la motivación del paciente durante las sesiones; **Personalización**, donde los escenarios pueden adaptarse meticulosamente a las necesidades específicas de cada paciente, creando entornos seguros y estimulantes donde los niños pueden practicar habilidades de comunicación de manera individualizada. Además, como lo señalan Bell et al. [9], la RV permite la recopilación de datos precisa y en tiempo real, facilitando un seguimiento detallado de la evolución del paciente, una ventaja crucial en terapias continuas. No obstante, es importante considerar los posibles desafíos al utilizar RV con niños pequeños (6-9 años). Entre estos se encuentran el riesgo de ciber-mareo (cinetosis) si la tasa de refresco no es óptima o si las experiencias son demasiado intensas, y la necesidad de gestionar adecuadamente los tiempos de exposición para evitar fatiga visual o sobreestimulación. Estas consideraciones, abordadas en la metodología del presente proyecto, son cruciales para garantizar una experiencia terapéutica segura y positiva.

## **Videojuegos en rehabilitación**

Aprovechando elementos del diseño de juegos, como los objetivos, desafíos, progresión, recompensas, estética y narrativa, se mejora la motivación y el compromiso en los procesos terapéuticos. Los videojuegos activos (AVGs), según Staiano y Flynn [10], han demostrado ser herramientas prometedoras en la rehabilitación, ya que pueden ser aceptables y disfrutables para diversas poblaciones en terapia. En el contexto de la rehabilitación del habla, los

videojuegos pueden transformar ejercicios repetitivos en actividades divertidas y desafiantes. El diseño de juegos terapéuticos implica la creación de mecánicas de juego que refuercen los objetivos de la terapia, como la pronunciación correcta o la fluidez verbal. Estudios han demostrado que el uso de videojuegos en rehabilitación puede aumentar significativamente la adherencia al tratamiento y mejorar los resultados, especialmente en poblaciones pediátricas, al proporcionar un entorno lúdico que fomenta la práctica continua.

### **Tecnología de reconocimiento de voz**

Según la enciclopedia de los sistemas de información [11] esta es una tecnología que permite a los usuarios ingresar información mediante comandos hablados o dictados. Por ello la tecnología de reconocimiento de voz es un componente crucial en la evaluación y mejora de la pronunciación en la terapia del habla. Esta tecnología se fundamenta en algoritmos avanzados de procesamiento de señales y aprendizaje automático para convertir el habla humana en texto o comandos interpretables por máquinas, permitiendo una evaluación objetiva y en tiempo real de la pronunciación y la articulación. Los sistemas modernos de reconocimiento de voz emplean una serie de técnicas sofisticadas: Procesamiento de señales de audio, que incluye métodos como la Transformada Rápida de Fourier (FFT) para el análisis espectral y los Coeficientes Cepstrales en las Frecuencias de Mel (MFCC) para la extracción de características acústicas relevantes; Modelos acústicos, que utilizan Modelos Ocultos de Markov (HMM) o Redes Neuronales Profundas (DNN) para modelar con precisión los sonidos del habla y sus transiciones; Modelos de lenguaje, que emplean técnicas estadísticas o basadas en redes neuronales para predecir secuencias de palabras probables, mejorando la precisión del reconocimiento; y Adaptación al hablante, que involucra técnicas para ajustar los modelos a las características específicas de la voz del usuario, optimizando el rendimiento para cada individuo.

## **Desarrollo de videojuegos**

El desarrollo de videojuegos, un proceso multifacético y técnicamente complejo, ofrece un potencial significativo para la creación de herramientas terapéuticas innovadoras en el campo del habla y el lenguaje. Este proceso involucra varios componentes clave: Motores de juego como Unity, Unreal Engine y Godot, que proporcionan frameworks robustos para el desarrollo multiplataforma; Gráficos y renderización, que incluyen técnicas avanzadas como shaders personalizados y optimizaciones de rendimiento (por ejemplo, Level of Detail y occlusion culling); Simulación física, utilizando motores como PhysX o Bullet para manejar colisiones y dinámicas realistas; y Inteligencia Artificial, que emplea algoritmos de pathfinding, máquinas de estados finitos, árboles de comportamiento y aprendizaje por refuerzo para crear experiencias adaptativas [12].

## **Diseño de interfaces para niños**

El diseño de interfaces para niños en el contexto de la terapia del habla requiere un enfoque multidisciplinario que integre principios de diseño centrado en el usuario, tecnologías avanzadas y consideraciones terapéuticas específicas. Este proceso debe tener en cuenta las capacidades cognitivas, motoras y emocionales en desarrollo de los usuarios jóvenes, enfatizando la simplicidad, la interactividad y el uso de elementos visuales atractivos. Como señalan Gossen et al. [13], los niños requieren interfaces que brinden soporte emocional, cognitivo e interactivo, con elementos visuales grandes y fáciles de manipular, lo que resulta crucial en aplicaciones terapéuticas. La usabilidad en aplicaciones infantiles para terapia del habla implica la creación de controles intuitivos, instrucciones claras y feedback inmediato, mientras que la accesibilidad debe considerar posibles limitaciones físicas o cognitivas. Las consideraciones éticas, incluyendo la protección de la privacidad y la promoción de un uso saludable de la tecnología, son fundamentales. Esto se manifiesta en la creación de escenarios de práctica inmersivos y dinámicos utilizando RV para simular situaciones donde se practica el habla.

## 2.2. Trabajos relacionados

En el campo de la rehabilitación del habla para niños, se han realizado importantes avances en la integración de tecnología para mejorar la eficacia de las terapias. Estos desarrollos han sentado las bases para nuestro proyecto, que buscó crear un prototipo de sistema interactivo basado en videojuegos y realidad virtual para la mecanización de palabras en niños de 6 a 9 años.

Un proyecto particularmente relevante para nuestro trabajo es Talking to Teo, desarrollado por Navarro-Newball et al. en 2014 [4] en colaboración con el Instituto para Niños Ciegos y Sordos del Valle del Cauca (INCSVC). Este sistema, diseñado para niños con implantes cocleares y otros dispositivos auditivos, combina ejercicios de mecanización de palabras con un enfoque gamificado. Talking to Teo utiliza reconocimiento de voz basado en coeficientes cepstrales en la frecuencia Mel (MFCC) para analizar la pronunciación en tiempo real, ofreciendo retroalimentación visual y auditiva. El sistema consta de tres módulos principales: reconocimiento de voz, interfaz para terapeutas y un videojuego interactivo. Los resultados de la evaluación mostraron un aumento significativo en el interés y compromiso de los niños, mejor adherencia a la terapia y una evaluación positiva por parte de los terapeutas. Sin embargo, se identificaron limitaciones como la confusión en el sistema de retroalimentación visual y la falta de integración completa entre el juego y el sistema del terapeuta. Nuestro proyecto buscó construir sobre los principios exitosos de Talking to Teo, retomando sus limitaciones mediante la incorporación de realidad virtual y una mejor integración entre la interfaz del terapeuta y el juego, con el objetivo de proporcionar un entorno más inmersivo y apoyar a la terapia del habla.

Un ejemplo destacado de estos esfuerzos es el sistema Pronuntia, desarrollado por Desolda et al. en 2021 [14]. Este sistema representa un paso significativo en la evolución de las herramientas de apoyo para la terapia del habla, abordando muchas de las limitaciones de los enfoques tradicionales. Pronuntia fue diseñado como un sistema integrado de rehabilitación remota, con el objetivo principal de asistir a los terapeutas del habla en el manejo de pacientes infantiles con trastornos del lenguaje primario. Una de sus características más innovadoras

es su enfoque en hacer que los ejercicios de terapia sean más atractivos para los niños, transformando los ejercicios tradicionales en juegos interactivos. Esta gamificación de la terapia es un aspecto que nuestro proyecto también busca incorporar y mejorar aún más.

La evolución de la gamificación y la realidad virtual en la terapia del habla ha sido significativa. Rodríguez et al. (2023) desarrollaron un sistema basado en realidad virtual y gamificación para la rehabilitación lingual, empleando un sistema de realidad virtual con cámaras externas para capturar movimientos de la lengua y ejercicios gamificados que mejoran la motivación de los pacientes [15]. Este enfoque facilita sesiones más interactivas y entretenidas, lo que resulta en una mayor adherencia a la terapia y una experiencia de usuario más envolvente. Esta propuesta está alineada con el objetivo de nuestro proyecto, que también busca utilizar un enfoque lúdico para mejorar la rehabilitación del habla en niños, aumentando la motivación y el compromiso de los usuarios.

El diseño de juegos terapéuticos para la rehabilitación del habla ha demostrado ser una estrategia efectiva. Estudios como el de Duval et al. (2017) han evidenciado que los juegos interactivos pueden transformar la práctica repetitiva en una actividad divertida, lo que motiva a los niños a participar más activamente en la terapia [16]. Desarrollaron un prototipo de juego llamado SpokeIt, que emplea técnicas de reconocimiento de voz para ayudar a los niños con problemas de habla. Los resultados preliminares mostraron que los niños estaban más motivados para practicar el habla con el juego en comparación con los métodos tradicionales. En nuestro proyecto, este enfoque es relevante, ya que se integraron mecanismos de retroalimentación inmediata y recompensas lúdicas para mejorar la adherencia de los niños a las sesiones de mecanización de palabras.

Los estudios longitudinales en juegos para la terapia del habla han proporcionado evidencia sólida sobre la efectividad de estas intervenciones a largo plazo. Un estudio sobre el juego de terapia del habla Apraxia World demostró que los niños lograron mejoras significativas en la pronunciación tras varias semanas de uso del juego [17]. Este juego basado en una plataforma 2D fue diseñado para motivar a los niños mediante niveles de juego que integran ejercicios de pronunciación con retroalimentación automática o evaluaciones por parte de cuidadores. Los resultados sugieren que las intervenciones a largo plazo que incorporan juegos pueden ser tan

efectivas como las terapias tradicionales, lo que resalta el potencial de nuestro proyecto para brindar una herramienta de rehabilitación efectiva a través de juegos interactivos.

Las ventajas de la gamificación en la rehabilitación del habla han sido ampliamente documentadas. Según diversos estudios, la gamificación y el uso de entornos virtuales han demostrado ser estrategias efectivas para mejorar la implicación de los pacientes en sesiones de rehabilitación. La investigación de Rodríguez et al. (2023) destaca la importancia de adaptar las tecnologías a las necesidades de los terapeutas, ofreciendo interfaces fáciles de usar y entornos que permiten a los pacientes realizar ejercicios de manera autónoma [15]. Esto es clave para nuestro proyecto, ya que buscamos crear un sistema accesible tanto para terapeutas como para los niños que usan la plataforma.

Un enfoque innovador y prometedor en el campo de la terapia del habla es el uso de asistentes de voz (VA, por sus siglas en inglés). Qiu y Abdullah (2021) exploraron el potencial de los VA para la entrega remota de terapias del habla a escala [18]. Los VA utilizan el reconocimiento de voz y el procesamiento del lenguaje natural para entender las expresiones del usuario y responder en consecuencia, lo que los hace particularmente adecuados para la terapia del habla, que se lleva a cabo como conversaciones interactivas entre clínicos y pacientes.

Los VA ofrecen varias ventajas únicas para la terapia del habla. En primer lugar, pueden proporcionar una interacción de voz bidireccional más natural y realista en comparación con las interacciones de pantalla táctil. Esto puede conducir a resultados más prácticos en términos de habilidades de comunicación cotidiana. Además, la interacción por voz puede mejorar significativamente la accesibilidad de las tecnologías para personas con diversos niveles de discapacidad (por ejemplo, lenguaje, cognición y función motora).

Otro aspecto crucial que los VA pueden abordar es la alianza terapéutica, un factor determinante en los resultados de la terapia del habla, el compromiso del usuario y la satisfacción. Los VA podrían aplicar técnicas de comunicación verificadas por los logopedas a lo largo del proceso de terapia para facilitar el compromiso del usuario y la eficacia de la terapia [18].

En conclusión, estos antecedentes demuestran el potencial y la efectividad de integrar tecnologías como la realidad virtual, los videojuegos y los asistentes de voz en la terapia del

habla para niños. Nuestro proyecto busca construir sobre estos fundamentos, incorporando las lecciones aprendidas de sistemas como Pronuntia y otros estudios relevantes, con el objetivo de crear una herramienta más efectiva y atractiva para la terapia del habla infantil.

## 3. METODOLOGÍA

### 3.1. Enfoque de Prototipado Evolutivo

#### 3.1.1. ¿En qué consiste?

El prototipado evolutivo es una estrategia de desarrollo iterativa en la que el producto final surge mediante la evolución sucesiva de prototipos funcionales. Cada versión parcial se construye, se valida con los usuarios y, con base en la retroalimentación recibida, se refina o amplía; así, el prototipo no se desecha, sino que crece hasta transformarse en el sistema definitivo.

#### 3.1.2. Principios fundamentales

1. **Iteración continua:** diseñar → construir → evaluar → ajustar, en ciclos cortos.
2. **Feedback temprano de usuarios:** las validaciones frecuentes reducen el riesgo de desviarse de las necesidades reales.
3. **Entrega incremental de valor:** cada versión agrega funcionalidad utilizable.
4. **Flexibilidad ante requisitos cambiantes:** el diseño se adapta conforme se aprende más del dominio.
5. **Documentación ligera y viva:** se actualiza al ritmo de los cambios del prototipo.

