

GRADO EN INGENIERÍA MULTIMEDIA



TRABAJO DE FIN DE GRADO

**SPEECHTEXTER ORIENTADO A SORDOS EN
METAVERSO Y REALIDAD AUMENTADA**

AUTOR:
KEVIN AGUDO MONTIL

TUTORÍA:
**MARCOS FERNANDEZ
MARIN**

JULIO, 2023



VNIVERSITAT
E VALÈNCIA



Escola Tècnica Superior
d'Enginyeria **ETSE-UV**

TRABAJO DE FIN DE GRADO

SPEECHTEXTER ORIENTADO A SORDOS EN METAVERSO Y REALIDAD AUMENTADA

AUTOR:

KEVIN AGUDO MONTIL

TUTORÍA:

MARCOS FERNANDEZ MARIN

TRIBUNAL

Presidente/a:

Vocal 1:

Vocal 2:

Fecha de Defensa:

Calificación:

Declaración de autoría:

Yo, Kevin Agudo Montil, declaro la autoría del Trabajo Fin de Grado titulado “Speechtexter orientado a sordos en Metaverso y Realidad Aumentada” y que el citado trabajo no infringe las leyes en vigor sobre propiedad intelectual. El material no original que figura en este trabajo ha sido atribuido a sus legítimos autores.

Valencia, 10 de julio de 2023

Fdo: Kevin Agudo Montil

Resumen:

En los últimos años, hemos sido testigos de un notable aumento en el uso de recursos multimedia en diferentes ámbitos, especialmente en el ocio y el entretenimiento. El surgimiento de tecnologías como la realidad virtual (RV) y la realidad aumentada (RA) ha permitido experiencias más inmersivas y envolventes, brindando a los usuarios la oportunidad de sumergirse en mundos virtuales y explorar nuevas formas de interacción.

Este crecimiento en el uso de recursos multimedia ha llevado consigo un importante enfoque en la inclusividad y la accesibilidad. Se ha reconocido la importancia de garantizar que todas las personas, independientemente de sus habilidades o discapacidades, puedan participar y disfrutar plenamente de estas experiencias multimedia. Sin embargo, a pesar de los avances en este campo, existe una falta notable de mecanismos de inclusión en la realidad virtual, lo que limita el acceso y la participación de ciertos grupos de personas.

El objetivo principal de este proyecto es desarrollar un paquete de Unity con la funcionalidad de Speech to Text en el entorno de la realidad virtual, simulando una interacción real, con un enfoque específico en la inclusión de personas con diversidad funcional, en particular aquellas que son sordas o tienen dificultades auditivas. El sistema permitirá convertir el habla en tiempo real en texto, brindando una herramienta poderosa que facilite la comunicación y la interacción para estas personas en entornos virtuales.

A través de la aplicación de técnicas de reconocimiento de voz y transcripción automática, mi TFG busca ofrecer las herramientas necesarias para fomentar la inclusión y su uso por parte de futuros desarrolladores, teniendo la base para la integración de personas con diversidad auditiva en metaverso. Además, se explorará la integración de esta tecnología en el entorno de la realidad aumentada para dispositivos móviles.

Espero proporcionar una herramienta útil, innovadora y accesible que fomente la inclusión y la participación activa de las personas con diversidad funcional en el ámbito de la realidad virtual.

Abstract:

In recent years, we have witnessed a remarkable increase in the use of multimedia resources in various fields, especially in leisure and entertainment. The emergence of technologies such as virtual reality (VR) and augmented reality (AR) has enabled more immersive and engaging experiences, giving users the opportunity to dive into virtual worlds and explore new forms of interaction.

This growth in the use of multimedia resources has brought with it a significant focus on inclusivity and accessibility. The importance of ensuring that all individuals, regardless of their abilities or disabilities, can participate in and fully enjoy these multimedia experiences has been recognized. However, despite advances in this field, there is a notable lack of inclusive mechanisms in virtual reality, which limits the access and participation of certain groups of people.

The main objective of this project is to develop a Unity package with Speech to Text functionality in the virtual reality environment, simulating a real interaction, with a specific focus on the inclusion of people with functional diversity, particularly those who are deaf or have hearing difficulties. The system will allow real-time speech-to-text conversion, providing a powerful tool that facilitates communication and interaction for these individuals in virtual environments.

Through the application of voice recognition and automatic transcription techniques, my TFG seeks to provide the necessary tools to promote inclusion and its use by future developers, laying the foundation for the integration of people with hearing diversity in the metaverse. Additionally, the integration of this technology into the augmented reality environment for mobile devices will be explored.

I hope to provide a useful, innovative, and accessible tool that fosters inclusion and active participation of individuals with functional diversity in the field of virtual reality.

Resum:

En els últims anys, hem estat testimonis d'un notable augment en l'ús de recursos multimèdia en diferents àmbits, especialment en l'oci i l'entreteniment. L'aparició de tecnologies com la realitat virtual (RV) i la realitat augmentada (RA) ha permès experiències més immersives i envolventes, donant als usuaris l'oportunitat d'immersir-se en mons virtuals i explorar noves formes d'interacció.

Aquest creixement en l'ús de recursos multimèdia ha portat amb si un important enfocament en la inclusió i l'accessibilitat. S'ha reconegut la importància de garantir que totes les persones, independentment de les seves habilitats o discapacitats, puguin participar i gaudir plenament d'aquestes experiències multimèdia. No obstant això, malgrat els avenços en aquest camp, hi ha una notable manca de mecanismes d'inclusió en la realitat virtual, la qual cosa limita l'accés i la participació de certs grups de persones.

L'objectiu principal d'aquest projecte és desenvolupar un paquet d'Unity amb la funcionalitat de Speech to Text en l'entorn de la realitat virtual, simulant una interacció real, amb un enfocament específic en la inclusió de persones amb diversitat funcional, particularment aquelles que són sordes o tenen dificultats auditives. El sistema permetrà convertir la parla en temps real en text, oferint una eina poderosa que facilita la comunicació i la interacció per a aquestes persones en entorns virtuals.

Mitjançant l'aplicació de tècniques de reconeixement de veu i transcripció automàtica, el meu TFG busca proporcionar les eines necessàries per fomentar la inclusió i el seu ús per part de futurs desenvolupadors, establint les bases per a la integració de persones amb diversitat auditiva en el metavers. A més, s'explorà la integració d'aquesta tecnologia en l'entorn de la realitat augmentada per a dispositius mòbils.

Espero proporcionar una eina útil, innovadora i accessible que fomenti la inclusió i la participació activa de les persones amb diversitat funcional en l'àmbit de la realitat virtual.

Agradecimientos:

En primer lugar, me gustaría agradecer a mis compañeros de clase en especial mis tres amigos más cercanos, que me han apoyado siempre, aunque tuviera una visión pesimista. En segundo lugar, agradecer a mi familia y amigos. Y por último agradecer a mi tutor y los profesores que han estado acompañándome durante estos cuatro años de carrera.

Índice general

1. Introducción.....	18
1.1. Introducción.....	18
1.1.1 Aumento del uso de XR.....	18
1.1.2 Inclusión y accesibilidad.....	18
1.2. Motivación	19
1.3. Objetivos.....	20
1.4. Organización de la memoria.....	20
2. Estado del arte.....	23
2.1. Análisis de aplicaciones similares.....	23
2.2. Tecnologías.....	26
2.2.1 Unity.....	26
2.2.2 Vuforia.....	27
3. Requisitos, especificaciones, coste, riesgos, viabilidad.....	30
3.1. Requisitos.....	30
3.2. Especificaciones.....	31
3.3. Costes.....	32
3.3.1 Coste temporal.....	32
3.3.2 Coste monetario.....	42
3.3.2.1 Costes directos.....	42
3.3.2.2 Costes indirectos.....	46
3.3.2.3 Costes totales.....	46
3.4. Riesgos.....	47
3.4.1 Detección de Riesgos.....	47
3.4.2 Amortiguación y contingencia.....	49
3.5. Viabilidad.....	50
4. Análisis	53
4.1 Realidad Virtual.....	53
4.1.1 Casos de uso.....	53
4.1.2 Diagramas de secuencia.....	57
4.1.3 Diagrama de actividad general.....	61
4.1.4 Diagrama de estado.....	62
4.2 Realidad Aumentada.....	63
4.1.1 Casos de uso.....	63
4.1.2 Diagramas de secuencia.....	65
4.1.3 Diagrama de actividad general.....	66
4.1.4 Diagrama de estado.....	67
5. Diseño	69
5.1 Interfaz.....	69
5.2 Otros elementos.....	71
6. Implementación y pruebas.....	74
6.1. Implementación.....	74
6.2. Pruebas funcionales.....	80
6.3. Pruebas de usabilidad.....	86
7. Conclusiones.....	89
7.1. Revisión de costes.....	89
7.2. Conclusiones.....	91
7.3. Trabajo futuro.....	92
Bibliografía.....	94