#### 3.1.3. Fases típicas

**Fase 1 Exploratoria:** Bocetos de la idea, definición de requerimientos esenciales y construcción del primer prototipo mínimo.

**Fase 2** Expansión: Incorporación iterativa de funciones clave; cada incremento se somete a pruebas de usabilidad.

**Fase 3** Consolidación: Refactorización del código, optimización de rendimiento y aseguramiento de la calidad.

**Fase 4** Validación final: Pruebas formales con usuarios meta, cierre de requisitos y preparación para despliegue.

### **3.1.4. Ventajas y desventajas**

#### **Ventajas**

- Riesgo técnico reducido: problemas se detectan en prototipos tempranos.
- Alta alineación con usuarios: ajustes inmediatos según su retroalimentación.
- Motivación del equipo y de los stakeholders al ver avances tangibles.

#### **Desventajas**

- Posible crecimiento desorganizado del código si no se planifican refactorizaciones.
- Gestión de alcance más compleja: tentación de “añadir y añadir” requisitos.
- Difícil estimar esfuerzo total al inicio, pues el alcance se define progresivamente.

Para mitigar riesgos como el crecimiento desorganizado del código y la complejidad en la gestión del alcance, en el desarrollo de "Fonetic Tower", se adoptó un enfoque de trabajo basado en sprints definidos y complementado con refactorizaciones de código periódicas.

### **3.1.5. Adecuación al presente proyecto**

La terapia del habla asistida con realidad virtual presenta retos muy específicos que hacen aconsejable un enfoque de prototipado evolutivo:

1. **Validación continua con usuarios finales.** El público objetivo (niños de 6–9 años) tiene ventanas de atención reducidas, variabilidad cognitiva y distinta tolerancia al vértigo por RV. Probar versiones tempranas permite medir aceptación, usabilidad y síntomas de ciber-mareo antes de invertir esfuerzos en funcionalidades avanzadas. Cada iteración recoge métricas objetivas (tiempo de sesión, número de intentos correctos, trayectorias de mirada) y subjetivas (encuestas de diversión, escala de presencia), vitales para decidir qué mantener, descartar o refinar.
2. **Refinamiento clínico incremental.** Los ejercicios de mecanización deben alinearse con los objetivos terapéuticos fijados por el profesional; sin retroalimentación temprana podrían volverse inadecuados o demasiado fáciles. Mediante ciclos cortos, los terapeutas revisan pronunciaciones, nivel de inmersión y carga cognitiva, y solicitan variaciones por ejemplo, cambiar la secuencia de fonemas o añadir apoyos visuales para un fonema difícil. Así, el contenido evoluciona al ritmo del tratamiento, no al ritmo del calendario de desarrollo.
3. **Gestión de riesgos técnicos y de seguridad.** La RV exige 90 fps estables para evitar molestias físicas. Iterar sobre un solo prototipo “vivo” posibilita perfilar rendimiento, ergonomía de controles y consumo de memoria conforme se introducen nuevos minijuegos.
4. **Adaptación a requisitos éticos y normativos.** Al tratar con población pediátrica se exige aprobación de comités de ética y consentimiento informado. Un prototipo funcional pero acotado facilita presentar evidence packages graduados (capturas, registros anonimizados, bitácoras de cambios) y obtener autorizaciones progresivas. Cambios regulatorios o recomendaciones del comité pueden incorporarse en el siguiente ciclo sin rehacer desde cero.
5. **Evolución del sistema a largo plazo.** Se proyecta que el videojuego amplíe vocabulario, integre reconocimiento automático de voz y exporte reportes clínicos estandarizados (FHIR). El enfoque evolutivo permite introducir estos módulos gradualmente, mante-

niendo compatibilidad descendente y testeando cada componente en aislamiento antes de acoplarlo al núcleo jugable.

En síntesis, el prototipado evolutivo *encaja* con este proyecto porque:

- Alinea desarrollo técnico y exigencias terapéuticas bajo un bucle de build measure learn.
- Reduce costes de rectificación: los fallos se descubren cuando aún son baratos de corregir.
- Fomenta la co-creación con fonoaudiólogos y con los propios niños, elevando la pertinencia clínica y la motivación lúdica.
- Garantiza que el producto final sea seguro, efectivo y escalable, sustentado en evidencia empírica recogida a lo largo de todo el proceso.

## **3.2. Actividades planificadas**

### **3.2.1. Análisis de la terapia de mecanización de palabras**

1. Revisión sistemática de literatura sobre aprendizaje del habla infantil.
2. Entrevistas semiestructuradas con terapeutas del INCSVC.
3. Observación de sesiones clínicas para detectar actividades interactivas efectivas.

### **3.2.2. Investigación sobre RV y videojuegos en terapia**

1. Estudio comparativo de casos de uso de RV en rehabilitación del habla.
2. Análisis de ventajas y limitaciones de hardware/software existentes.
3. Consultas con profesores expertos en RV y game-based learning.

4. Formulación de un marco conceptual para integrar RV, gamificación y mecanización de palabras.

### **3.2.3. Implementación del prototipo**

1. Diseño arquitectónico basado en hallazgos de las fases 1 y 2.
2. Desarrollo iterativo:
  - a) Creación de un entorno RV básico.
  - b) Implementación de mecánicas de juego para ejercicios fonéticos.
  - c) Diseño de actividades motivadoras y sistema de recompensas.
  - d) Integración de retroalimentación supervisada.
  - e) Diseño de UI accesible para niños y terapeutas.

### **3.2.4. Validación de la efectividad**

1. Pruebas internas de usabilidad y refinamiento continuo.
2. Reclutamiento de participantes junto con el INCSVC.
3. Recopilación de datos:
  - a) Registros de uso: precisión, tiempo y flujo de navegación.
  - b) Recopilación de percepciones de niños y profesionales:
    - Se realizaron entrevistas semiestructuradas individuales con cada participante (niños y terapeutas) durante y después de las sesiones de prueba. A los niños se les hicieron preguntas abiertas sobre su experiencia con el minijuego (e.g., ¿Qué te pareció el juego? ¿Entendiste lo que tenías que hacer? ¿Te resultó cómodo el visor?), además de registrar su edad y nombre. A los terapeutas se les consultó sobre la adecuación de los niveles para las edades y necesidades de los niños, la claridad de las mecánicas, la comodidad del dispositivo

desde su perspectiva, y la utilidad general del prototipo como herramienta terapéutica.

### 3.2.5. Fases del desarrollo

El desarrollo del prototipo se organizó en etapas iterativas tales que:

**Investigación y diseño conceptual:** Se realizó una revisión bibliográfica de la terapia de mecanización y de casos previos (TalkingToTeo, prototipos del INCSVC). Se definieron los requisitos principales y la temática central (una torre de mago) basada en el análisis del modelo de Ling. Paralelamente, se esbozó un storyboard en papel, priorizando los elementos clave del juego.

**Modelado 3D y assets:** Se exploraron bibliotecas de objetos (BlenderKit, Unity Asset Store) en busca de posibles modelos de torre y elementos mágicos. Al no encontrar assets que cumplieran con los requisitos, se decidió modelar la torre y los elementos en Blender. Esto incluyó la creación de una torre escalonada con texturas simples y estilizadas.

**Prototipado de mecánica básica:** En Unity se implementó inicialmente el Nivel 1. Se configuró la escena con el escenario de laboratorio mágico, el sistema de triggers de voz simulada y las primeras animaciones. Se validó que la lógica respondiera correctamente a los comandos de voz supervisados, simulando el comportamiento de un terapeuta por un miembro del equipo de desarrollo.

**Desarrollo de niveles sucesivos:** Se programaron los niveles restantes incrementando la dificultad. Cada nuevo nivel integró más textos y reglas (sílabas compuestas, control de entonación).

**Implementación de la interfaz de usuario y retroalimentación:** Se diseñaron paneles de feedback. Además de menús sencillos para inicio y selección de nivel. Se incluyeron señales visuales (destellos mágicos, partículas) al completar ejercicios.

**Pruebas preliminares y ajuste:** Una vez integrado el prototipo básico, se realizaron sesiones de prueba con terapeutas del INCSVC. Se recopiló su retroalimentación sobre la usabilidad y la adecuación del nivel de dificultad. Con base en estos comentarios, se ajustaron parámetros

del juego (tiempo de repetición, cantidad de intentos, sensaciones visuales) para mejorar la experiencia.

**Iteración final y optimización:** En la fase final se corrigieron errores menores y se optimizaron modelos y scripts para asegurar un rendimiento fluido en Meta Quest 3. Se dejó listo el prototipo completo con los siete niveles funcionales, preparado para su validación formal en entornos clínicos.

Cada una de estas fases se documentó detalladamente en la bitácora del proyecto (fechas, tareas realizadas, decisiones de diseño), lo que permitió un desarrollo ordenado y controlado.

### **3.2.6. Interacciones con Usuarios Durante el Desarrollo y Validación**

A lo largo del desarrollo del prototipo, se llevaron a cabo múltiples interacciones con los usuarios. En las fases iniciales, se realizaron entrevistas semiestructuradas con terapeutas del Instituto para Niños Ciegos y Sordos del Valle del Cauca (INCSVC) para analizar la terapia de mecanización de palabras y recopilar requisitos. Adicionalmente, se efectuaron observaciones de sesiones clínicas. Una vez desarrollado un prototipo básico, se llevaron a cabo sesiones de prueba preliminares con estos terapeutas para obtener retroalimentación sobre la usabilidad y la adecuación de los niveles de dificultad, lo que permitió realizar ajustes tempranos al diseño.

Finalmente, se realizó una prueba piloto formal del prototipo con la participación de cuatro niños (con edades entre 7 y 11 años y diversas características, incluyendo algunos con discapacidad auditiva y una con baja visión) y la presencia de terapeutas del INCSVC. Durante esta sesión de prueba piloto, cada niño interactuó con el sistema durante un tiempo aproximado de 10-15 minutos dedicados a la realización de 2 a 4 ejercicios específicos, precedidos por una explicación de la dinámica. La metodología empleada incluyó la observación directa de la interacción de los niños con el juego, y la recopilación de datos cualitativos mediante entrevistas semiestructuradas individuales con los niños y los terapeutas, realizadas durante y después de las sesiones de prueba.

### **3.2.7. Alcance de la Documentación**

Para mantener la claridad del presente documento, las siguientes secciones se centran en describir y analizar la versión final y consolidada del prototipo. Si bien el desarrollo siguió un proceso evolutivo con múltiples iteraciones intermedias, este documento omite los detalles de dichas versiones para concentrarse en los resultados y validaciones del producto final.

#### 4. REQUERIMIENTOS DEL SISTEMA

Los requerimientos constituyen la fuente principal de verificación del proyecto: definen el alcance, orientan el diseño y sirven como criterio objetivo para las pruebas de aceptación. Si están bien redactados, claros, completos y medibles, minimizan el riesgo de malinterpretaciones entre los desarrolladores, los terapeutas y los demás interesados.

Existen dos grandes categorías:

- **Requisitos funcionales (RF)** – especifican qué debe hacer el sistema: servicios, flujos de trabajo y reglas de negocio visibles para el usuario.
- **Requisitos no funcionales (RNF)** – describen cómo debe comportarse el sistema: atributos de calidad (rendimiento, seguridad, ergonomía, etc.) o restricciones tecnológicas y normativas.

A continuación en el cuadro 4.1 se listan los requisitos identificados a partir de la revisión clínica, las entrevistas con terapeutas y los primeros prototipos.

## Requisitos funcionales

Cuadro 4.1: Requisitos funcionales (RF)

ID	Descripción
RF-01	<b>Interacción por voz mediada por terapeuta.</b> El juego permitirá la interacción basada en la pronunciación de fonemas o palabras clave por el usuario. La validación de dicha pronunciación y la consecuente activación de comandos en el juego (ej.: decir “co” para activar un objeto) será gestionada por el terapeuta en esta fase del prototipo.
RF-02	<b>Progresión de niveles.</b> El sistema debe desbloquear automáticamente el siguiente nivel al completar los ejercicios del nivel actual, incorporando nuevos fonemas, sílabas o reglas de entonación.
RF-03	<b>Tres minijuegos temáticos.</b> El prototipo debe incluir, al menos uno de los minijuegos sombrero mágico, preparación de pociones o lanzamiento de hechizos, vinculados a un tipo de tarea fonética.
RF-04	<b>Retroalimentación inmediata.</b> Tras cada intento, el sistema debe mostrar animaciones o sonidos mágicos de éxito, o sugerir la repetición cuando la pronunciación sea incorrecta.
RF-05	<b>Registro de datos.</b> El sistema debe almacenar aciertos, errores y tiempo por ejercicio, generando un expediente por sesión para uso del terapeuta.

El Cuadro 4.2 presenta los requisitos no funcionales.

## Requisitos no funcionales

Cuadro 4.2: Requisitos no funcionales (RNF)

ID	Descripción
RNF-01	<b>Rendimiento en RV.</b> El juego debe mantener al menos 90 fps estables en Meta Quest 3 para evitar mareo cinético.
RNF-02	<b>Seguridad y ergonomía.</b> La interfaz debe ser segura y cómoda para niños: controles accesibles, ausencia de estímulos excesivamente luminosos y límites de tiempo sugeridos por el terapeuta.
RNF-03	<b>Escalabilidad de contenido.</b> La arquitectura debe permitir añadir nuevos niveles, palabras y minijuegos sin reescribir componentes existentes.
RNF-04	<b>Compatibilidad multiplataforma.</b> Debe compilarse una versión de prueba en Windows y tener posibilidad de exportación a Android (OpenXR) para Meta Quest 3.

### 4.1. Verificación Interna de Requerimientos

Antes de proceder a las pruebas de validación con los usuarios finales (niños), se llevó a cabo un proceso de verificación interna para asegurar que el prototipo desarrollado cumplía con los requisitos funcionales (Ver Tabla 4.1) y no funcionales (Ver Tabla 4.2) definidos en este capítulo. Este proceso incluyó:

- **Pruebas unitarias y de integración (implícitas en el desarrollo evolutivo):** *El enfoque de prototipado evolutivo implicó la construcción y prueba incremental de componentes. Si bien no se detalla un plan formal de pruebas unitarias, la metodología conllevó la depuración y ajuste continuo de las mecánicas del juego y la lógica del sistema a*

*medida que se desarrollaban.*

- ***Validación temprana con terapeutas:*** *Como parte de las fases de desarrollo (sección 3.2.5), el prototipo básico fue presentado a terapeutas del INCSVC. Su retroalimentación sobre la usabilidad, la adecuación de los niveles de dificultad iniciales y la alineación con los objetivos terapéuticos sirvió como una forma temprana de verificación de los requisitos clave, especialmente aquellos relacionados con la pertinencia clínica y la interacción básica.*
- ***Cumplimiento de requisitos técnicos:*** *Se prestó especial atención al requisito de rendimiento en RV (RNF-01) durante la optimización de modelos y scripts para asegurar un funcionamiento fluido en Meta Quest 3.*

## 5. ANÁLISIS Y DISEÑO

### 5.1. Modelo terapéutico de Daniel Ling y adaptación al juego

El diseño de Fonetic Tower se basa en los principios de la terapia auditivo-verbal de Daniel Ling, la cual establece una jerarquía de habilidades auditivas a desarrollar en niños con deficiencias auditivas. Este modelo terapéutico comprende cuatro niveles principales: detección, discriminación, identificación y comprensión del sonido. En la detección, el niño reconoce la presencia o ausencia de un sonido; en la discriminación, distingue si dos sonidos son iguales o diferentes; en la identificación, reconoce qué sonido específico es (por ejemplo, asociándolo a un objeto o reproducirlo verbalmente); finalmente, la comprensión implica entender el significado del sonido o palabra en contexto. Adicionalmente, Ling propuso la conocida prueba de los seis sonidos (/a/, /i/, /u/, /sh/, /s/, /m/) para verificar si el niño percibe todo el espectro de frecuencias del habla. La Figura 5.1 reproducida de [4], ilustra los estados de aprendizaje de consonantes según este enfoque.

Stg.	occlus.	nasal	trill	lat.	fri.	affri.
I	/b//p/	/m/			/b//f//y/	
II	/d//t/	/n/		/l/	/d//s//sh/	
III	/g//k//b/	/ñ/	/r/		/g//j//b//d//g/	/ch/
IV	/p//t//k/		/r/		/p//t//k/	

Figura 5.1: Estados de aprendizaje de consonantes

En Fonetic Tower se adaptó este modelo auditivo-verbal al contexto de un juego de realidad virtual. Inicialmente se planificó que cada etapa de la jerarquía de Ling estuviera representada por un tipo de ejercicio o minijuego dentro del entorno lúdico.

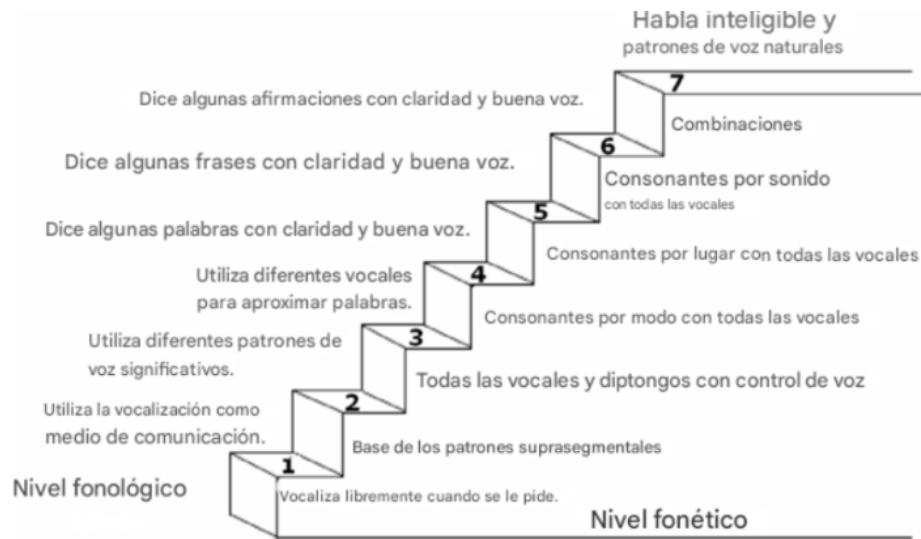


Fig. 1. Orden natural de adquisición de los sonidos, según [5].

Figura 5.2: Orden natural de adquisición de sonidos

Por ejemplo, se propuso un ejercicio de detección donde el niño simplemente reaccionara a la aparición de un texto mágico (cumpliendo la tarea básica de notar un estímulo visual); otro de discriminación donde tendría que distinguir entre dos textos o estímulos visuales presentados por criaturas o elementos del juego; y ejercicios de identificación donde el niño leería un fonema o sílaba en un texto mágico flotante y debería repetirlo o señalar un objeto correspondiente. Finalmente, en un nivel avanzado se contempló incorporar comprensión lectora y auditiva (si el terapeuta lee la instrucción), donde el jugador seguiría instrucciones verbales o escritas más complejas dentro del juego. Todos estos ejercicios buscan reforzar las habilidades fonéticas y auditivas descritas por Ling pero de forma interactiva y motivadora. El videojuego convierte los fonemas y sílabas en "palabras mágicas", o comandos de voz que el niño debe pronunciar tras leerlos en un texto mágico, integrando así la práctica terapéutica dentro de la narrativa del juego. Cada nivel del juego fue concebido para enfocarse en ciertos sonidos o sílabas específicas (por ejemplo, iniciar con vocales y sonidos nasales como /m/ de baja frecuencia, y progresivamente incluir sonidos de frecuencias más altas como /s/) tal como sugiere el modelo de Ling para cubrir todo el rango del habla. Esto aseguraría que el niño practicara desde sonidos simples hasta combinaciones más complejas, alineado con un

programa de rehabilitación fonética integral. Este orden se alinea con la adquisición natural de sonidos, como se observa en la Figura 5.2 reproducida de [4].

## 5.2. Concepto inicial del juego: torre mágica y múltiples minijuegos

En la fase de diseño conceptual, Fonetic Tower se imaginó como una torre mágica con múltiples salas o niveles, cada uno albergando un minijuego terapéutico distinto. La metáfora de la torre permitía representar de forma lúdica el progreso del niño a través de distintas etapas de rehabilitación: a medida que "sube", en la torre, el jugador domina habilidades fonéticas cada vez más avanzadas. Cada sala de la torre estaría tematizada según un ejercicio fonético específico, manteniendo al jugador motivado con variedad de desafíos. Un esquema inicial de estos niveles se presenta en la Figura 5.3

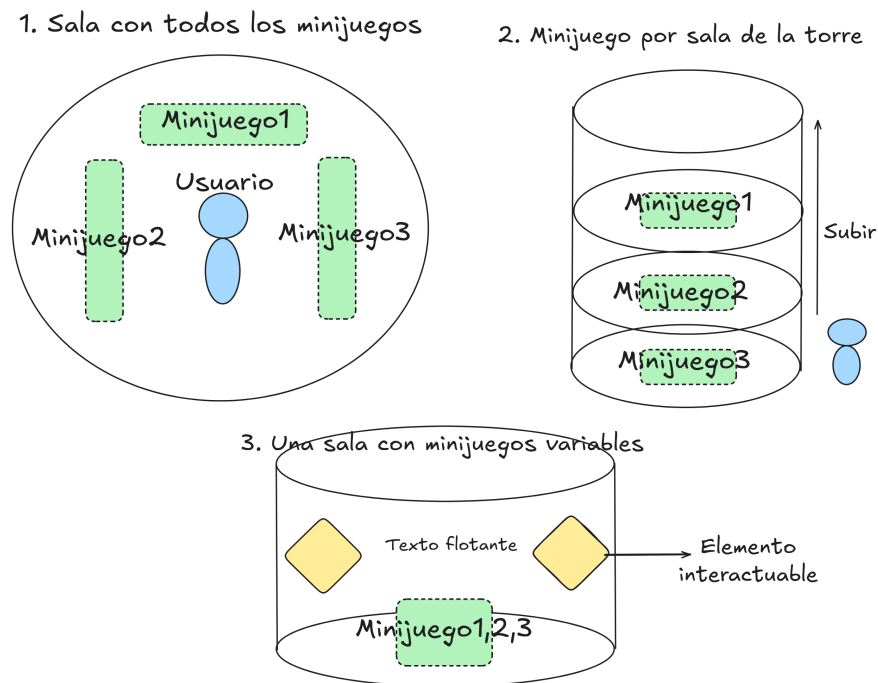


Figura 5.3: Esquema de niveles del prototipo

Por ejemplo, una sala inicial podía ser un bosque encantado para ejercicios de detección (el

niño debe ver un texto mágico aparecer entre los árboles y señalar que lo vio); otra sala podía ser el salón de los ecos para ejercicios de discriminación (donde diferentes objetos mostrarían textos y el jugador debe indicar si son iguales o diferentes); también se conceptualizó una sala de los espejos parlantes donde el niño leería un fonema de un texto mágico y debería identificarlo repitiéndolo correctamente, y el terapeuta validaría la respuesta. Finalmente, la cúspide de la torre podría haber presentado un desafío más complejo que implicara seguir instrucciones escritas o verbales del terapeuta (ejercitando la comprensión). La elección de una ambientación de fantasía (magos, pociones, criaturas mágicas) no fue casual: este entorno busca incrementar la motivación y atención del niño durante la terapia. La narrativa de la torre mágica convierte tareas terapéuticas en misiones de un juego. Por ejemplo, aprender a pronunciar la /s/, se transforma en "dominar el hechizo de la serpiente", dentro del juego, haciendo que el ejercicio fonético sea divertido y significativo para el jugador. Así mismo, un diseño multi-sala proporcionaría variedad de mecánicas para prevenir el aburrimiento: un minijuego de lectura y repetición en un nivel y otro de habla activa en el siguiente, manteniendo el interés. Desde el punto de vista del análisis, esta diversidad también permitiría adaptar el juego a distintas necesidades terapéuticas: dependiendo del diagnóstico del infante, se podrían enfatizar más ciertos minijuegos

### **5.3. Decisiones de diseño finales: enfoque en el minijuego de las pociones**

La implementación final, Fonetic Tower se simplificó a un único minijuego principal: Preparación de pociones. Es decir, en lugar de desarrollar múltiples salas y mecánicas diferentes, el juego final presenta una sola escena o entorno (el laboratorio del mago con su caldero) en la cual se llevan a cabo los ejercicios fonéticos. Esta decisión respondió a consideraciones prácticas de tiempo de desarrollo y complejidad, priorizando tener un minijuego pulido y funcional que abarque los objetivos terapéuticos esenciales, en lugar de varios minijuegos incompletos. El entorno virtual final con esta temática se muestra en la Figura 5.4.



Figura 5.4: Entorno virtual del juego con temática de torre de mago

El minijuego de preparación de pociones amalgama varios de los conceptos terapéuticos planteados en una sola mecánica central. En este juego, el niño asume el rol de un aprendiz de mago que debe preparar pociones pronunciando "palabras mágicas", (fonemas o sílabas objetivo de la terapia) que lee de un texto mágico flotante, usualmente presentado en un libro sobre una mesa mágica. La dinámica funciona así: el texto mágico le "pide", un ingrediente especial mostrando una sílaba o palabra (por ejemplo, "ma"). El niño debe pronunciar correctamente esa sílaba en voz alta; al hacerlo, y una vez que el terapeuta pulsa un botón para confirmar que la pronunciación fue correcta o incorrecta, luego si fue correcta el juego reconoce esta validación y ocurre una reacción mágica: suena un sonido de correcto, y se ilumina el ingrediente solicitado (por ejemplo, se ilumina una manzana, por haber pronunciado "ma" de manzana) que el jugador toma con sus manos virtuales y lo echa al caldero. Si el terapeuta indica que la pronunciación fue incorrecta, sonará un sonido distintivo de "fallo", y el juego registrará este intento. Cada ingrediente añadido acerca al jugador a completar los niveles. Cuando el niño logra agregar todos los ingredientes requeridos (es decir, pronuncia correctamente todas las sílabas de ese nivel, validadas por el terapeuta), la poción se completa con un efecto visual y se da paso al siguiente nivel. De esta forma, el minijuego de las pociones integra la identificación (lectura del texto mágico) y producción de sonidos como núcleo (el

jugador lee la sílaba objetivo y la produce correctamente, con validación del terapeuta).