Índice de ilustraciones

Ilustración 1: Visualización de texto de las Xrai Glass	24
Ilustración 2: Subtítulos de Vacation Simulator al apartar la mirada	25
Ilustración 3: Subtítulos de Vacation Simulator.....	25
Ilustración 4: Escena de la película “Ironman” (escritorio holográfico)	70
Ilustración 5: Mockup de la interfaz principal (VR).....	70
Ilustración 6: Mockup de la interfaz (AR)	71
Ilustración 7: Marca de hablante.....	72
Ilustración 8: Elementos de la escena (VR)	75
Ilustración 9: Elementos de la escena (AR)	76
Ilustración 10: Eventos del botón	76
Ilustración 11: Eventos de la slider.....	77
Ilustración 12: Sistema encendido sin hablante en el campo de visión	78
Ilustración 13: Sistema con el hablante en el campo de visión	78
Ilustración 14: Botón de cambiar hablante.....	79
Ilustración 15:Script de "No Mirando"	79
Ilustración 16: Sistema encendido sin jugadores en la zona de actuación.....	80
Ilustración 17: Sistema apagado (VR)	81
Ilustración 18: Sistema encendido con el hablante cambiado	81
Ilustración 19: Sistema encendido hablante marcado.....	82
Ilustración 20: Sistema transcribiendo voz de supuesto jugador.....	82
Ilustración 21: Sistema al quedarse sin jugadores en la zona de actuación	83
Ilustración 22: Sistema transcribiendo una frase muy extensa.....	84
Ilustración 23: Sistema encendido (AR)	84
Ilustración 24: Sistema apagado (AR)	85
Ilustración 25: Sistema transcribiendo voz de hablante (AR)	85

Índice de tablas

Tabla 1: Tareas del proyecto	33
Tabla 2: Coste temporal del proyecto (estimación de Kevin Agudo).....	35
Tabla 3: Coste temporal del proyecto (estimación Pablo Mocholí)	36
Tabla 4: Coste temporal del proyecto (estimación Diego Ruíz)	37
Tabla 5: Coste temporal esperado final (calculado teniendo en cuenta 8h de trabajo diarios).....	39
Tabla 6: Calculo del sueldo base por tarea.....	43
Tabla 7: Gasto total por trabajador (teniendo en cuenta un 30% de gastos).....	44
Tabla 8: Precio del software	44
Tabla 9: Amortización del coste de hardware	45
Tabla 10: Análisis del riesgo en la fase de inicio del proyecto	48
Tabla 11: Análisis del riesgo en la fase de planificación	48
Tabla 12: Análisis de riesgo en la fase de ejecución	48
Tabla 13: Análisis del riesgo en la fase de cierre	48
Tabla 14: Caso de uso de "Encender sistema" (VR).....	53
Tabla 15: Flujo de eventos principal de "Encender sistema" (VR)	54
Tabla 16: Caso de uso de "Apagar sistema" (VR)	54
Tabla 17: Flujo de eventos principal de "Apagar sistema" (VR)	54
Tabla 18: Caso de uso de "Cambiar radio" (VR)	54
Tabla 19: Flujo principal de "Cambiar radio" (VR)	55
Tabla 20: Caso de uso de "Entrar en la zona de actuación" (VR)	55
Tabla 21: Flujo principal de "Entrar en la zona de actuación" (VR)	55
Tabla 22: Caso de uso de "Salir de la zona de actuación" (VR)	56
Tabla 23: Flujo principal de "Salir de la zona de actuación" (VR)	56
Tabla 24: Caso de uso de "Hablar" (VR)	56
Tabla 25: Flujo principal de "Hablar" (VR)	57
Tabla 26: Caso de uso de "Encender sistema" (AR).....	63
Tabla 27: Flujo principal de "Encender sistema" (AR)	64
Tabla 28: Caso de uso de "Apagar sistema" (AR).....	64
Tabla 29: Flujo principal de "Apagar sistema" (AR).....	64
Tabla 30: Caso de uso de "Hablar"	64
Tabla 31: Flujo principal de "Hablar" (AR).....	65
Tabla 32: Diagrama de estados (AR).....	67
Tabla 33: Pruebas de usabilidad (VR)	87
Tabla 34: Pruebas de usabilidad (AR)	87
Tabla 35: Tiempo real de las tareas	90

Índice de diagramas

Diagrama 1: Diagrama de Gantt del proyecto (Microsoft Projects)	41
Diagrama 2: Diagrama de casos de uso del proyecto (VR)	53
Diagrama 3: Diagrama de secuencia de "Encender sistema" (VR)	57
Diagrama 4: Diagrama de secuencia de "Apagar sistema" (VR)	58
Diagrama 5: Diagrama de secuencia de "Cambiar radio" (VR)	58
Diagrama 6: Diagrama de secuencia de "Entrar en la zona de acción" (VR)	59
Diagrama 7: Diagrama de secuencia de "Salir de la zona de acción" (VR)	59
Diagrama 8: Diagrama de secuencia de "Hablar" (VR)	60
Diagrama 9: Diagrama de actividad de la aplicación (VR)	61
Diagrama 10: Diagrama de estados del sistema (VR)	62
Diagrama 11: Diagrama de casos de uso (AR)	63
Diagrama 12: Diagrama de secuencia de "Encender sistema" (AR)	65
Diagrama 13: Diagrama de secuencia de "Apagar sistema" (AR)	65
Diagrama 14: Diagrama de secuencia de "Hablar" (AR)	66
Diagrama 15: Diagrama de actividad general (AR)	66

Capítulo 1

Introducción

1.1. Introducción

1.1.1 Aumento del uso de XR

1.1.2 Inclusión y Accesibilidad

1.2. Motivación

1.3. Objetivos

1.4. Organización de la memoria

1. Introducción

1.1 Introducción

1.1.1 *Aumento del uso de XR*

El aumento del uso de la realidad virtual y la realidad aumentada en la actualidad se debe a diversos factores que han contribuido a su popularidad y adopción creciente en diferentes ámbitos.

La realidad virtual y la realidad aumentada proporcionan experiencias altamente inmersivas que permiten a los usuarios sumergirse en entornos virtuales o combinar elementos digitales con el mundo real. Esta capacidad de ofrecer experiencias envolventes y realistas ha despertado un gran interés en los usuarios y ha generado un aumento en su uso. Los avances tecnológicos en áreas como la potencia de procesamiento, la resolución de pantallas, la detección de movimiento y los gráficos 3D han impulsado el desarrollo de dispositivos de realidad virtual y aumentada más accesibles y efectivos. Estos avances han permitido una mayor calidad visual y una mejor interacción con los entornos virtuales.

Estas tecnologías se han aplicado con éxito en una amplia gama de sectores, incluyendo la industria del entretenimiento, la educación, la medicina, el diseño arquitectónico, la simulación de entrenamiento y la industria de los videojuegos. Estos sectores han experimentado un aumento en la demanda de soluciones basadas en realidad virtual y aumentada debido a los beneficios que ofrecen, como la mejora de la experiencia del usuario, el aprendizaje interactivo y la capacidad de visualizar conceptos complejos.

Con el tiempo, los dispositivos de realidad virtual y aumentada han ido volviéndose más accesibles en términos de costo y disponibilidad. A medida que los precios han disminuido y la oferta de dispositivos ha aumentado, más personas tienen la oportunidad de experimentar la realidad virtual y aumentada en sus propios hogares o en entornos comerciales. Esto ha contribuido a un mayor uso y popularidad de estas tecnologías. Los desarrolladores y diseñadores están explorando constantemente nuevas formas de utilizar estas tecnologías para crear experiencias únicas y emocionantes. A medida que se descubren nuevas aplicaciones y usos, se amplía el alcance y la demanda de la realidad virtual y aumentada.

Por lo tanto, se puede afirmar que el aumento del uso de la realidad virtual y la realidad aumentada se debe a su capacidad para ofrecer experiencias inmersivas, los avances tecnológicos que han mejorado su rendimiento, las aplicaciones prácticas en diversos sectores, la accesibilidad creciente y el impulso constante de la innovación y la creatividad. Estos factores han contribuido a su popularidad y han establecido un escenario prometedor para el futuro desarrollo y adopción de estas tecnologías.

1.1.2 *Inclusión y Accesibilidad*

Promover la inclusión y la accesibilidad significa proporcionar igualdad de oportunidades para que todas las personas puedan participar plenamente en la sociedad. Esto implica eliminar las barreras físicas, cognitivas y sociales que puedan dificultar disfrute de los derechos y servicios disponibles.

La inclusión y la accesibilidad son derechos humanos fundamentales reconocidos internacionalmente. Según la Declaración Universal de Derechos Humanos de las Naciones Unidas, todas las personas tienen derecho a participar en la vida cultural, social y política de su comunidad, así como a acceder a los servicios básicos y a la información sin discriminación. De igual forma este derecho a la inclusión se ve protegido por la Constitución Española en el Artículo 9.

La sociedad es diversa y cada persona tiene su propio conjunto de habilidades, características y necesidades. Promover la inclusión y la accesibilidad implica respetar y valorar esta diversidad, reconociendo que todas las personas tienen algo que aportar y merecen ser incluidas en todos los aspectos de la vida. La inclusión y la accesibilidad generan beneficios tanto a nivel individual como colectivo. A nivel individual, permite a las personas desarrollar su pleno potencial, mejorar su calidad de vida y tener una participación activa en la sociedad al igual que el resto de sus integrantes. A nivel colectivo, fomenta la cohesión social, la diversidad de ideas y perspectivas, y contribuye al desarrollo sostenible de la sociedad en su conjunto.