Adicionalmente, incorpora algo de discriminación en la práctica, pues el sistema solo reacciona positivamente cuando el terapeuta da el visto bueno, lo que obliga al niño a afinar su escucha de sí mismo y corregir diferencias con el objetivo presentado en el texto. Varias mecánicas propuestas inicialmente quedaron fuera en la versión final. Por ejemplo, la idea de un juego de detección pura (solo reaccionar a un estímulo visual) no se implementó explícitamente como un minijuego separado; tampoco se desarrollaron los minijuegos independientes de discriminación auditiva o visual sin producción. Sin embargo, el diseño final trató de compensar estas ausencias dentro del propio minijuego de las pociones. En vez de salas separadas, se incorporaron elementos de esas mecánicas al flujo, con lo cual el niño practica la detección (debe ver el texto mágico) y la discriminación (distinguir ese texto de otros elementos visuales) antes de repetirlo. Así, aunque la variedad de escenarios se redujo, el núcleo terapéutico se mantuvo: el jugador sigue pasando por las fases de leer, identificar y producir el fonema, obteniendo retroalimentación mediada por el terapeuta. Finalmente, el concepto de la torre persiste de forma simbólica en la progresión por niveles. Si bien físicamente el juego ocurre en un solo cuarto (el del caldero), narrativamente se puede interpretar que cada nuevo conjunto de ingredientes o nuevas sílabas representa subir un peldaño en la torre del conocimiento fonético.

Esto conserva el sentido de avance y logro. El enfoque en un único minijuego permitió dedicar más detalle a la calidad de la interacción: se afinó el sistema de presentación de textos, se agregaron animaciones atractivas al caldero e ingredientes, y se garantiza una experiencia estable en RV - aspectos que podrían haberse resentido de haberse dispersado el esfuerzo en muchos minijuegos. Los elementos principales para la interacción dentro de este minijuego se pueden apreciar en la Figura 5.5.



Figura 5.5: Elementos principales para la interacción (items)

## 5.4. Estructura de niveles: planificación vs. producto final

Durante el análisis se definió una estructura de niveles que correlacionan con las etapas terapéuticas y los sonidos a practicar. En la implementación final, la estructura de niveles se centra en la producción verbal guiada por texto, validada por el terapeuta, y consta de siete niveles con complejidad creciente:

- Producción de consonantes aisladas o en una sílaba:
  - Ejercicio: Se presenta al niño un texto mágico flotante con una sílaba o una palabra simple (por ejemplo, "da"). Al leerlo y pronunciarlo, el terapeuta valida.
  - Reforzador: Si la pronunciación es correcta (validada por el terapeuta), suena un efecto de sonido positivo y se ilumina una fruta, verdura o poción que el niño puede echar al caldero. Este mecanismo de reforzador se aplica a todos los niveles.
- Repetición de una consonante con la misma vocal:
  - Ejercicio: El texto mágico muestra una secuencia repetitiva (por ejemplo, "di-di-di"). El niño debe leer y repetir la secuencia.
- Repetición con diferentes vocales:

- Ejercicio: El texto mágico presenta una secuencia de sílabas donde la consonante se mantiene pero cambia la vocal (por ejemplo, "di-du-da"). El niño debe leer y pronunciar la secuencia en orden.
- Alternancia de sílabas con diferentes consonantes:
  - Ejercicio: El texto mágico muestra una combinación de sílabas con diferentes consonantes y vocales (por ejemplo, "do-mo"). El niño debe leer y repetir la combinación.
- Producción con intensidad variable:
  - Ejercicio: El texto mágico podría indicar la necesidad de variar la intensidad (por ejemplo, "PA-PA-PA", con una indicación visual de tono creciente). El niño debe pronunciar la secuencia modulando la intensidad de su voz.
- Repetición con posición intervocálica y acento:
  - Ejercicio: El texto mágico presenta una palabra, posiblemente con sílabas resaltadas para indicar el acento (por ejemplo, "PA-to", y luego "pa-TO"). El niño debe leer y repetir, prestando atención al acento indicado.
- Variación de tono:
  - Ejercicio: El texto mágico presenta una frase o palabra con una indicación de emoción o entonación (por ejemplo, "Poción de fuego", en tono alegre, triste o enojado). El niño debe leer y repetir la frase aplicando la variación de tono solicitada. El terapeuta evalúa la correcta aplicación del tono.

El progreso de un nivel a otro incrementa el desafío al añadir sílabas más complejas, secuencias más largas, o requerir mayor control sobre aspectos prosódicos de la voz. La información sobre aciertos y errores en cada nivel se almacena vinculada al perfil del niño. El diagrama de la arquitectura general del sistema se presenta en la Figura 5.6.

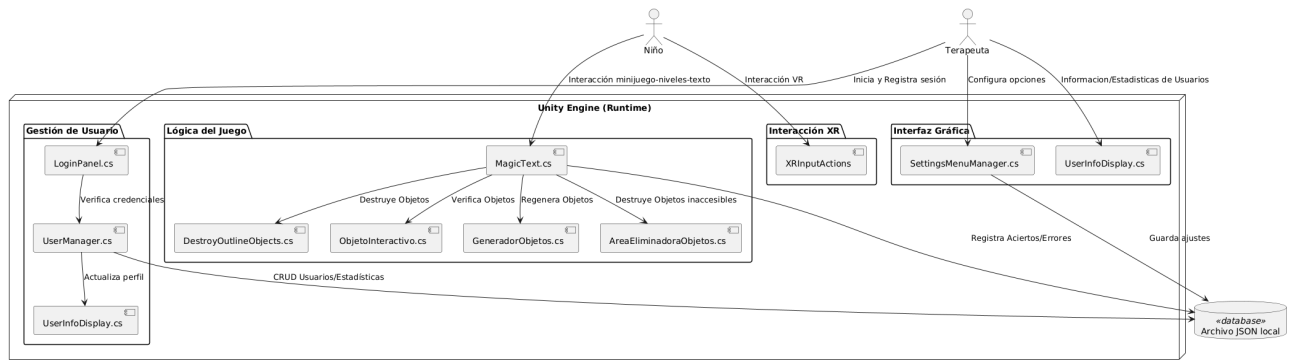


Figura 5.6: Diagrama de la arquitectura del sistema

El flujo detallado del minijuego de preparación de pociones se ilustra en la Figura 5.7.

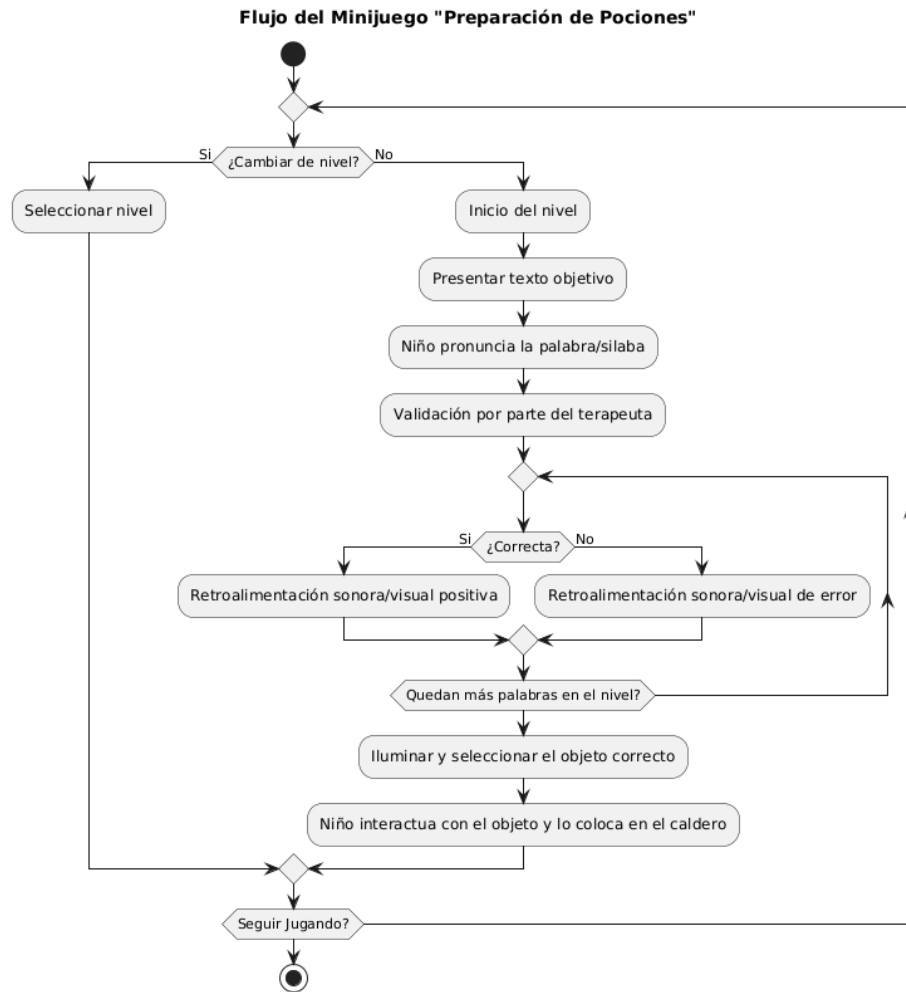


Figura 5.7: Diagrama de flujo del minijuego

## 5.5. Interfaz de usuario, interacción RV y retroalimentación

En el diseño de la interfaz de usuario (UI) de Phonetic Tower se tuvo en cuenta la población objetivo (niños en terapia) y las particularidades de la realidad virtual con Meta Quest 3. Se optó por una interfaz minimalista y altamente visual, evitando sobrecargar de texto o instrucciones complicadas. Al iniciar el juego, el usuario ve un menú de inicio sencillo

con el panel de login (para seleccionar o ingresar el nombre del jugador) y botones grandes y claros para comenzar la sesión. Este menú aparece como un panel flotante dentro del mundo virtual, colocado a una distancia cómoda y a la altura de los ojos del niño para facilitar la lectura y selección mediante los controles de RV. La elección de tipografías legibles y colores contrastantes fue importante para la accesibilidad: por ejemplo, las sílabas o palabras a pronunciar se muestran en un texto mágico flotante, en letras grandes y de color llamativo, de manera que el niño las pueda ver fácilmente dentro del casco de RV. Un esquema de la interfaz gráfica del prototipo se muestra en la Figura 5.8 y el esquema final puede ser visto en la Figura 5.9.

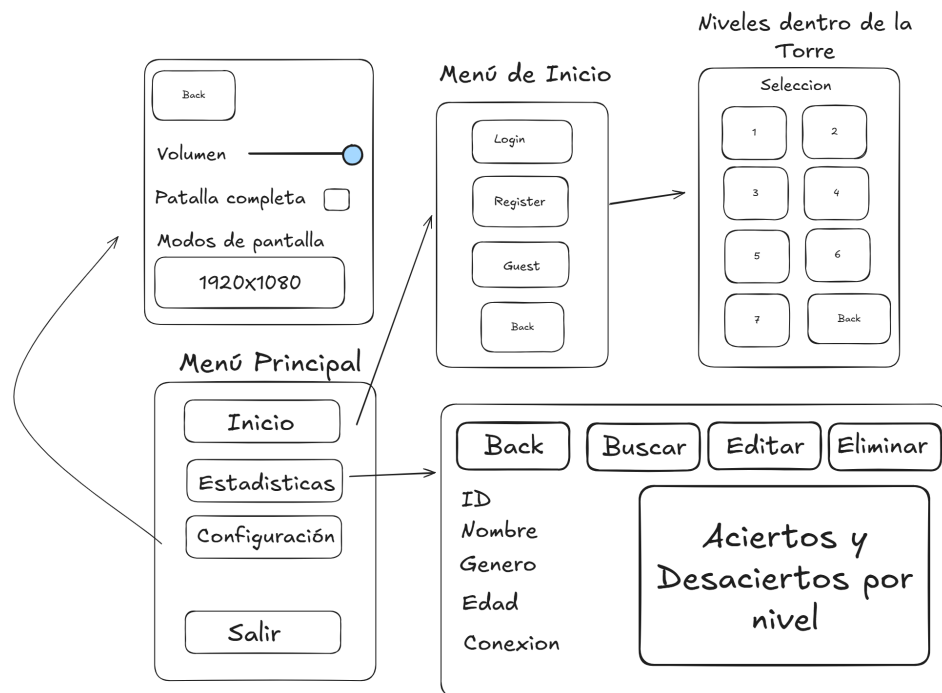


Figura 5.8: Interfaz grafica prototipo



Figura 5.9: Paneles de la interfaz gráfica de usuario con menús de inicio y registro

En cuanto a la interacción, *Fonetic Tower* aprovecha las capacidades del Meta Quest 3 tanto a nivel de hardware como de SDK. El jugador utiliza los controladores de movimiento del Quest 3 (uno en cada mano) para interactuar con los objetos virtuales. El sistema de seguimiento posicional 6DOF del Quest permite que el niño se agache, estire o mueva libremente por el espacio virtual limitado, lo cual enriquece la experiencia lúdica. Por ejemplo, cuando un ingrediente mágico aparece (tras la validación del terapeuta), el niño puede físicamente mover su brazo virtual (reflejado por el controlador) para agarrar ese objeto; esto se diseñó para ser lo más intuitivo posible: basta con acercarse

la mano virtual al ingrediente y apretar el gatillo del controlador para tomarlo, emulando el acto de agarrar un objeto real. Para indicar qué elementos son manipulables, el juego implementa un resaltado visual: al mirar o apuntar con la mano hacia un ingrediente, este se ilumina con un borde brillante u outline, señalando interactivamente que está listo para ser tomado.