Es por esto por lo que creo que, con los avances tecnológicos, se han abierto nuevas oportunidades para aumentar la inclusión y la accesibilidad en diferentes ámbitos. Las tecnologías de asistencia, la accesibilidad web, las aplicaciones móviles y otras soluciones tecnológicas pueden ayudar a superar barreras y permitir una participación plena e igualitaria de las personas con discapacidades o necesidades específicas.

1.2 Motivación

La motivación que me impulsa a llevar a cabo este Trabajo de Fin de Grado (TFG) se fundamenta en mi firme convicción como activista y defensor de los derechos de las minorías. Como miembro de una minoría y consciente de las injusticias que a menudo enfrentan, siempre he buscado movilizarme cuando percibo situaciones de desigualdad o la falta de mejoras debido a inercias culturales arraigadas. La diversidad es también un tema que me afecta directamente, ya que, aunque no por la sordera (que es el enfoque de este proyecto), tengo la fortuna de tener amigos y familiares que conviven con diversas discapacidades funcionales, como parálisis, trastorno por déficit de atención e hiperactividad (TDAH) o autismo.

Esta experiencia personal me ha llevado a tomar conciencia de la importancia de la inclusión y la accesibilidad en nuestra sociedad. He sido testigo de cómo las barreras y los estigmas pueden limitar la participación plena y la calidad de vida de las personas con diversidad funcional. Este proyecto de TFG se presenta como una oportunidad valiosa para contribuir a la superación de estas barreras, centrándome específicamente en la comunidad de personas con dificultades auditivas.

Al abordar el desarrollo de un sistema de Speech to Text en el entorno de la realidad virtual, aspiro a mejorar la inclusión de las personas sordas o con dificultades auditivas o por lo menos mostrar mis ideas al respecto y proporcionar una herramienta que les permita comunicarse e interactuar de manera efectiva en entornos virtuales y promover el avance tecnológico en una dirección más igualitaria e inclusiva, rompiendo las barreras comunicativas y promoviendo la inclusión social.

1.3 Objetivos

El objetivo de mi proyecto es conseguir una herramienta fácil de usar y eficaz que se pueda integrar en cualquier juego o sistema de realidad virtual:

- Facilidad de uso: El objetivo principal es desarrollar un sistema de Speech to Text (STT) en la realidad virtual que sea intuitivo y fácil de usar. Se buscará una interfaz amigable que permita a las personas con diversidad auditiva acceder y utilizar el sistema de manera sencilla, sin requerir conocimientos técnicos especializados.
- Eficiencia en la ayuda a personas con diversidad auditiva: El objetivo es garantizar que el sistema de STT cumpla su propósito de manera efectiva, proporcionando transcripciones precisas y en tiempo real. Se utilizarán técnicas avanzadas de reconocimiento de voz y transcripción automática para lograr una alta precisión en la conversión del habla en texto, mejorando así la comunicación y la interacción de las personas con diversidad auditiva en entornos virtuales.

Para lograr esto es necesario completar unos objetivos secundarios como crear una interfaz agradable a la vista, diseñar un modelo que muestre el radio de efecto del *speechtexter* e implementar el código que permita variar su tamaño en función de las necesidades del usuario.

En el caso de la realidad aumentada los objetivos son los mismos, conseguir una aplicación fácil de usar, pero eficiente.

1.4 Organización de la Memoria

La memoria está organizada en capítulos (siendo este el Capítulo 1) enumerados a continuación:

- Capítulo 2: Revisión de otras aplicaciones y proyectos que abordan la inclusión y la accesibilidad en la realidad virtual y aumentada y una descripción de las tecnologías y herramientas empleadas, como Unity (y sus herramientas para la implementación de aplicaciones de realidad virtual) y Vuforia, y su relevancia en el contexto del proyecto.
- Capítulo 3: Identificación y descripción detallada de los requisitos funcionales y no funcionales del sistema, la definición de las especificaciones técnicas y funcionales del sistema, incluyendo los aspectos relacionados con la usabilidad y la accesibilidad, incluirá también el análisis de los costes asociados al desarrollo e implementación del proyecto la identificación y evaluación de los posibles riesgos y obstáculos que podrían surgir durante el desarrollo del proyecto así como una valoración de la viabilidad técnica, económica y temporal del proyecto.
- Capítulo 4: Análisis de los requisitos funcionales y no funcionales (desglose detallado de los requisitos y su importancia en el diseño y desarrollo del sistema).

- Capítulo 5: Diseño detallado de la realidad virtual y aumentada (descripción de las decisiones de diseño y la implementación del entorno y las interacciones en la realidad virtual y cómo se utiliza la extensión Vuforia para la realidad aumentada y cómo se integra con el sistema de Speech to Text).
- Capítulo 6: Implementación del sistema en Unity así como la descripción del proceso de implementación del sistema, incluyendo la configuración de Unity y la integración de los componentes necesarios. Se verá también las pruebas funcionales en realidad virtual y las pruebas funcionales en realidad aumentada.
- Capítulo 7: Presentación de los resultados y logros alcanzados durante el desarrollo del proyecto, se realiza un análisis de las conclusiones obtenidas en relación con los objetivos establecidos inicialmente, así como de las limitaciones encontradas durante el proyecto y las lecciones aprendidas para futuros desarrollos y sugerencias para mejorar y ampliar el sistema en futuras iteraciones.

Capítulo 2

Estado del arte

2.1. Análisis de aplicaciones similares

2.2. Tecnologías

2.2.1 Unity

2.2.2 Vuforia

2. Estado del arte

2.1 Análisis de aplicaciones similares

En la actualidad, existen diversas aplicaciones y tecnologías que se centran en el reconocimiento de voz y transcripción en tiempo real. Estas aplicaciones están diseñadas para mejorar la accesibilidad y la comunicación para personas sordas o con dificultades auditivas. A continuación, se presentará algunas de las aplicaciones más relevantes en este campo:

Google Live Transcribe

Esta aplicación para dispositivos móviles utiliza el reconocimiento de voz en tiempo real para transcribir lo que se dice en una conversación. Proporciona una transcripción en tiempo real en la pantalla del dispositivo y es capaz de detectar y etiquetar diferentes hablantes.

Ava

Ava es una aplicación de transcripción en tiempo real que permite a las personas sordas seguir conversaciones en situaciones grupales. Permite a los usuarios ver las transcripciones en tiempo real en sus dispositivos móviles y también muestra quién está hablando. Permite también escribir un texto para que el dispositivo reproduzca la frase de manera que las personas que no son capaces de expresarse verbalmente también se puedan comunicar.

Microsoft Translator

Microsoft Translator ofrece capacidades de traducción y transcripción en tiempo real. Permite a los usuarios grabar y transcribir conversaciones, y también ofrece la opción de traducir la transcripción a diferentes idiomas. Aunque no está centrada en la realidad virtual o la realidad aumentada, proporciona una base tecnológica sólida para la implementación de transcripciones en tiempo real.

Estas aplicaciones están pensadas para smartphone y/o ordenadores, son aplicaciones muy útiles, pero no logran atravesar la barrera que puede ocasionar la realidad virtual en las personas sin audición o con problemas de audición. Y tampoco ofrecen una solución mediante la realidad aumentada que permitiría al usuario relacionarse con más normalidad al no tener que desviar la mirada de los hablantes. A continuación, voy a hablar de una aplicación más novedosa:

Xrai Glass

Un ejemplo a destacar es Xrai Glass, una aplicación diseñada para smartphones que permite su funcionamiento en realidad aumentada mediante el uso de tres tipos de gafas compatibles. Aunque es importante destacar que estas gafas suelen presentar un costo elevado para los usuarios. Además, la aplicación ofrece diferentes planes de suscripción mensual que implican un desembolso adicional. Sin embargo, a pesar de estas

consideraciones, Xrai Glass se ha reconocido como una solución muy beneficiosa en el día a día de las personas con diversidad funcional auditiva.

Cabe mencionar que Xrai Glass, además de proporcionar una experiencia inmersiva, posee características específicas orientadas a mejorar la comunicación y la interacción de las personas con diversidad auditiva. Estas características pueden incluir opciones de transcripción de voz en tiempo real, visualización de subtítulos y capacidades de reconocimiento de lenguaje de señas, entre otras funcionalidades. Se puede apreciar en la foto el formato de texto que aparece en las gafas, simple y poco vistoso, aunque no es posible entender bien con las imágenes disponible por la compañía como se visualiza el texto desde las gafas en realidad.



Ilustración 1: Visualización de texto de las Xrai Glass

Vacation Simulator

El juego Vacation Simulator es un excelente ejemplo de adaptación en términos de accesibilidad para jugadores con dificultades auditivas. Aunque utiliza subtítulos predefinidos en lugar de un sistema de transcripción en tiempo real, implementa características innovadoras que mejoran la experiencia de juego y la comprensión de la narrativa.

Una de las características de la adaptación en Vacation Simulator es la ubicación de los subtítulos. En lugar de presentarlos en la parte inferior de la pantalla, los subtítulos se superponen a la imagen del personaje que está hablando. Sin embargo, lo más destacado es que, al desviar la mirada del personaje, los subtítulos siguen la mirada del jugador para mantenerlos siempre a la vista. Esto asegura que los jugadores no pierdan información importante mientras exploran el entorno virtual.