La interacción principal, la “voz” del jugador, se logra mediante la pronunciación del niño después de leer el texto mágico. La validación de esta pronunciación no es automática por el sistema, sino que depende de la intervención del terapeuta, quien utiliza un control (botón) para indicar si la producción fue correcta o incorrecta. En el diseño se consideró la importancia de que esta mecánica fuera robusta, pero a la vez no frustrante. En caso de que el terapeuta marque un intento como fallido, el juego puede registrarlo y, opcionalmente, repetir la instrucción textual, siguiendo principios de retroalimentación constructiva, permitiendo motivar al niño a intentarlo de nuevo, y permite al terapeuta guiar el proceso. Los datos de aciertos y errores por nivel se almacenan para el seguimiento del progreso del niño.

La retroalimentación al jugador es un componente crítico del diseño de la UI/UX. Fonetic Tower implementa múltiples canales de feedback: visual, sonoro y cinestésico, muchos de ellos condicionados a la validación del terapeuta. En el plano visual, cada vez que el terapeuta valida una pronunciación correcta: el ingrediente asociado aparece con un destello y animación. Si la pronunciación es marcada como incorrecta por el terapeuta, simplemente no ocurre la interacción con los objetos; en su lugar, se reproduce un sonido específico de fallo. Durante el juego, hay música ambiental suave y efectos de sonido mágicos (por ejemplo, el chisporroteo del caldero) para mantener la inmersión; estos elementos auditivos también sirven para mantener la atención del niño y recompensar su participación activa. Por último, desde el punto de vista ergonómico, el diseño de la interacción RV consideró la seguridad y comodidad del niño. Se delimitaron las acciones dentro de un espacio reducido para que el jugador no necesite desplazarse físicamente más que un par de pasos (reduciendo riesgos de choques en el mundo real).

El Quest 3 permite definir un guardián (área segura), y el juego mantiene los objetos de interés dentro de ese rango. La torre mágica está ambientada de forma acogedora y no amenaza, evitando elementos que pudieran asustar o abrumar a niños pequeños; los colores son vivos pero equilibrados, y los personajes o elementos mágicos tienen apariencia amable. De esta forma, la interfaz y la interacción en Fonetic Tower fueron cuidadosamente diseñadas para que la experiencia terapéutica en RV sea intuitiva, motivadora y segura para los niños. La Figura 5.10 ofrece una vista de este entorno con un ingrediente resaltado y la sílaba propuesta.



Figura 5.10: Vista del entorno del caldero con un ingrediente resaltado listo para ser agarrado y la sílaba propuesta

## 6. IMPLEMENTACIÓN

### 6.1. Herramientas y tecnologías

El ecosistema tecnológico se eligió con el fin de equilibrar rendimiento en RV, curva de aprendizaje para un equipo universitario y disponibilidad de herramientas educativas de bajo costo. La Figura 6.1 resume la pila empleada; a continuación se describen sus componentes y la justificación de cada elección.

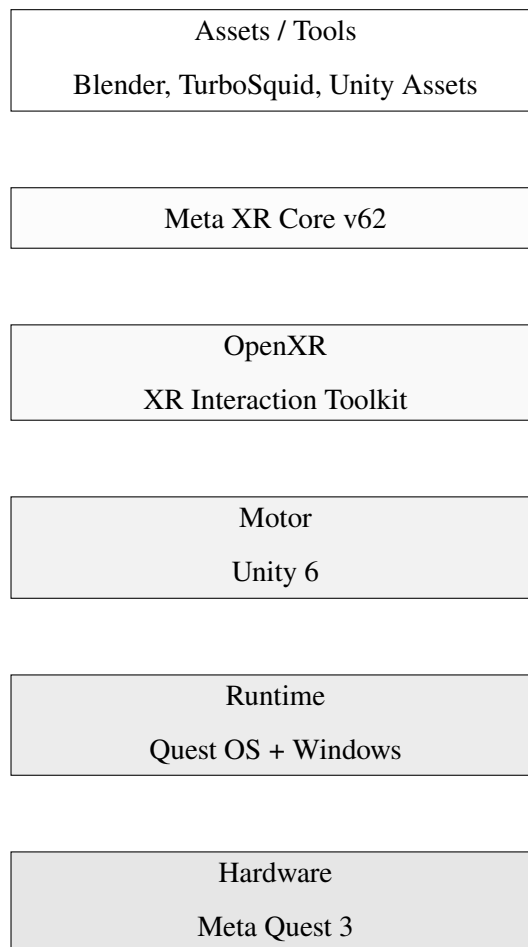


Figura 6.1: Pila tecnológica (hardware → contenido).

## Motor principal – Unity 6

**Motivación de la elección** Unity ofrece soporte nativo para Meta Quest 3, flujo de edición en caliente (play-mode) y un lenguaje asequible (C#) que la mayoría de los integrantes ya dominaba. Comparado con otros motores como se evidencia en el cuadro 6.1.

Cuadro 6.1: Comparación de motores RV

Criterio	Unity	Unreal 5	Godot 4
Lenguaje base	C#	C++ / Blueprints	GScript / C#
Curva de aprendizaje para prototipos	Baja	Alta	Media
Catálogo de plugins RV	>1 900	~600	>150
Herramientas clínicas pre-hechas (heatmaps, analíticas)	Amplias	Medias	Escasas
Licencia para proyectos académicos	Gratuita	Gratuita (royalties)	MIT

### Ventajas

- Motor probado en Quest: plantillas optimizadas para 72/90 fps y repositorio oficial de muestras RV.
- Asset Store con paquetes pedagógicos (Audio Visualizer, RV-Locomotion Toolkit, etc.) que reducen esfuerzo de desarrollo.
- C# y su sistema de componentes facilitan iterar mecánicas con pequeños scripts independientes.

### Desventajas

- Requiere atención al garbage collector; los stutters son notables en realidad virtual si no se optimiza la memoria.
- Motor cerrado: algunas optimizaciones de bajo nivel (render pipeline) no son accesibles sin licencias Enterprise.

El anexo 11 muestra como configurar el Meta Quest para uso en el sistema. El código fuente es accesible en [URL](#)

## 6.2. Desarrollo en Unity y Arquitectura General del Prototipo

El videojuego "Fonetic Tower", se desarrolló utilizando el motor Unity, con programación en C#. La plataforma objetivo fue Meta Quest 3 (ver Figura 6.2), empleando el XR Interaction Toolkit de Unity para las funcionalidades de Realidad Virtual, como el seguimiento de movimiento y la interacción con los controladores.



Figura 6.2: Gafas de realidad virtual Meta Quest 3 utilizadas

La arquitectura del software se adhiere a la organización modular típica de los proyectos en Unity, estructurándose en escenas y GameObjects con scripts (componentes lógicos) adjuntos. El sistema se compone de varios módulos principales que interactúan para ofrecer la experiencia de juego terapéutica:

- **Módulo de Lógica de Juego y Progresión:** Orquesta la secuencia de niveles, la presentación de los estímulos fonéticos (sílabas/palabras) y la mecánica central del minijuego de las pociones. Este módulo es el encargado de gestionar el avance del niño a través de los contenidos terapéuticos y de procesar la validación de la pronunciación realizada por el terapeuta.
- **Módulo de Interacción RV:** Dota a los elementos del juego (como frutas, verduras e ingredientes) de la capacidad de ser manipulados por el jugador en el entorno virtual. Incluye la gestión de la selección de objetos, el agarre y la respuesta a las acciones del niño, como depositar ingredientes en el caldero.
- **Módulo de Gestión de Usuarios e Interfaz:** Administra los perfiles de los jugadores, permitiendo el ingreso del nombre del niño y el almacenamiento de su progreso individual (intentos, aciertos, nivel alcanzado). Se encarga también de la interfaz de inicio y selección de perfil.
- **Módulo de Retroalimentación y Efectos:** Responsable de generar las respuestas visuales y sonoras del juego, como la iluminación de ingredientes, efectos de partículas, y sonidos de éxito o fallo. Estos elementos son cruciales para la gamificación y la motivación del usuario.

Esta arquitectura modular facilita la interconexión de los componentes. Donde, la lógica de juego activa elementos del módulo de interacción (como la iluminación de un ingrediente tras una pronunciación correcta validada) y registra el progreso a través del módulo de gestión de usuarios.

### 6.3. Componentes Clave Implementados:

- **Coordinador del Juego (Controlado por MagicText.cs):** Este script, adjunto a un objeto central en la escena (como la mesa del libro mágico), es el núcleo lógico de "Fonetic Tower". Sus responsabilidades incluyen:
  - **Definición y Presentación de Contenidos Terapéuticos:** Codifica el currículo de sílabas, palabras y frases objetivo para cada uno de los siete niveles. Utiliza estructuras de datos, como listas o arrays, para almacenar los estímulos textuales que se mostrarán en el libro mágico virtual (usando TextMeshPro) en cada fase del juego.
  - **Control del Flujo de Niveles y Avances:** Gestiona el estado interno del juego, indicando el nivel actual y el estímulo fonético correspondiente. Al iniciar o tras el login, presenta el primer texto y espera la pronunciación del niño, seguida de la validación del terapeuta mediante una señal externa (ej. pulsación de un botón).
  - **Gestión de Eventos y Retroalimentación:** Tras la validación del terapeuta, si la pronunciación es correcta, el script instancia el prefab del ingrediente asociado al estímulo y activa efectos visuales (como partículas) y sonoros positivos. Si es incorrecta, registra el intento fallido. Esta información (acierto/fallo) se almacena.
  - **Integración con Datos de Usuario:** Comunica los resultados de cada interacción (aciertos, fallos, nivel completado) al UserManager.cs para que se registren en el perfil del niño, facilitando el seguimiento del progreso.
- **Gestión de Interactividad de Objetos (Mediante ObjetoInteractivo.cs y DestroyOutlineObjects.cs):** Para que el entorno virtual sea manipulable, los objetos del juego, especialmente los ingredientes, cuentan con el componente ObjetoInteractivo.cs.
  - **Mecánicas de Interacción RV:** Este script gestiona la detección de la proximidad o el apuntado del controlador RV del niño hacia un objeto, activando un resaltado visual (outline) para indicar que es interactuable. Utiliza el XR Interaction Toolkit

para las mecánicas de agarre, permitiendo al niño tomar los ingredientes y, por ejemplo, echarlos al caldero, el cual posee un collider de tipo trigger para detectar esta acción.

- Optimización de Recursos: DestroyOutlineObjects.cs se encarga de eliminar de la escena los objetos interactivos que ya no son necesarios y sus efectos visuales asociados, lo cual es crucial para mantener un rendimiento fluido en la Meta Quest 3, que tiene recursos limitados.
- Manejo de Perfiles de Usuario (Controlado por LoginPanel.cs y UserManager.cs): Este sistema es fundamental para el uso terapéutico con múltiples niños y sesiones.
  - Interfaz de Inicio: LoginPanel.cs gestiona la pantalla de inicio, implementada como un World Space Canvas en RV, que presenta un campo para el nombre del jugador y un botón de inicio. Escucha las interacciones del usuario, como la entrada del nombre mediante un teclado virtual, y valida esta entrada.
  - Gestión Centralizada de Datos de Usuario: Al pulsar "Iniciar", LoginPanel.cs invoca a UserManager.cs, que actúa como el gestor centralizado para guardar y recuperar la información de los perfiles de usuario (incluyendo aciertos, errores por nivel y estímulo) desde una base de datos local en formato JSON. Una vez procesado el login, LoginPanel.cs oculta la interfaz de inicio y da paso al juego.
- Generación de Objetos y Efectos Audiovisuales:
  - Objetos Dinámicos: La aparición de ingredientes se realiza mediante la instancia-ción de prefabs (objetos preconfigurados en Unity). Cada prefab de ingrediente incluye su modelo 3D, un collider para la física, un componente Rigidbody para la gravedad, y el script ObjetoInteractivo.cs. Se definieron posiciones específicas de aparición (posicionSpawn) para estos objetos.
  - Efectos Visuales: Se utilizó el sistema de partículas de Unity para crear efectos como burbujas y destellos en el caldero cuando se añade un ingrediente, los cuales están pre-colocados y se activan por script.

- Efectos Sonoros: Se implementaron mediante componentes AudioSource de Unity. Se utilizaron múltiples fuentes: una para sonidos ambientales del caldero (hervor, llamas) y otras para los sonidos de correcto, o fallo, activados por la validación del terapeuta.

### **Assets Utilizados**

Para la creación del entorno visual y los elementos interactivos, se recurrió a una combinación de assets personalizados y recursos de la Unity Asset Store (ver anexo 11.1 Listado de assets). Se utilizó un skybox personalizado para representar un cielo nocturno estrellado, contribuyendo a la atmósfera mágica del juego. Diversos objetos 3D fueron descargados de la Unity Asset Store, incluyendo principalmente mobiliario (bancos, mesas, librerías), pociones, el caldero, lámparas, libros, alfombras, cestos para guardar frutas y verduras, así como los modelos de las propias frutas y verduras que funcionan como ingredientes.

Un elemento central del escenario, la torre del mago (ver Figura 6.3), fue modelada en 3D personalmente utilizando el software Blender. En cuanto al terreno, inicialmente se intentó realizar un diseño personalizado; sin embargo, tras encontrar diversos errores de compatibilidad y problemas gráficos, se optó por utilizar el sistema de terreno estándar que ofrece Unity para asegurar la estabilidad y el rendimiento del entorno.



Figura 6.3: Modelo 3D de la torre de mago utilizada en el entorno de juego

## **Framework XR y SDK de Meta Quest 3**

**Meta Quest 3 SDK (Meta XR Core v62)** Añade passthrough, estabilización de manos, foveated-rendering y perfiles de controladores específicos del dispositivo.

**Ventajas** *Quest Link* permite depurar a tiempo real; los paquetes se instalan desde el gestor de paquetes de Unity y reciben actualizaciones mensuales.

## **Modelado y optimización 3D – Blender 3.6 LTS**

- Creación de la torre <100 k vértices para cada modelo

## **Control de versiones y CI/CD – Git & GitHub**

- Flujo GitFlow con ramas develop y main para mantener versión del código mantenibles.

## **Lenguaje de programación**

Toda la lógica se implementó en C#. Se descartó Visual Scripting (Bolt) porque añade sobrecarga de CPU;

## 7. PLAN DE VALIDACIÓN Y RESULTADOS

### 7.1. Plan de Validación

Inicialmente, se llevaron a cabo pruebas preliminares con dos terapeutas expertos del INCSVC, quienes brindaron retroalimentación sobre la usabilidad, el contenido y la relevancia terapéutica de las actividades. Sus observaciones sirvieron para refinar tanto la mecánica de juego como la interfaz y la progresión de dificultad: se sustituyeron los escenarios caricaturescos por entornos más realistas que facilitan la orientación de niños con baja visión o audición; se adoptó una paleta de colores de alto contraste y gran saturación para mejorar la percepción visual de los elementos clave, como objetos interactivos, caldero, texto e interfaz; se ajustó el diseño sonoro para evitar la saturación de estímulos, manteniendo una densidad auditiva equilibrada; se estableció una curva de aprendizaje gradual con niveles claramente definidos que aumentan la complejidad de forma progresiva; y se simplificó la navegación mediante botones grandes y símbolos claros, asegurando que tanto los niños como los terapeutas puedan interactuar con el videojuego sin asistencia técnica adicional.

Este trabajo involucró sujetos humanos o animales en su investigación. La aprobación de todos los procedimientos y protocolos éticos y experimentales fue otorgada por el Comité de Ética de la Investigación del Instituto para Niños Ciegos y Sordos del Valle del Cauca (Research Ethics Committee INCS) bajo la solicitud N° INV-2020-007, el 30 de junio de 2020, y se realizó conforme a las resoluciones 8430 (1994) y 2378 (2008) del Ministerio de Salud y Protección Social de Colombia. Se realizaron pruebas con un pequeño grupo de niños usuarios finales (pacientes en terapia del habla), tal como se detalla en la sección 7.2. Durante estas sesiones supervisadas, se recopilaron datos de desempeño (p. ej., número de intentos correctos, tiempo de respuesta) y se observó el nivel de compromiso y motivación de los niños al utilizar el sistema. También se recogieron impresiones de los terapeutas sobre los cambios percibidos en la disposición de los niños hacia la terapia. Los resultados de estas

evaluaciones permitieron una medición inicial de la percepción del entorno gamificado de RV y guiaron las conclusiones del estudio.

## **7.2. Resultados de las pruebas**

Para la realización de las pruebas con los participantes menores de edad, se contó con la aprobación formal y previa del Comité de Ética en Salud del Instituto para Niños Ciegos y Sordos del Valle del Cauca (INCSVC). A continuación, se describen los resultados de dichas pruebas:

Se llevó a cabo una prueba piloto del prototipo en las instalaciones del Instituto para Niños Ciegos y Sordos del Valle del Cauca, con la participación de cuatro niños de entre 7 y 11 años. Los participantes incluyeron dos niñas con discapacidad auditiva (usuarios de audífonos), un niño sin ninguna discapacidad y una niña con discapacidad visual parcial. Cada uno de ellos interactuó con distintos niveles del videojuego según su edad y capacidad. En concreto, la niña de 8 años con baja visión completó exitosamente los niveles 1, 2 y 3 (repitiendo algunos para refuerzo); la niña de 7 años sin discapacidad superó los niveles 3 y 4; el niño de 11 años con hipoacusia neurosensorial bilateral completó los niveles 4 y 5; y la niña de 10 años con discapacidad auditiva realizó los niveles 2 y 3. En todos los casos, los niños lograron pronunciar correctamente las palabras objetivo de cada ejercicio, cumpliendo con los criterios establecidos (marcados como “lo hizo bien” en los registros del sistema). Esto sugiere que los diferentes niveles y desafíos del juego fueron adecuadamente superados por los usuarios, validando el diseño técnico y pedagógico de las actividades para las edades objetivo.

Cabe destacar que, si bien el sistema está diseñado para registrar aciertos y errores, durante estas pruebas piloto se observó que, tras una explicación inicial sobre la dinámica del ejercicio (que tomó en promedio entre 10-15 minutos por niño para la realización de 2 a 4 ejercicios), todos los niños lograron las pronunciaciones correctas requeridas para avanzar. Aunque dos participantes necesitaron intentos de práctica durante la fase explicativa inicial antes de comenzar formalmente el minijuego, una vez comprendida la tarea, el desempeño dentro del juego fue exitoso para completar los ejercicios propuestos en esta sesión.

Cualitativamente, la respuesta de los niños durante la prueba fue muy positiva. Los participantes mostraron entusiasmo y asombro por el entorno virtual: incluso antes de comenzar los ejercicios específicos de habla, exploraban con curiosidad el mundo 3D de la torre del juego, mirando a su alrededor e interactuando con los elementos visuales. Uno de los niños exclamó emocionado que “se siente increíble, parece que estoy en una torre de bruja”, destacando el alto grado de inmersión que proporciona la experiencia de realidad virtual. Además, se observó motivación entre pares: la primera niña que jugó animó a la siguiente participante diciéndole que el juego “era muy divertido” y que “tenía que jugar”. En general, los niños se mostraron muy interesados en participar y disfrutaron de la experiencia lúdica, lo que indica un nivel de compromiso y motivación superior al que suele lograrse con ejercicios tradicionales de terapia de habla.

Por razones de tiempo, los niveles más avanzados (niveles 6 y 7) no se llegaron a probar de forma interactiva con los niños durante esta sesión. Sin embargo, estos niveles finales fueron presentados a cinco terapeutas en modo demostración sin el visor (es decir, mostrando en pantalla su contenido y dinámica). Dichas etapas avanzadas del juego causaron una muy buena impresión entre las profesionales. Las terapeutas señalaron que estos niveles adicionales se veían atractivos e interesantes, y comentaron que podrían ser de gran utilidad en futuras sesiones, dado que mantenían la línea lúdica y pedagógica de los anteriores, pero con mayor desafío. La reacción positiva del personal al observar los niveles 6 y 7 sugiere que el diseño del contenido avanzado es igualmente llamativo y potencialmente eficaz, aun cuando no pudo ser probado directamente con los menores en esta ocasión.

Al finalizar la prueba piloto, se recopilaron comentarios cualitativos de las fonoaudiólogas y terapeutas del instituto que presenciaron la sesión. En general, la recepción del prototipo por parte de las profesionales fue muy positiva: manifestaron que la aplicación les pareció “interesante, inmersiva y entretenida” para los niños, enfatizando el valor añadido que la realidad virtual y la gamificación aportan a la terapia de habla. Además de este feedback general, brindaron recomendaciones importantes para optimizar futuras sesiones. En primer lugar, aconsejaron ofrecer una explicación previa a los niños sobre los controles del juego y la dinámica de cada nivel, preferiblemente mostrándolas en una pantalla convencional cómo

interactuar antes de usar el visor de RV. Esta sugerencia surgió porque notaron que una breve inducción inicial podría reducir la confusión y ayudar a los niños a sentirse más seguros al empezar a jugar. En segundo lugar, las terapeutas resaltaron la utilidad de los datos registrados por el sistema durante la sesión. Consideraron valioso que la plataforma almacene información sobre cada intento y resultado, ya que esto permite verificar objetivamente si el ejercicio de pronunciación fue realizado correctamente. Indicaron que es positivo que, en la versión actual, un terapeuta esté a cargo de validar si la pronunciación de las palabras fue correcta (lo que garantiza un control de calidad terapéutico durante la sesión). Asimismo, sugirieron que sería beneficioso exportar los datos recopilados a formatos que faciliten su análisis fuera de la aplicación, de modo que el equipo clínico pueda monitorizar el progreso de cada niño a lo largo del tiempo y analizar patrones de mejora en su habla.

### **7.3. Evaluación de Rendimiento**

Debido a la decisión de implementar un alto nivel de realismo gráfico, se optó por ejecutar el videojuego en PC y transmitirlo a las gafas Meta Quest 3S mediante Meta Link. Esta configuración permite aprovechar la potencia de procesamiento del computador para mantener la calidad visual requerida mientras se utiliza el sistema de visualización inmersiva de las gafas RV. A continuación se presentan las pruebas de rendimiento obtenidas durante la ejecución normal del videojuego.



Figura 7.1: Imagen de rendimiento del videojuego



Figura 7.2: Imagen de la correcta interacción con los objetos



Figura 7.3: Imagen de la correcta eliminación de los objetos



Figura 7.4: Imagen del correcto funcionamiento de la progresión en los niveles

## 8. DISCUSIÓN

### 8.1. Evaluación crítica del diseño y metodología del prototipo

La validación del prototipo se realizó mediante pruebas piloto con un grupo reducido de cuatro niños y la observación de terapeutas. Si bien este tamaño muestral es adecuado para obtener retroalimentación inicial valiosa en el contexto del desarrollo de un prototipo y para identificar áreas de mejora en usabilidad y diseño, es importante reconocer que no permite extraer conclusiones generalizables sobre la efectividad terapéutica del sistema. Los hallazgos presentados deben interpretarse como preliminares y específicos al contexto de esta prueba inicial.

El prototipo desarrollado integró elementos de gamificación, realidad virtual y un modelo terapéutico establecido para la rehabilitación del habla infantil. Desde el punto de vista metodológico, basar la estructura del juego en la teoría de adquisición fonética de Daniel Ling permitió una progresión lógica de la dificultad. Los niveles iniciales se enfocaron en vocalizaciones simples (fonemas o sílabas básicas) asentando una base articulatoria, mientras que los niveles avanzados introdujeron combinaciones más complejas de sonidos y palabras, siguiendo una secuencia gradual y acorde al desarrollo auditivo-verbal propuesto por Ling. Esta alineación teórica es un acierto, ya que garantiza que las actividades lúdicas mantengan un sustento terapéutico sólido y apropiado para la edad y capacidades de los niños. No obstante, podría considerarse una mayor personalización de la progresión para adaptarse a ritmos individuales de aprendizaje, dado que la literatura sugiere la importancia de ajustar los ejercicios a las necesidades específicas de cada niño.

En cuanto al diseño del juego y su interfaz gráfica, la elección de una narrativa lúdica ambientada en un mundo de fantasía (por ejemplo, un laboratorio mágico en la torre de un mago)

resultó muy efectiva para captar la atención. La estética colorida y escenarios 3D inmersivos contribuyeron a que los niños se sintieran transportados al entorno del juego, aumentando su motivación intrínseca. Un participante describió que “se siente increíble, parece que estoy de verdad en una torre de bruja”, reflejando un alto nivel de inmersión y suspensión de la realidad durante la actividad terapéutica. Este nivel de involucramiento es altamente positivo, pues un entorno atractivo puede convertir ejercicios repetitivos de pronunciación en experiencias divertidas. Diversos estudios han reportado que el uso de videojuegos en contextos terapéuticos incrementa la motivación y el compromiso de los pacientes, lo cual coincide con lo observado en nuestro prototipo. En particular, se ha documentado en niños con discapacidad auditiva que los juegos serios mejoran la atención y participación en la terapia, validando la estrategia de emplear elementos lúdicos para apoyar la mecanización del habla.

Las mecánicas de juego implementadas fueron otro punto fuerte, pero con espacio para mejoras. El prototipo incorporó retroalimentación inmediata: cuando el niño pronunciaba correctamente la palabra o sonido pedido, el sistema respondía con animaciones o efectos sonoros “mágicos” de refuerzo positivo; en caso de fallo, se indicaba de forma amable que lo intentara nuevamente. Este feedback instantáneo es fundamental para el aprendizaje, ya que refuerza las respuestas correctas en el momento y guía al usuario en caso de error. No obstante, se identificó que la precisión del reconocimiento de voz y la validación de las respuestas son aspectos críticos. En la versión actual, el criterio de éxito dependía en última instancia de la verificación por parte de un terapeuta humano, quien confirmaba si la pronunciación había sido correcta. Si bien este enfoque garantiza la calidad terapéutica (evitando que el juego refuerce pronunciaciones incorrectas), también limita la automatización del sistema. A largo plazo, sería deseable añadir algoritmos de reconocimiento de voz para población infantil (incorporando quizás modelos de machine learning entrenados con voces de niños y con pronunciaciones atípicas propias de trastornos del habla). De esta forma, el juego podría brindar retroalimentación aún más precisa y autónoma, reduciendo la carga del terapeuta y permitiendo sesiones de práctica independientes. Cualquier sistema automático, sin embargo, deberá calibrarse cuidadosamente para mantener un bajo margen de error, ya que retroalimentar equivocadamente podría frustrar al niño o reforzar patrones incorrectos.