Además, el juego utiliza la transparencia en los subtítulos para evitar obstrucciones visuales innecesarias. Esto permite que los jugadores sigan disfrutando de una experiencia inmersiva mientras leen los subtítulos sin distracciones. Al combinar los subtítulos con la imagen del personaje y el seguimiento del movimiento, Vacation Simulator logra proporcionar una experiencia más envolvente y comprensible para los jugadores con dificultades auditivas.

Aunque los subtítulos en *Vacation Simulator* no son generados en tiempo real, me parece una referencia muy importante gracias a que su diseño permite una experiencia más accesible y cómoda para los jugadores con dificultades auditivas. La combinación de subtítulos dinámicos y la capacidad de seguir la mirada del jugador aseguran que la información relevante siempre esté disponible sin interrupciones visuales.



Ilustración 2: Subtítulos de Vacation Simulator al apartar la mirada



Ilustración 3: Subtítulos de Vacation Simulator

En términos de realidad virtual y realidad aumentada, hay aplicaciones y tecnologías emergentes que se están desarrollando para mejorar la comunicación y la accesibilidad. Sin embargo, la mayoría de ellas se centran en la visualización y la interacción en entornos virtuales, y no tanto en el reconocimiento de voz y la transcripción en tiempo real orientado a la comunicación para sordos. En es el caso de IIElevenLabs, ha desarrollado una voz para metaverso centrándose en las personas con diversidad mudas, pero no ha desarrollado un sistema que funcione de manera inversa.

Dado este panorama, este Speech to Text en realidad virtual y realidad aumentada aborda una necesidad importante al ofrecer una solución específicamente diseñada para personas sordas o con dificultades auditivas. Al combinar la tecnología de realidad virtual y realidad aumentada con el reconocimiento de voz en tiempo real, este proyecto tiene el potencial de mejorar significativamente la comunicación y la experiencia de las personas

con diversidad funcional en entornos virtuales ya que el resultado final de este será un paquete de Unity fácil de integrar en otros proyectos, de manera que otros desarrolladores puedan utilizarlo, expandir y mejorar mi trabajo.

2.2 Tecnologías

Después de realizar una exhaustiva investigación sobre las tecnologías disponibles en el mercado, se ha constatado una notable ausencia de referencias relacionadas con la aplicación de la tecnología *speechtexter* en el contexto de la realidad virtual. Por consiguiente, se ha decidido utilizar una plataforma que se considere la más idónea para trabajar con entornos XR. En este sentido, Unity es la opción más adecuada debido a su amplia compatibilidad con diversos dispositivos, librerías y recursos adicionales.

Unity ha destacado por su capacidad de integración con dispositivos de realidad virtual y aumentada, lo cual es fundamental para lograr una experiencia inmersiva y efectiva. Al aprovechar las funcionalidades de Unity, espero una mayor compatibilidad con los dispositivos XR disponibles en el mercado, así como con las librerías y assets (elementos predefinidos) proporcionados por la comunidad de desarrolladores.

2.2.1 Unity

Unity es una plataforma de desarrollo de videojuegos y aplicaciones interactivas ampliamente reconocida en la industria. Su popularidad se debe a su versatilidad, facilidad de uso y capacidades multiplataforma, lo que la convierte en una herramienta preferida tanto por desarrolladores principiantes como por profesionales experimentados.

Unity permite crear aplicaciones en 2D y 3D, ofreciendo un entorno de desarrollo integrado (IDE) que facilita la creación de escenas, modelos, animaciones, efectos visuales y lógica de juego. Utiliza el lenguaje de programación C# para desarrollar scripts y comportamientos interactivos.

La fama de Unity se ha consolidado principalmente debido a sus ventajas clave:

- Unity es compatible con una amplia variedad de plataformas, como PC, Mac, consolas de videojuegos, dispositivos móviles y realidad virtual/aumentada. Esto permite a los desarrolladores crear una vez y publicar en múltiples plataformas, ahorrando tiempo y esfuerzo.
- Cuenta con una comunidad activa y una amplia biblioteca de recursos, tutoriales y assets predefinidos que facilitan el desarrollo. Además, ofrece una Asset Store donde los desarrolladores pueden encontrar elementos gráficos, scripts y herramientas adicionales para enriquecer sus proyectos, este punto es de especial importancia ya que se ofrece de manera gratuita scripts destinados a la realidad virtual, aumentada e incluso speech to text.
- Unity ofrece un potente motor gráfico que permite crear escenas y efectos visuales impresionantes. También proporciona una amplia gama de simulaciones físicas para crear comportamientos realistas en los objetos y personajes del juego.

Esta plataforma cuenta con compatibilidad con las gafas VR de Oculus, de manera muy sencilla se pueden crear proyectos ejecutables en las mismas. XR Toolkit (XR significa Extended Reality, que abarca la Realidad Virtual y la Realidad Aumentada), es una extensión oficial de Unity que proporciona herramientas y funcionalidades específicas para el desarrollo de experiencias de realidad virtual y realidad aumentada.

El XR Toolkit simplifica el desarrollo de aplicaciones XR al ofrecer una serie de scripts y componentes predefinidos que facilitan la interacción con dispositivos de realidad virtual y aumentada, como controladores, cámaras y sistemas de detección de movimiento. Además, incluye características de detección y seguimiento de objetos, soporte para renderizado estereoscópico y opciones para adaptar la interfaz de usuario a entornos inmersivos.

2.2.2 Vuforia

Vuforia es una plataforma de desarrollo de realidad aumentada (RA) líder en la industria. Proporciona herramientas y capacidades avanzadas para crear experiencias de realidad aumentada de alta calidad. Vuforia es compatible con una amplia gama de dispositivos, incluidos smartphones, tablets y gafas de realidad aumentada.

Vuforia tiene la capacidad de detectar y rastrear imágenes en el entorno del usuario en tiempo real. Esto permite vincular la imagen capturada por la cámara del dispositivo con contenido virtual, lo que resulta esencial para superponer la transcripción del habla en el entorno visual, sin mucho problema esta tecnología es capaz de generar una textura donde se muestra la grabación de la cámara del dispositivo. Vuforia está diseñado para trabajar perfectamente con el motor de desarrollo Unity. Proporciona un kit de desarrollo de software (SDK) que se integra fácilmente en Unity, permitiéndote crear aplicaciones de realidad aumentada de manera eficiente y con una gran interoperabilidad. Es compatible con la plataforma Android, lo que permite desarrollar y desplegar el proyecto de realidad aumentada en dispositivos Android de manera efectiva. Esto garantiza una amplia compatibilidad con una amplia gama de smartphones y tablets, lo que a su vez amplía el alcance y la accesibilidad de la aplicación.

En conclusión, después de realizar una investigación exhaustiva sobre las tecnologías disponibles en el mercado, se ha determinado que Unity y Vuforia son las herramientas más adecuadas para el desarrollo de las aplicaciones de realidad aumentada (AR) y realidad virtual (VR) en este proyecto.

Unity, una plataforma reconocida en la industria, destaca por su versatilidad, facilidad de uso y capacidades multiplataforma. Permite desarrollar aplicaciones en 2D y 3D, ofreciendo un entorno integrado que facilita la creación de escenas, modelos, animaciones y efectos visuales. Su compatibilidad con una amplia variedad de plataformas y su comunidad activa, que proporciona recursos y assets predefinidos, contribuyen a agilizar el proceso de desarrollo.

Además, Unity cuenta con el XR Toolkit, una extensión oficial que simplifica el desarrollo de aplicaciones XR al ofrecer scripts y componentes predefinidos para interactuar con dispositivos de realidad virtual y aumentada. También brinda características de detección y seguimiento de objetos, soporte para renderizado estereoscópico y opciones de adaptación de la interfaz de usuario a entornos inmersivos.

Por otro lado, Vuforia destaca como una plataforma líder en el desarrollo de realidad aumentada. La integración de Vuforia con Unity, a través de su kit de desarrollo de software (SDK), facilita el desarrollo eficiente de aplicaciones de realidad aumentada con una gran interoperabilidad. Además, su compatibilidad con dispositivos Android amplía el alcance y la accesibilidad de la aplicación.

En resumen, la combinación de Unity y Vuforia proporciona un conjunto de herramientas sólidas y compatibles que permitirán el desarrollo eficiente de las aplicaciones de realidad aumentada y realidad virtual en este proyecto. Estas tecnologías ofrecen funcionalidades avanzadas, como la superposición de transcripciones en tiempo real y la detección de objetos, lo que contribuirá a brindar experiencias inmersivas y accesibles para las personas con diversidad funcional auditiva.

Capítulo 3

Requisitos, especificaciones, coste, riesgos, viabilidad

3.1. Requisitos

3.2. Especificaciones

3.3. Costes

3.3.1 Coste temporal

3.3.2 Coste monetario

3.3.2.1 Costes directos

3.3.2.2 Costes indirectos

3.3.2.3 Costes Totales

3.4. Riesgos

3.4.1 Detección de riesgos

3.4.2 Mitigación y Contingencia

3.5. Viabilidad

3. Requisitos, especificaciones, coste, riesgos, viabilidad

En este capítulo, se abordarán los requisitos fundamentales necesarios para el desarrollo de las aplicaciones en el marco de este proyecto. La identificación y definición precisa de los requisitos es una etapa esencial para garantizar que las aplicaciones cumplan con las expectativas y necesidades de los usuarios, así como para orientar el proceso de desarrollo de manera efectiva.

El objetivo es proporcionar una descripción detallada de los requisitos funcionales y no funcionales que se deben tener en cuenta durante el diseño y la implementación de las aplicaciones. Estos requisitos abarcarán diversas dimensiones, como la funcionalidad, la usabilidad, la accesibilidad y el rendimiento, entre otros aspectos cruciales para garantizar una experiencia satisfactoria para los usuarios.

3.1 Requisitos

Requisitos de Realidad Aumentada (AR):

- **Requisitos No Funcionales (RNF):**

RNF 1. El sistema debe ser desarrollado en Unity.

RNF 2. El sistema debe usar la cámara de Vuforia.

RNF 3. El sistema debe ser compatible con PC y dispositivos móviles.

- **Requisitos Funcionales (RF):**

RF 1. El sistema debe ser translúcido para no opacar la imagen.

RF 2. El usuario debe poder apagar/encender el sistema en cualquier momento.

RF 2. El sistema debe mostrar la imagen de la cámara correctamente.

RF 3. El sistema debe mostrar su estado en todo momento para evitar confusiones.

Requisitos de Realidad Virtual (VR):

- **Requisitos No Funcionales (RNF):**

RNF 1. El sistema debe ser desarrollado en Unity.

RNF 2. El sistema debe ser compatible con las gafas de Realidad Virtual y sus mandos.

- **Requisitos Funcionales (RF):**

RF 1. El usuario debe poder acceder siempre al interfaz 3D.

RF 2. El interfaz debe seguir siempre al usuario.

RF 3. El usuario debe poder encender/apagar el sistema en cualquier momento.

RF 4. El sistema deberá detectar cuando se está hablando para ponerse en funcionamiento.

RF 5. El sistema speechtexter no debe estar en funcionamiento si no hay nadie en el rango de actuación.

RF 6. El sistema debe detectar la entrada de un jugador en el rango de actuación.

RF 7. El usuario debe poder cambiar el radio de escucha en cualquier momento.

RF 8. El sistema debe cambiar el tamaño del texto según su longitud.

RF 9. El usuario debe poder visualizar el rango de actuación del sistema para decidir su distancia según sus necesidades.

RF 10. El sistema debe señalar correctamente al hablante para su identificación.

RF 11. El sistema debe mostrar su estado en todo momento para evitar confusiones.

RF 12. El sistema no debe interferir con la visión del usuario.

3.2 Especificaciones

Teniendo en cuenta los requisitos anteriores:

Funcionalidades de la Aplicación de Realidad Aumentada (AR):

- Superposición translúcida: La aplicación superpondrá la transcripción de voz en tiempo real de manera translúcida sobre la imagen capturada por la cámara de Vuforia, evitando opacar la imagen de la realidad aumentada.
- Encendido/apagado del sistema: El usuario tendrá la capacidad de activar o desactivar el sistema de transcripción en cualquier momento, lo que permitirá controlar la visualización de la transcripción de voz.
- Visualización correcta de la imagen de la cámara: La aplicación asegurará que la imagen de la cámara capturada por Vuforia se muestre correctamente en la interfaz, permitiendo una integración efectiva de la realidad aumentada y la transcripción de voz.
- Indicación clara del estado: La aplicación mostrará de manera visible y comprensible el estado del sistema de transcripción en todo momento, evitando cualquier confusión por parte del usuario.
- Transcripción: La transcripción de voz debe de hacerse de manera correcta y mostrarse en el interfaz.

Funcionalidades de la Aplicación de Realidad Virtual (VR):

- Acceso constante al interfaz 3D: El usuario podrá acceder y visualizar el interfaz 3D en todo momento durante la experiencia de realidad virtual, lo que permitirá interactuar con las funcionalidades de transcripción de voz.
- Transcripción: La transcripción de voz debe de hacerse de manera correcta y mostrarse en el interfaz.
- Seguimiento del interfaz al usuario: El interfaz 3D se mantendrá constantemente orientado hacia el usuario, asegurando que la transcripción de voz sea legible y

esté siempre visible, independientemente de la dirección en la que el usuario esté mirando.

- Encendido/apagado del sistema: El usuario tendrá la posibilidad de activar o desactivar el sistema de transcripción en cualquier momento durante la experiencia de realidad virtual, permitiendo un control total sobre la visualización de la transcripción.
- Activación basada en el habla y rango de actuación: El sistema de transcripción se activará cuando se detecte que el usuario está hablando y se encuentre dentro del rango de actuación establecido. Si no hay nadie en el rango de actuación, el sistema no estará en funcionamiento.
- Cambio del radio de escucha: El usuario tendrá la capacidad de ajustar el radio de escucha del sistema de transcripción según sus necesidades y preferencias, lo que le permitirá adaptar el alcance de captura de voz.
- Ajuste automático del tamaño del texto: La aplicación ajustará automáticamente el tamaño del texto de la transcripción según la longitud del contenido, garantizando una legibilidad óptima en la interfaz de realidad virtual.
- Visualización del rango de actuación: El usuario podrá visualizar un indicador o representación gráfica del rango de actuación del sistema de transcripción en la realidad virtual, lo que le permitirá seleccionar la distancia adecuada según sus necesidades y preferencias.
- Señalización clara del hablante: La aplicación proporcionará una señalización clara o indicador visual para identificar correctamente al hablante en el entorno de realidad virtual, facilitando la interacción y la comunicación.
- Indicación visible del estado: La aplicación mostrará de manera constante y comprensible el estado del sistema de transcripción en la interfaz de realidad virtual, evitando cualquier confusión o malentendido por parte del usuario.

3.3 Costes

Al abordar el análisis de costes, es fundamental contar con una planificación exhaustiva de las tareas, desglosada y detallada, que sirva como base de trabajo. En el contexto de este proyecto, utilizaremos el propio índice del trabajo como punto de partida para realizar un análisis de costes tanto en términos temporales como monetarios.

3.3.1 Costes temporales

Tabla de tareas del proyecto:

ID	TAREAS
1	Introducción
1.1	Introducción
1.1.1	<i>Aumento del uso de XR</i>
1.1.2	<i>Inclusión y accesibilidad</i>
1.2	Motivación
1.3	Objetivos
1.4	Organización de la memoria
2	Estado del arte
2.1	Análisis de aplicaciones similares
2.2	Tecnologías
2.2.1	<i>Unity</i>
2.2.2	<i>Vuforia</i>
3	Plan de trabajo
3.1	Requisitos
3.2	Especificaciones
3.3	Costes
3.4	Riesgos
3.5	Viabilidad
4	Análisis
5	Diseño
6	Implementación y pruebas
6.1	Implementación
6.2	Pruebas funcionales
6.3	Pruebas de rendimiento
6.4	Pruebas de usabilidad
7	Conclusiones
7.1	Revisión de costes
7.2	Conclusiones
7.3	Trabajo futuro

Tabla 1: Tareas del proyecto

Para realizar la estimación temporal del proyecto, se emplea una técnica de estimación mixta por horas que combina dos métodos distintos:

- Asesoramiento de expertos: Se solicita la opinión de uno o varios expertos en el campo del desarrollo de software. Se les proporciona una explicación detallada del proyecto junto con la planificación de las actividades, y se les pide que estimen de manera fiable los tiempos aproximados necesarios para cada actividad.

- Estimación por tres valores: Se utiliza un enfoque estadístico basado en tres tipos de tiempos que se combinan para obtener una estimación más precisa. Estos tiempos son los siguientes:

- Tiempo más optimista (t_o): Representa la duración mínima que podría requerir una actividad bajo las mejores circunstancias.
- Tiempo más probable (t_m): Indica la duración más probable de una actividad basada en un análisis realista de los recursos y las condiciones.
- Tiempo más pesimista (t_p): Representa la duración máxima que podría requerir una actividad en situaciones adversas o con dificultades inesperadas.

El tiempo estimado (t_e) se calcula mediante la siguiente fórmula:

$$t_e = \frac{t_o + 4t_m + t_p}{6}$$

Esta fórmula pondera el tiempo más probable más significativamente, mientras que los tiempos más optimista y pesimista se consideran con menor peso. La combinación de estos dos métodos de estimación, basados en la experiencia de expertos y en un análisis estadístico de los tiempos, proporciona una aproximación más precisa de la duración total del proyecto y de las actividades individuales.

He consultado a dos expertos del IRTIC (Instituto de Robótica y Tecnologías de la Información y las Comunicaciones), un reconocido centro de investigación universitario afiliado a la Universidad de Valencia. Los expertos que proporcionaron estimaciones son Pablo Mocholí y Diego Ruiz. Se dará un mayor peso a las estimaciones de Pablo Mocholí debido a su mayor experiencia en el campo de la realidad virtual y realidad aumentada.