La estructura de niveles y minijuegos diseñada presentó fortalezas claras en términos de variedad y escalabilidad. Cada nivel del prototipo estaba asociado a un contexto lúdico y objetivo específico. Este diseño modular facilitó adaptar el juego a distintas etapas: de hecho, durante las pruebas, niños de diferentes edades y condiciones pudieron interactuar con niveles distintos adecuados a sus capacidades, sin mostrar desajuste en la dificultad. Esto sugiere que la gradación de retos fue en general acertada. Como acierto adicional, el juego registró datos de desempeño (aciertos, errores, fecha de realización) en cada nivel, lo que ofrece un mecanismo valioso para monitorizar el progreso de los pacientes a lo largo del tiempo. Las terapeutas valoraron positivamente esta característica, destacando que dichos datos facilitan la verificación objetiva de los ejercicios y podrían usarse para un análisis detallado de avances y dificultades. Una posible mejora a futuro es explotar estos datos para implementar una adaptabilidad del juego: por ejemplo, ajustando dinámicamente la dificultad (repetiendo más ciertos ejercicios, cambiando el ritmo de la narración o adaptando los estímulos) según el desempeño individual de cada niño, personalizando aún más la experiencia terapéutica.

## **8.2. Percepción de los usuarios: respuesta de niños y terapeutas al enfoque gamificado**

Los resultados de las pruebas piloto reflejan una recepción muy positiva tanto por parte de los niños usuarios finales como de los terapeutas encargados de su rehabilitación. En primer lugar, los niños demostraron un alto grado de satisfacción, entusiasmo e involucramiento al interactuar con el juego. Incluso antes de iniciar formalmente los ejercicios de habla, se observó que los participantes infantiles exploraban con curiosidad el entorno virtual, mirando a su alrededor y comentando entre ellos sobre lo que veían. Esto contrasta con la actitud frecuentemente reservada o ansiosa que pueden mostrar los niños durante terapias tradicionales, indicando que el componente lúdico disminuyó barreras iniciales y generó una actitud positiva hacia la sesión. Un caso ilustrativo fue el de una niña de 8 años con discapacidad visual parcial, quien tras usar el juego exclamó sentirse “como si estuviera de

verdad en una torre de bruja”, evidenciando no solo diversión sino también inmersión total en la narrativa propuesta. Esta inmersión logró que la niña olvidara en cierto modo que estaba realizando un ejercicio terapéutico estructurado, viviéndolo más bien como una aventura, lo que es un objetivo central de la gamificación terapéutica.

Asimismo, surgieron manifestaciones espontáneas de motivación extrínseca e intrínseca entre los niños: por ejemplo, una participante animó a otra diciéndole que debía jugar porque “era muy divertido”, fomentando un ambiente de cooperación y emoción compartida. Este tipo de feedback social sugiere que el juego puede convertir la rehabilitación en una experiencia hasta cierto punto contagiosa en motivación, donde los mismos pacientes se incentivan mutuamente a participar. En términos generales, ninguno de los niños se mostró renuente a continuar con las tareas; al contrario, varios completaron los niveles asignados con éxito e incluso repitieron algunos ejercicios por iniciativa propia, deseosos de seguir interactuando con el entorno lúdico. Esta disposición a practicar repetidamente es un logro notable, dado que la mecanización del habla requiere muchas repeticiones que usualmente resultan monótonas. El prototipo consiguió que la repetición se diera en un contexto de juego, atenuando la percepción de esfuerzo y potenciando la recompensa psicológica de lograr objetivos en el juego. La literatura reporta aumentos similares en compromiso cuando se integra gamificación a la terapia, hallando mejoras en la motivación y en la continuidad de las sesiones por parte de los niños. Por ejemplo, estudios como el de [16] (2017) con el juego "SpokeIt", evidenciaron mayor motivación para practicar el habla en comparación con métodos tradicionales, y el trabajo de [15] (2023) destacó cómo la gamificación en RV para rehabilitación lingual mejora la adherencia a la terapia. Estos hallazgos coinciden con lo observado en nuestro prototipo. En nuestro caso, la combinación de niveles, narrativas y refuerzos visuales/sonoros mantuvo a los usuarios jóvenes interesados de principio a fin de cada prueba.

Desde la perspectiva de las terapeutas y especialistas, la respuesta al prototipo también fue favorable, con comentarios que avalan el potencial del enfoque. Desde la perspectiva clínica, la experiencia piloto proporcionó una validación terapéutica preliminar positiva del sistema y evidenció una aceptación general favorable por parte del personal especializado. Las fonoaudiólogas y terapeutas presentes señalaron que el prototipo complementa de manera efectiva

las estrategias tradicionales de terapia del habla, ya que consigue que los niños practiquen la pronunciación y mecanización de palabras de forma activa y correcta, al mismo tiempo que se divierten. Este equilibrio entre eficacia terapéutica y entretenimiento fue destacado como uno de los mayores aciertos del proyecto. El hecho de que las profesionales calificaran la aplicación como interesante e innovadora para sus pacientes sugiere que el sistema podría integrarse en las sesiones reales de rehabilitación del habla sin resistencia, e incluso con entusiasmo, por parte de los terapeutas. Asimismo, la funcionalidad de registro de datos incorporada en el juego (que documenta el desempeño del niño en cada ejercicio) fue bien recibida, ya que provee información objetiva que puede respaldar el trabajo del terapeuta en la evaluación de los progresos del paciente. En términos generales, la respuesta positiva del personal clínico indica que el prototipo tiene viabilidad práctica en entornos terapéuticos y aporta un valor añadido real al proceso de rehabilitación del habla.

A pesar de los logros alcanzados, se identificaron algunas limitaciones importantes en la versión actual del prototipo, las cuales ofrecen oportunidades claras de mejora. Primero, la validación de la voz es manual: el sistema no cuenta todavía con un reconocimiento automático del habla, por lo que depende de que un terapeuta (o usuario supervisor) escuche la pronunciación del niño y marque manualmente si esta fue correcta o incorrecta. Esta falta de automatización implica una carga de trabajo adicional para el terapeuta y evita que el niño reciba retroalimentación inmediata directamente del software. Segundo, existe una falta de automatización en ciertas tareas del flujo de la aplicación; donde la progresión entre niveles y la evaluación del desempeño están supeditadas a la intervención del terapeuta, en lugar de realizarse de forma autónoma por el sistema. Esto limita parcialmente la escalabilidad del uso del juego en sesiones sin supervisión constante. Tercero, el repertorio de actividades es limitado en este prototipo inicial: se dispone únicamente de los siete niveles o minijuegos desarrollados, con un conjunto fijo de palabras y ejercicios. Si bien estos cubren algunos fonemas y dificultades comunes, la variedad de contenido sigue siendo reducida en comparación con la amplia gama de ejercicios que se emplean en terapia de habla tradicional. Esta restricción de contenido podría llevar a que, tras usos repetidos, los niños se familiaricen demasiado con las mismas actividades, disminuyendo potencialmente el desafío y la novedad.

Reconocer estas limitaciones es crucial, ya que orientan las mejoras necesarias para futuras versiones del sistema, y sientan las bases para el trabajo futuro que permitirá robustecer tanto la autonomía tecnológica de la aplicación como la variedad y alcance de sus componentes terapéuticos.

## 9. CONCLUSIONES

En conclusión, el desarrollo del prototipo de aplicación de software basado en videojuegos y realidad virtual alcanzó con éxito los objetivos técnicos propuestos. Se logró implementar un entorno interactivo 3D inmersivo en el dispositivo Meta Quest 3, incorporando una temática lúdica (la exploración de una torre mágica) dividida en siete niveles de dificultad progresiva. La experiencia piloto proporcionó una validación terapéutica preliminar positiva del sistema, reconociendo que estos hallazgos se basan en una muestra pequeña (N=4) propia de una fase de prototipado, y que se requieren estudios más amplios para determinar la efectividad generalizable. Cada nivel funciona a modo de minijuego que integra ejercicios de pronunciación específicos, permitiendo así mecanizar palabras de forma guiada y entretenida. El sistema demostró un desempeño estable durante las pruebas piloto, con una correcta sincronización entre las acciones del jugador y la retroalimentación visual/auditiva del juego en tiempo real. Además, el prototipo fue capaz de registrar datos de uso (por ejemplo, fecha de intento realizado, aciertos y errores en la pronunciación), sentando las bases para un seguimiento objetivo del progreso de cada usuario. Estos logros técnicos evidencian la viabilidad de combinar la tecnología de realidad virtual con ejercicios terapéuticos del habla, creando una herramienta innovadora que cumple con los requerimientos funcionales esperados.

En este sentido, se considera que los objetivos específicos del proyecto fueron abordados satisfactoriamente. El análisis de la terapia de mecanización de palabras (OE1) se realizó mediante la revisión bibliográfica y la consulta con especialistas, fundamentando el diseño en el modelo de Daniel Ling. La investigación sobre el uso de RV y videojuegos para el diseño del prototipo (OE2) se materializó en el marco teórico y las decisiones de diseño que buscaron un entorno inmersivo y lúdico. El desarrollo e implementación del prototipo (OE3) se concretó en la aplicación "Fonetic Tower", para Meta Quest 3, con sus siete niveles funcionales y mecánicas interactivas. Finalmente, la evaluación de la efectividad del prototipo (OE4) se llevó a cabo a través de las pruebas piloto con niños y la retroalimentación de profesionales

del INCSVC, cuyos resultados cualitativos iniciales sobre motivación y usabilidad fueron positivos

Los resultados de la evaluación también permiten concluir que el prototipo tiene un alto impacto motivacional e inmersivo en los niños, cumpliendo uno de los objetivos principales del proyecto. Se observó que la ambientación de videojuego en RV incrementó notablemente la motivación de los participantes para realizar los ejercicios de articulación. Los niños experimentaron la terapia como un juego divertido más que como una obligación clínica, mostrando altos niveles de atención, entusiasmo y persistencia en las tareas. El grado de inmersión alcanzado hizo que los usuarios se sintieran “transportados” al mundo virtual de la historia, lo que a su vez fomentó una actitud positiva frente a la práctica repetitiva de palabras. En comparación con metodologías tradicionales de rehabilitación del habla (que a veces resultan monótonas o poco atractivas para los menores), esta aproximación lúdica demostró ser eficaz para aumentar la participación activa de los niños y mantener su interés durante la sesión terapéutica. En síntesis, la incorporación de elementos de juego e inmersión logró transformar ejercicios de pronunciación potencialmente rutinarios en actividades estimulantes, potenciando el compromiso del paciente infantil con su proceso de rehabilitación.

Desde la perspectiva clínica, la experiencia piloto proporcionó una validación terapéutica preliminar positiva del sistema y evidenció una aceptación general favorable por parte del personal especializado. Las fonoaudiólogas y terapeutas presentes señalaron que el prototipo complementa de manera efectiva las estrategias tradicionales de terapia del habla, ya que consigue que los niños practiquen la pronunciación y mecanización de palabras de forma activa y correcta, al mismo tiempo que se divierten. Este equilibrio entre eficacia terapéutica y entretenimiento fue destacado como uno de los mayores aciertos del proyecto. El hecho de que las profesionales calificaran la aplicación como “interesante e innovadora” para sus pacientes sugiere que el sistema podría integrarse en las sesiones reales de rehabilitación del habla sin resistencia, e incluso con entusiasmo, por parte de los terapeutas. Asimismo, la funcionalidad de registro de datos incorporada en el juego (que documenta el desempeño del niño en cada ejercicio) provee información objetiva que puede respaldar el trabajo del terapeuta en la evaluación de los progresos del paciente. En términos generales, la respuesta

positiva del personal clínico indica que el prototipo tiene viabilidad práctica en entornos terapéuticos y aporta un valor añadido real al proceso de rehabilitación del habla.

A pesar de los logros alcanzados, se identificaron algunas limitaciones en la versión actual del prototipo, las cuales deberán abordarse en trabajos futuros. La validación de la pronunciación aún depende de la supervisión humana directa, lo que implica una carga adicional para el terapeuta y limita la automatización del sistema. Asimismo, la progresión entre niveles y la evaluación del desempeño requieren intervención del terapeuta en lugar de realizarse de forma autónoma, lo cual reduce la escalabilidad del uso independiente del juego. Finalmente, el repertorio de ejercicios incorporado es limitado; aunque cubre ciertos fonemas y dificultades comunes, sería necesario ampliarlo para abarcar una gama más amplia de contenidos terapéuticos. Reconocer estas limitaciones orienta las mejoras para iteraciones futuras del sistema, sirviendo de base para proponer estrategias que fortalezcan la autonomía técnica de la aplicación y amplíen su alcance terapéutico.