Tabla de tareas y coste temporal en horas (Kevin Agudo):

ID	TAREAS	t_o	t_m	t_p	t_e
1	Introducción	16	22	40	24
1.1	Introducción	4	8	16	9
1.1.1	<i>Aumento del uso de XR</i>	2	4	8	4
1.1.2	<i>Inclusión y accesibilidad</i>	2	4	8	4
1.2	Motivación	2	4	8	4
1.3	Objetivos	4	7	12	7
1.4	Organización de la memoria	2	3	4	3
2	Estado del arte	12	23	42	24
2.1	Análisis de aplicaciones similares	6	11	20	12
2.2	Tecnologías	6	12	22	13
2.2.1	<i>Unity</i>	3	8	17	9
2.2.2	<i>Vuforia</i>	3	4	5	4
3	Plan de trabajo	44	70	109	72
3.1	Requisitos	12	18	25	18
3.2	Especificaciones	7	12	20	13
3.3	Costes	7	12	20	13
3.4	Riesgos	9	14	22	15
3.5	Viabilidad	9	14	22	15
4	Análisis	20	32	45	32
5	Diseño	30	40	55	41
6	Implementación y pruebas	388	555	888	583
6.1	Implementación	350	500	800	525
6.2	Pruebas funcionales	20	25	35	26
6.3	Pruebas de rendimiento	10	18	28	18
6.4	Pruebas de usabilidad	8	12	25	14
7	Conclusiones	25	37	80	42
7.1	Revisión de costes	8	12	25	14
7.2	Conclusiones	8	12	25	14
7.3	Trabajo futuro	9	13	30	15
	TOTAL	535	779	1259	818

Tabla 2: Coste temporal del proyecto (estimación de Kevin Agudo)

Tabla de tareas y coste temporal en horas (Pablo Mocholí):

ID	TAREAS	t_o	t_m	t_p	t_e
1	Introducción	15	24	38	25
1.1	Introducción	5	8	16	9
1.1.1	<i>Aumento del uso de XR</i>	3	4	8	5
1.1.2	<i>Inclusión y accesibilidad</i>	2	4	8	4
1.2	Motivación	2	4	6	4
1.3	Objetivos	6	9	12	9
1.4	Organización de la memoria	2	3	4	3
2	Estado del arte	18	27	38	27
2.1	Análisis de aplicaciones similares	7	12	15	12
2.2	Tecnologías	11	15	23	16
2.2.1	<i>Unity</i>	7	10	17	11
2.2.2	<i>Vuforia</i>	4	5	6	5
3	Plan de trabajo	44	64	98	66
3.1	Requisitos	13	20	25	20
3.2	Especificaciones	5	10	17	10
3.3	Costes	10	12	19	13
3.4	Riesgos	9	12	20	13
3.5	Viabilidad	7	10	17	11
4	Análisis	19	30	50	32
5	Diseño	35	40	60	43
6	Implementación y pruebas	356	570	902	590
6.1	Implementación	310	510	820	528
6.2	Pruebas funcionales	24	28	35	29
6.3	Pruebas de rendimiento	12	18	22	18
6.4	Pruebas de usabilidad	10	14	25	15
7	Conclusiones	32	40	65	43
7.1	Revisión de costes	10	12	20	13
7.2	Conclusiones	12	15	25	16
7.3	Trabajo futuro	10	13	20	14
	TOTAL	519	795	1251	825

Tabla 3: Coste temporal del proyecto (estimación Pablo Mocholí)

Tabla de tareas y coste temporal en horas (Diego Ruíz):

ID	TAREAS	t_o	t_m	t_p	t_e
1	Introducción	22	26	42	28
1.1	Introducción	8	8	16	9
1.1.1	<i>Aumento del uso de XR</i>	4	6	8	6
1.1.2	<i>Inclusión y accesibilidad</i>	4	5	7	5
1.2	Motivación	4	5	7	5
1.3	Objetivos	7	9	14	10
1.4	Organización de la memoria	3	4	5	4
2	Estado del arte	23	30	40	31
2.1	Análisis de aplicaciones similares	9	12	16	12
2.2	Tecnologías	14	18	24	18
2.2.1	<i>Unity</i>	9	12	17	12
2.2.2	<i>Vuforia</i>	5	6	7	6
3	Plan de trabajo	55	69	99	72
3.1	Requisitos	17	21	25	21
3.2	Especificaciones	7	12	17	12
3.3	Costes	12	14	18	14
3.4	Riesgos	10	12	22	13
3.5	Viabilidad	9	10	17	11
4	Análisis	24	32	54	34
5	Diseño	39	40	58	43
6	Implementación y pruebas	446	564	1020	754
6.1	Implementación	400	500	900	683
6.2	Pruebas funcionales	23	30	70	36
6.3	Pruebas de rendimiento	12	18	24	18
6.4	Pruebas de usabilidad	11	16	26	17
7	Conclusiones	32	50	68	50
7.1	Revisión de costes	12	16	20	16
7.2	Conclusiones	12	19	26	19
7.3	Trabajo futuro	8	15	22	15
	TOTAL	641	811	1199	1011

Tabla 4: Coste temporal del proyecto (estimación Diego Ruíz)

Al examinar detenidamente las tablas anteriores, se puede apreciar que Kevin ha realizado una estimación total de 818 horas para la duración del proyecto. Esto equivale a aproximadamente 102 días de trabajo, lo que se traduce en alrededor de 4 meses y 14 días.

Por otro lado, Pablo Mocholí ha llevado a cabo una estimación distinta, proyectando un total de 825 horas para la duración del proyecto. Esto equivale a aproximadamente 103 días de trabajo, lo que se traduce en alrededor de 4 meses y 15 días. Es interesante observar que esta estimación presenta similitudes significativas con la estimación previa realizada por Kevin. La estimación de 825 horas elaborada por Pablo Mocholí también ha considerado cuidadosamente las diferentes tareas y actividades involucradas en el proyecto. Ha tenido en cuenta la complejidad de implementar el sistema de reconocimiento de voz, realizar pruebas funcionales y de rendimiento, y desarrollar el diseño y la implementación de la interfaz de usuario.

Y, por último, la estimación de Diego Ruíz ha sido de 1011 horas de duración, esto significa que se prolongaría a 126 días de trabajo, 5 meses y 16 días, una estimación bastante más elevada que las dos anteriores.

Para poder calcular el resultado final es necesario asignar una ponderación (p_i) a cada tiempo estimado, en este caso la estimación de Diego Ruíz será menor por ser notablemente distante a las otras dos y la de Pablo Mocholí la mayor por su experiencia en este ámbito. La ponderación de Kevin Agudo será de 0,3 la de Diego Ruíz de 0.2 y la de Pablo Mocholí de 0.5, utilizando estas en la siguiente formula:

$$t_e = \frac{\sum p_i * t_e^i}{\sum p_i}$$

Resultado final teniendo en cuenta las tres estimaciones ponderadas:

ID	TAREAS	t_e horas	t_e días
1	Introducción	24,8	3,1
1.1	Introducción	8,9	1,1125
1.1.1	<i>Aumento del uso de XR</i>	4,5666667	0,570833
1.1.2	<i>Inclusión y accesibilidad</i>	4,4	0,55
1.2	Motivación	4,4	0,55
1.3	Objetivos	7,6666667	0,958333
1.4	Organización de la memoria	3,2	0,4
2	Estado del arte	25,466667	3,183333
2.1	Análisis de aplicaciones similares	11,866667	1,483333
2.2	Tecnologías	13,9	1,7375
2.2.1	<i>Unity</i>	9,5	1,1875
2.2.2	<i>Vuforia</i>	4,4	0,55
3	Plan de trabajo	72,016667	9,002083
3.1	Requisitos	18,683333	2,335417
3.2	Especificaciones	12,55	1,56875
3.3	Costes	13,016667	1,627083
3.4	Riesgos	14,416667	1,802083
3.5	Viabilidad	13,95	1,74375
4	Análisis	32,55	4,06875
5	Diseño	41,283333	5,160417
6	Implementación y pruebas	616,96667	77,12083
6.1	Implementación	556,66667	69,58333
6.2	Pruebas funcionales	27,816667	3,477083
6.3	Pruebas de rendimiento	18,166667	2,270833
6.4	Pruebas de usabilidad	14,316667	1,789583
7	Conclusiones	43,683333	5,460417
7.1	Revisión de costes	14,15	1,76875
7.2	Conclusiones	14,75	1,84375
7.3	Trabajo futuro	15,083333	1,885417
	TOTAL	856,76667	107,0958

Tabla 5: Coste temporal esperado final (calculado teniendo en cuenta 8h de trabajo diarios)

El diagrama de Gantt es una herramienta gráfica utilizada en la gestión de proyectos para planificar y programar las tareas a lo largo del tiempo. Proporciona una representación visual de las actividades del proyecto, mostrando las fechas de inicio y finalización de cada tarea, su duración y las dependencias entre ellas.

Este tipo de diagrama permite a los equipos de proyecto y a los interesados tener una visión clara y estructurada de la secuencia de actividades y el cronograma del proyecto. También ayuda a identificar posibles superposiciones o retrasos en las tareas y facilita la asignación eficiente de recursos.

A continuación, se adjunta el diagrama de Gantt que muestra los resultados finales teniendo en cuenta las estimaciones previas realizadas por Kevin, Pablo Mocholí y Diego Ruíz. Este diagrama representa de manera visual las actividades planificadas, su duración y la secuencia en la que se llevarán a cabo. Se ha tenido en cuenta las dependencias entre las tareas y las estimaciones de tiempo proporcionadas por los expertos para crear un plan de proyecto completo.

El diagrama de Gantt proporciona una representación clara y estructurada del cronograma del proyecto, permitiendo una mejor comprensión de las actividades y su secuencia. Además, facilita la identificación de posibles retrasos o superposiciones en las tareas, lo que contribuye a una gestión más eficiente del proyecto.

Es importante tener en cuenta que el diagrama de Gantt es una representación visual y que el proyecto real puede variar en función de factores externos y ajustes necesarios durante su desarrollo.

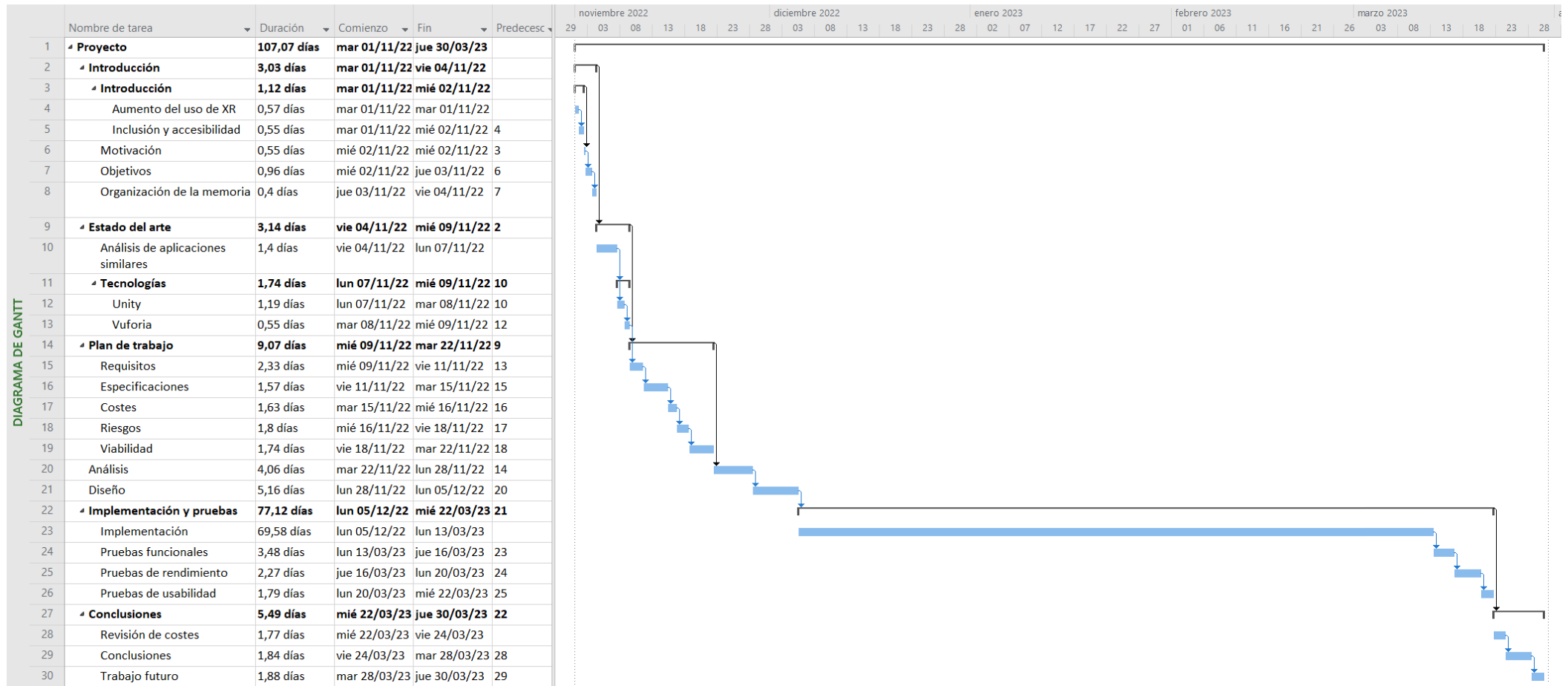


Diagrama 1: Diagrama de Gantt del proyecto (Microsoft Projects)

3.3.2 Coste monetario

En el marco de cualquier proyecto, es fundamental realizar una evaluación y análisis exhaustivo de los costes monetarios involucrados. Estos costes desempeñan un papel crucial en la planificación y ejecución exitosa de un proyecto, y su comprensión adecuada es esencial para garantizar una gestión financiera eficiente.

En este apartado, se abordará el análisis de los costes monetarios asociados al desarrollo del proyecto. Para ello, es importante diferenciar entre dos tipos principales de costes: costes directos y costes indirectos.

Los costes directos se refieren a aquellos gastos específicos y fácilmente atribuibles al proyecto en sí. Estos costes están directamente relacionados con las actividades y recursos utilizados en la ejecución del proyecto. Algunos ejemplos comunes de costes directos pueden incluir la contratación de personal especializado, la adquisición de materiales y equipos necesarios, los gastos de licencias de software o herramientas específicas utilizadas en el desarrollo, entre otros.

Por otro lado, los costes indirectos son aquellos gastos que no pueden atribuirse directamente a una actividad o recurso en particular, pero que son necesarios para el funcionamiento general del proyecto. Estos costes son más generales y abarcan aspectos como los gastos administrativos, los costes de alquiler de instalaciones, los servicios públicos, los seguros, los impuestos y otros gastos generales asociados al proyecto en su conjunto.

Es fundamental tener en cuenta tanto los costes directos como los costes indirectos al realizar una evaluación completa de los costes monetarios del proyecto. Esto permitirá una visión más precisa de los recursos financieros necesarios y contribuirá a una planificación más sólida y realista.

3.3.2.1 Costes directos

Coste de personal

- **Jefe de Proyecto:** El jefe de proyecto es el responsable principal de planificar, coordinar y supervisar la ejecución de un proyecto. Su función es garantizar que todas las tareas se realicen de manera eficiente, dentro del presupuesto y dentro del plazo establecido. Algunas de sus responsabilidades incluyen la definición de los objetivos del proyecto, la asignación de recursos, la gestión del equipo, la resolución de problemas, la comunicación con los interesados y la supervisión del progreso general del proyecto.

En cuanto al sueldo de un jefe de proyecto, puede variar según varios factores, como la ubicación geográfica, el nivel de experiencia, la industria y el tamaño de la empresa. El salario promedio de un jefe de proyecto en España puede oscilar entre los 40.000 y 70.000 euros anuales, usaremos como valor de salario **19,62€/hora** (Talent.com, 2023).

- **Analista Programador:** El analista programador es el profesional encargado de analizar los requisitos del sistema, diseñar soluciones técnicas, desarrollar el software y realizar pruebas para garantizar su funcionalidad y calidad. Su rol implica comprender las

necesidades del cliente, traducirlos en especificaciones técnicas, escribir el código necesario y colaborar con el equipo de desarrollo en la implementación y resolución de problemas.

En cuanto al sueldo de un analista programador, nuevamente, puede variar dependiendo de factores como la experiencia, las habilidades técnicas, la ubicación y el tipo de empresa. El salario promedio puede oscilar entre los 25.000 y 45.000 euros anuales, usaremos como salario **16,41€ por hora** (Talent.com, 2023).

- **Programador:** El Programador es el profesional encargado de escribir y desarrollar el código informático para crear software y aplicaciones. Su tarea principal es traducir los requisitos y diseños técnicos en código funcional, realizar pruebas y corregir errores. Los programadores pueden especializarse en diferentes lenguajes de programación y tecnologías según el proyecto y las necesidades específicas.

El sueldo varía según factores como la experiencia, las habilidades técnicas, la ubicación y el tipo de empresa. Aproximadamente, el salario de un programador puede rondar entre los 20.000 y 35.000 euros anuales, aproximadamente **14,62€ por hora** (Talent.com, 2023).

Tareas	Trabajador	Horas	Sueldo base
Introducción	Jefe de Proyecto	24,8	486,58 €
Estado del Arte	Jefe de Proyecto	25,47	499,72 €
Planificación	Jefe de Proyecto	72,01	1.412,84 €
Análisis	Analista-Programador	32,55	534,15 €
Diseño	Programador	41,28	603,51 €
Implementación y pruebas	Programador	616,97	9.020,10 €
Conclusiones	Jefe de Proyecto	43,68	857,00 €

Tabla 6: Cálculo del sueldo base por tarea

Basándonos en los cálculos realizados para determinar los sueldos base asignados a cada tarea, podemos concluir que el Jefe de Proyecto recibirá un salario base de **3.256,14 €** para la duración total del proyecto. Por otro lado, el Analista Programador tendrá un total de **534,14 €** como sueldo base, mientras que el Programador recibirá un sueldo base de **9.623,61 €**. Sumando estos sueldos base, el gasto total en concepto de salarios asciende a **13.413,89 €**.

Es importante destacar que estos valores corresponden únicamente a los sueldos base y no incluyen las contribuciones de seguridad social u otros beneficios adicionales que pueden ser requeridos por la legislación laboral vigente. Teniendo en cuenta un gasto adicional del 30% por la seguridad social y otros seguros adicionales podemos calcular el gasto total en personal:

Trabajador	Sueldo base total	SS	Gasto/trabajador
Jefe de Proyecto	3.256,14 €	976,84 €	4.232,98 €
Analista-Programador	534,14 €	160,24 €	694,38 €
Programador	9.623,61 €	2.887,08 €	12.510,69 €
Total	13.413,89 €	4.024,17 €	17.438,06 €

Tabla 7: Gasto total por trabajador (teniendo en cuenta un 30% de gastos)

Coste del material

El análisis y gestión de los costes de material desempeña un papel fundamental en cualquier proyecto. Estos costes se refieren a los recursos necesarios para llevar a cabo el proyecto, como equipos, dispositivos, herramientas, software y cualquier otro elemento físico utilizado en el desarrollo.