## 10. TRABAJO FUTURO

A partir de las conclusiones, las limitaciones identificadas y, de manera crucial, la valiosa retroalimentación proporcionada por los terapeutas del INCSVC durante las pruebas piloto, se proponen varias líneas de trabajo futuro para extender y mejorar el prototipo desarrollado: **Ampliación de contenidos y minijuegos:** Expandir el prototipo incorporando más niveles y minijuegos dentro de la temática de la torre, con el fin de cubrir un espectro más amplio de ejercicios de rehabilitación del habla. Esto incluiría agregar nuevas palabras, frases y desafíos articulatorios en cada nivel, aumentando la variedad de actividades lúdicas para mantener el interés de los niños a largo plazo y abordar diferentes aspectos de la pronunciación (por ejemplo, niveles enfocados en distintos fonemas, dificultades silábicas o velocidad del habla). Una mayor diversidad de minijuegos permitirá adaptar la terapia a las necesidades individuales de cada paciente y evitar la monotonía tras sesiones reiteradas.

**Integración de otros métodos terapéuticos:** Explorar e incorporar en el sistema nuevas modalidades de rehabilitación del habla más allá de la mecanización de palabras aisladas. Por ejemplo, se podrían desarrollar ejercicios de entonación y modulación de la voz, prácticas de pronunciación de frases completas o juegos de interacción verbal que impliquen construir oraciones y dialogar con personajes virtuales. Asimismo, podrían añadirse actividades que refuercen aspectos complementarios de la terapia del lenguaje, como ejercicios de respiración, ritmo o discriminación auditiva, integrándolos en la narrativa del juego. Estas extensiones permitirían que la aplicación abarque un rango más amplio de habilidades comunicativas, ofreciendo una intervención más holística en la rehabilitación del habla.

**Reconocimiento automático de voz:** Implementar un sistema de reconocimiento de voz automático directamente en la aplicación para permitir la validación autónoma de la pronunciación por parte del software. Con esta mejora, el juego sería capaz de analizar en tiempo real la palabra o frase pronunciada por el niño y determinar si coincide con el objetivo esperado, sin necesidad de que un terapeuta intervenga manualmente en cada intento. Esto brindaría

al usuario una retroalimentación inmediata (por ejemplo, confirmando con animaciones o sonidos cuando la pronunciación es correcta, o indicando de forma amigable que lo intente de nuevo cuando no lo es). La incorporación de reconocimiento de voz haría posible que el sistema operara de manera más independiente, abriendo la puerta a que el niño pueda practicar también en contextos domiciliarios o con menor supervisión, al tiempo que el terapeuta podría dedicar más atención a aspectos cualitativos de la sesión. Como parte de esta línea de trabajo, sería necesario entrenar y adaptar modelos de reconocimiento de voz al habla infantil y, particularmente, a las variaciones de pronunciación propias de niños con trastornos del habla, garantizando así una detección precisa y sensible a sus necesidades.

**Evaluaciones clínicas a mayor escala:** Llevar a cabo estudios y evaluaciones con una muestra más amplia de participantes, así como con sesiones más prolongadas en el tiempo. Un siguiente paso importante será realizar pruebas piloto extendidas que involucren a un número significativo de niños de la población objetivo (por ejemplo, de diferentes edades o tipos de trastorno del habla) para obtener datos estadísticamente relevantes sobre la eficacia terapéutica del sistema. Adicionalmente, se propone evaluar el uso del juego en contextos clínicos reales a lo largo de varias semanas o meses, integrándolo como complemento regular de la terapia convencional. Esto permitiría medir el impacto a largo plazo de la aplicación en la mejora del habla de los niños, evaluando indicadores como la evolución en la pronunciación, la transferencia de las habilidades practicadas en RV a situaciones de la vida real, y el nivel de adherencia o motivación sostenida de los pacientes. Los estudios a mayor escala también proporcionarían retroalimentación valiosa para refinar el diseño del juego, identificando qué elementos funcionan mejor o qué ajustes son necesarios para diferentes grupos de usuarios. En resumen, una evaluación clínica más amplia y prolongada ayudará a validar científicamente la efectividad del prototipo y a pulir sus características antes de considerarlo para una adopción más generalizada en entornos terapéuticos.

**Accesibilidad sin gafas RV (versión teclado-ratón):** Desarrollar un modo “desktop” que permita ejecutar el juego en ordenadores convencionales, utilizando teclado y ratón (o gamepad) como dispositivos de entrada. Esto ampliará el alcance a centros que no disponen de visores RV y facilitará la práctica domiciliaria cuando el niño no pueda llevarse el casco. El

reto principal consiste en adaptar la mecánica de interacción — actualmente basada en gestos 3D — a controles bidimensionales sin perder la sensación lúdica. Se propone: a) implementar un cursor mágico controlado por el ratón para manipular objetos, b) sustituir la locomoción física por desplazamiento con teclas WASD y teletransporte con clic derecho, y c) añadir un sistema de zoom que simule la proximidad a los elementos interactivos. Un modo plano también serviría como herramienta de demostración para terapeutas y padres, reduciendo la curva de familiarización antes de pasar a la experiencia inmersiva completa.

**Pruebas de rendimiento exhaustivas:** Aunque el prototipo fue optimizado durante el desarrollo para asegurar un rendimiento fluido en Meta Quest 3, buscando cumplir con el requisito de 90 fps para evitar mareo cinético, se considera como trabajo futuro la ejecución de pruebas de rendimiento más formales y detalladas. Estas pruebas permitirían cuantificar métricas específicas como la tasa de frames por segundo promedio y mínima bajo diferentes condiciones de carga, el uso de CPU y GPU, y el consumo de memoria. Los resultados de estas pruebas ayudarían a identificar posibles cuellos de botella y a optimizar aún más el sistema, especialmente si se considera la adición de contenido más complejo o funcionalidades avanzadas.

**Exportación de datos en formatos adicionales a JSON:** Atendiendo directamente a las sugerencias de los profesionales clínicos, y para integrarse mejor con los flujos de trabajo clínicos, se plantea:

- Generar reportes CSV/Excel para análisis rápido en hojas de cálculo.
- Implementar exportación PDF con gráficos de progreso y firmas digitales, útil para incluir en historias clínicas.

## REFERENCIAS

- [1] M. Frutos, I. Bustos, B. G. Zaporain, and A. M. Zorrilla, “Computer game to learn and enhance speech problems for children with autism,” in *2011 16th International Conference on Computer Games (CGAMES)*, 2011, pp. 209–216. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/6000340>
- [2] “Evaluación de la efectividad de la rehabilitación basada en la comunidad en niños con retraso del lenguaje expresivo,” *Revista Médica Hondureña*, vol. 89, no. 2, pp. 6–12, 2021. [Online]. Available: <https://revistamedicahondurena.hn/assets/Uploads/Vol89-2-2021-6-2.pdf>
- [3] “Touchmark: Partial tactile feedback design for upper limb rehabilitation in virtual reality,” *IEEE Transactions on Rehabilitation Engineering*, vol. 11, no. 2, pp. 150–159, 2023. [Online]. Available: <https://ieeexplore.ieee.org/document/10673800>
- [4] A. A. Navarro-Newball *et al.*, “Talking to teo: Video game supported speech therapy,” *Entertainment Computing*, vol. 5, no. 4, pp. 401–412, 2014. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1875952114000408>
- [5] A. N. Khoirunnisa, Munir, L. Dewi, Rasim, N. N. Azizah, and Z. P. Alivia, “The use of augmented reality, virtual reality, and mixed reality in communication children’s with ASD: Systematic literature review,” in *Advances in Visual Informatics*. Cham, Switzerland: Springer, 2024, pp. 175–190. [Online]. Available: [https://link.springer.com/chapter/10.1007/978-981-99-7339-2\\_16](https://link.springer.com/chapter/10.1007/978-981-99-7339-2_16)
- [6] S. E. Crowe, B. Shahri, T. Piumsomboon, S. Hoermann, and A. Waller, “Modelling patient-therapist collaboration for brain injury rehabilitation in virtual reality,” in *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation*

- and Personalization*, Cagliari, Italy, 2024, pp. 437–444. [Online]. Available: <https://dl-acm-org.us1.proxy.openathens.net/doi/10.1145/3631700.3665239>
- [7] H. M. Sharp and K. Hillenbrand, “Speech and language development and disorders in children,” *Pediatric Clinics of North America*, vol. 55, no. 5, pp. 1159–1173, 2008. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0031395508001570>
- [8] Cleveland Clinic Medical Professional, “How do i know if i need speech therapy?” Cleveland Clinic, Feb. 2022. [Online]. Available: <https://my.clevelandclinic.org/health/treatments/22366-speech-therapy>
- [9] I. H. Bell, J. Nicholas, M. Alvarez-Jimenez, A. Thompson, and L. Valmaggia, “Virtual reality as a clinical tool in mental health research and practice,” *Dialogues Clin. Neurosci.*, vol. 22, no. 2, pp. 169–177, Jun. 2020. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7366939/>
- [10] “Therapeutic uses of active videogames: A systematic review,” PubMed Central (PMC). [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4688462/>
- [11] J. DeCamp, “Globalization,” in *Encyclopedia of Information Systems*, H. Bidgoli, Ed. New York, NY, USA: Elsevier, 2003, pp. 459–474. [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/B0122272404000800>
- [12] S. Aleem, L. F. Capretz, and F. Ahmed, “Game development software engineering process life cycle: a systematic review,” *Journal of Software Engineering Research and Development*, vol. 4, no. 1, p. 6, Nov. 2016. [Online]. Available: [https://www.researchgate.net/publication/348446595\\_Game\\_Development\\_Research](https://www.researchgate.net/publication/348446595_Game_Development_Research)
- [13] T. Gossen, M. Nitsche, and A. Nürnberger, “Search user interface design for children: Challenges and solutions,” 2012. [Online]. Available: [https://www.researchgate.net/publication/266886488\\_Search\\_User\\_Interface\\_Design\\_for\\_Children\\_Challenges\\_and\\_Solutions](https://www.researchgate.net/publication/266886488_Search_User_Interface_Design_for_Children_Challenges_and_Solutions)

- [14] G. Desolda, R. Lanzilotti, A. Piccinno, and V. Rossano, “A system to support children in speech therapies at home,” in *Proceedings of the 14th Biannual Conference of the Italian SIGCHI Chapter*, Bolzano, Italy, 2021. [Online]. Available: <https://dl-acm-org.us1.proxy.openathens.net/doi/10.1145/3464385.3464745>
- [15] A. Rodríguez, M. Chover, and I. Boada, “Gamification and virtual reality for tongue rehabilitation,” *IEEE Access*, vol. 11, pp. 124 975–124 984, 2023. [Online]. Available: <https://ieeexplore.ieee.org/document/10292556>
- [16] J. Duval *et al.*, “Designing towards maximum motivation and engagement in an interactive speech therapy game,” in *Proceedings of the 2017 Conference on Interaction Design and Children*, Stanford, CA, USA, 2017, pp. 589–594. [Online]. Available: <https://dl-acm-org.us1.proxy.openathens.net/doi/10.1145/3078072.3084329>
- [17] A. Hair *et al.*, “Preliminary results from a longitudinal study of a tablet-based speech therapy game,” in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, Honolulu, HI, USA, 2020, pp. 1–8. [Online]. Available: <https://dl-acm-org.us1.proxy.openathens.net/doi/10.1145/3334480.3382886>
- [18] L. Qiu and S. Abdullah, “Voice assistants for speech therapy,” in *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, Virtual, USA, 2021, pp. 211–214. [Online]. Available: <https://dl-acm-org.us1.proxy.openathens.net/doi/10.1145/3460418.3479336>

## 11. ANEXOS

Cuadro 11.1: Listado de Assets de Unity Utilizados en el Proyecto

<b>Nombre</b>	<b>Función</b>	<b>Fuente</b>
Torre de mago	Entorno principal para la interacción	Creación original usando técnicas de modelado en Blender
Glossy Bubbles	Íconos brillantes para interfaces gráficas 2D	link
Free Fire VFX - URP	Efectos visuales de fuego y explosiones para URP	link
Free Fire VFX	Efectos visuales de fuego y explosiones para HDRP	link
Match 3D Object Pack: Fruits & Vegetables	Objetos 3D de frutas y vegetales para juegos tipo Match 3D	link
Free Rug Pack	Alfombras 3D para decoración de interiores	link
Alchemy Lab Props	Props de laboratorio de alquimia para escenarios de fantasía	link
Continúa en la siguiente página		

Cuadro 11.1 – continuación de la página anterior

Nombre	Función	Fuente
Animated PBR Chest Demo	Cofre animado con materiales PBR para interiores	link
Low Poly Weapons	Armas de bajo poligonaje para juegos estilizados	link
Dark Wave Paint Table 01	Mesa de pintura con estilo oscuro para decoración	link
Unity Terrain URP Demo Scene	Escena de demostración de terreno utilizando URP	link
SteampunkUI	Interfaz de usuario con temática steampunk para juegos 2D	link

## **ANEXO 2: VIDEO DE REFERENCIA PARA CONFIGURACIÓN DE UNITY RV**

El siguiente video fue consultado como referencia para la configuración inicial del proyecto de realidad virtual en Unity con soporte para Meta Quest, utilizando OpenXR y XR Toolkit:

**Título:** Unity RV XR Development Meta Quest 2023 – Getting Started PC & Mac, Open XR, XR Toolkit

**URL:** <https://www.youtube.com/watch?v=7mAAkBIWGpk>

**Autor:** Rob Shocks

**Fecha de publicación:** 9 de Junio de 2023