En este apartado, se abordará en detalle el análisis de los costes de material asociados al proyecto, evaluando cuidadosamente los recursos requeridos y su impacto financiero. La correcta identificación y estimación de estos costes resulta esencial para una planificación precisa y una gestión eficiente del presupuesto del proyecto.

Software

Software	Precio	Cantidad	Total
Windows 10	145 €	3	435 €
Microsoft 365 Estándar	58,50 €	3	175,50 €
Unity	Gratuito		
Microsoft Project	28,10 €	1	28,10 €
Visual Paradigm	313,66	1	313,66
TOTAL	545,26 €		952,26 €

Tabla 8: Precio del software

Hardware

La amortización del hardware es un concepto financiero que se refiere al proceso de distribuir el costo de adquisición de un activo tangible a lo largo de su vida útil estimada. Se utiliza para asignar los costes directos de los equipos o dispositivos utilizados en el desarrollo del proyecto a lo largo del tiempo.

Cuando se adquiere hardware, como ordenadores, servidores, dispositivos móviles u otros equipos necesarios para el proyecto, estos tienen un costo inicial asociado. Sin embargo, en lugar de contabilizar ese costo en su totalidad en el año de adquisición, se distribuye a lo largo de la vida útil esperada del activo.

La vida útil del hardware se basa en una estimación de cuánto tiempo se espera que el equipo sea funcional y relevante para el proyecto. Esta estimación puede variar según el tipo de hardware y la rapidez con la que avanza la tecnología en ese campo específico.

La amortización del hardware permite reflejar de manera más precisa los costes reales del proyecto a lo largo del tiempo, en lugar de cargar el costo completo en un solo período contable. Esto ayuda a suavizar el impacto financiero y a distribuir los costes de manera más equitativa.

En el cálculo de los costes directos del proyecto, la amortización del hardware se tiene en cuenta al considerar el período de tiempo en el que se espera que el hardware sea utilizado en el proyecto.

$$\text{Amortización} = \text{Número de días de uso} * \% \text{ de uso en el proyecto} * \frac{\text{Precio}}{\text{Vida útil}}$$

Hardware	Precio	Vida útil (estimación en días)	N.º días (de vida útil) en el proyecto	% uso en el proyecto	Amortización
MSI Katana GF66	1.349 €	1460	107,09	100%	98,95 €
Oculus Quest 2	549,99 €	1277,5	77,12	72,01%	23,91 €
One plus 8t	350 €	1460	0	72,01%	0 €
TOTAL					122,86 €

Tabla 9: Amortización del coste de hardware

Coste directo final

En conclusión, al analizar los costes directos del proyecto, se ha identificado que los gastos relacionados con los trabajadores ascienden a un total de **17.438,06 €**. Este valor incluye los sueldos base estimados para el Jefe de Proyecto, el Analista Programador y el Programador. Estos costes representan una parte significativa de los recursos financieros destinados a la ejecución del proyecto.

En cuanto al material utilizado, se ha determinado que el coste del software necesario es de **952,26 €**, mientras que el coste del hardware se ha calculado en **122,86 €**. Estos valores reflejan los gastos incurridos en la adquisición de los recursos tangibles utilizados durante el desarrollo del proyecto.

Sumando estos costes directos, el total de los gastos de material asciende a **1.075,12 €**. Esta cifra representa el desembolso financiero necesario para obtener y utilizar los recursos materiales requeridos en el proyecto. Y al sumar la cantidad asociada a los trabajadores podemos afirmar que el gasto total directo en el proyecto es de **18.513,18€**.

3.3.2.2 Costes indirectos

Los costes indirectos de un proyecto son aquellos gastos que no están directamente asociados a una actividad o elemento específico del proyecto, pero que son necesarios para su ejecución. Estos costes son más generales y abarcan aspectos como la administración, la supervisión, los gastos generales y otros recursos compartidos que son utilizados en varias áreas del proyecto.

A diferencia de los costes directos, que se pueden asignar fácilmente a una tarea o componente específico del proyecto, los costes indirectos son más difíciles de rastrear y asignar directamente. Estos costes están relacionados con la infraestructura, los equipos de soporte, los servicios generales, los gastos administrativos, los costes de oficina, los seguros, la capacitación, luz, agua de las oficinas y otros elementos necesarios para el funcionamiento del proyecto en su conjunto.

Al ser tan complicado el rastreo de los costes de una empresa o proyecto se calculan los gastos indirectos mediante un concepto llamado “Overhead” esto es un porcentaje que representa una cantidad aproximada de costes indirectos proporcional a los costes directos del proyecto.

$$\text{Gastos indirectos} = \% \text{ Overhead} * \text{Gastos directos}$$

Dado que este proyecto cuenta únicamente con tres trabajadores, se estima que los gastos indirectos asociados no serán significativos. Teniendo en cuenta que cada miembro del equipo trabajará desde su hogar, los únicos costes reales se derivarán de la administración y los gastos específicos relacionados con reuniones y comunicación. Por lo tanto, se ha establecido un porcentaje de “Overhead” del 15% para representar estos costes indirectos.

La aplicación del “Overhead” del 15% resulta en un total estimado de gastos indirectos de **2.776,98€**. Estos costes reflejan los recursos financieros necesarios para respaldar las actividades generales del proyecto y garantizar su correcto funcionamiento.

3.3.2.2 Costes Totales

El coste total del proyecto ascendería a la cantidad de **21.290,16€**.

3.4 Riesgos

3.4.1 Detección de riesgos

La detección de riesgos es una etapa fundamental en la gestión de cualquier proyecto. Consiste en identificar y evaluar los posibles eventos o situaciones que podrían afectar negativamente el desarrollo y los resultados del proyecto. Estos riesgos pueden surgir de diversas fuentes, como el entorno, los recursos, las tecnologías utilizadas, el equipo de trabajo y los aspectos externos.

Para llevar a cabo la detección de riesgos de manera efectiva, se emplean diversas técnicas y métodos. Una de las aproximaciones comunes es realizar una evaluación exhaustiva de los diferentes componentes del proyecto, analizando las interacciones entre ellos y considerando los posibles escenarios adversos que podrían surgir, se suele dividir el proyecto en cuatro fases para poder analizar estos tramos de manera cerrada y de manera más eficaz, las fases son las siguientes:

- Inicio del proyecto: El inicio del proyecto comprende la introducción y el estado del arte. En esta fase, se establecen los fundamentos del proyecto y se proporciona una visión general de su contexto y objetivos. La introducción proporciona una descripción general del proyecto, su propósito y los beneficios esperados. Además, se presenta el estado del arte, que examina las tecnologías, investigaciones y soluciones existentes relacionadas con el proyecto. Esta fase sienta las bases para el desarrollo del proyecto y ayuda a comprender su relevancia en el contexto actual.

- Planificación: La fase de planificación se divide en análisis y diseño. Durante el análisis, se identifican y documentan los requisitos del proyecto, las necesidades de los interesados y los objetivos específicos que se pretenden lograr. Se lleva a cabo un estudio detallado de los elementos clave del proyecto, como los requisitos funcionales y no funcionales, el alcance, los plazos y los recursos necesarios. A continuación, en la etapa de diseño, se definen las soluciones técnicas y se desarrolla una arquitectura que cumpla con los requisitos establecidos en la fase de análisis. Esto implica la creación de diseños detallados, diagramas y especificaciones técnicas que guiarán la implementación del proyecto.

- Ejecución: La fase de ejecución se divide en implementación y pruebas. Durante la implementación, se lleva a cabo la construcción del proyecto según los diseños y especificaciones desarrollados en la fase de diseño. Se codifican los componentes, se integran y se realizan las configuraciones necesarias para hacer funcionar el sistema. Posteriormente, se procede a la etapa de pruebas, en la cual se verifican y validan las funcionalidades del proyecto. Se realizan pruebas funcionales, de rendimiento y de usabilidad para asegurar que el proyecto cumpla con los requisitos establecidos y funcione de manera óptima.

- Cierre: La fase de cierre del proyecto se enfoca en las conclusiones. En esta etapa, se evalúa el desempeño del proyecto en relación con los objetivos y los requisitos establecidos. Se realiza una revisión de los costes y recursos utilizados, se evalúa el grado de cumplimiento de los plazos y se analiza el resultado final del proyecto. Además, se resumen las lecciones aprendidas y se identifican las áreas de mejora para

proyectos futuros. Esta fase marca el final del proyecto y proporciona una evaluación integral de los resultados obtenidos.

Cada una de estas fases presenta sus propios riesgos y desafíos específicos. Es importante identificar y abordar los posibles riesgos en cada etapa del proyecto para minimizar su impacto. Para llevar a cabo esta tarea es necesario calcular el riesgo mediante una fórmula y después clasificarlo en diferentes niveles (nulo, bajo, moderado, alto e inaceptable).

$$\text{Nivel de riesgo} = \frac{\text{Probabilidad de que ocurra}}{100} * \text{Impacto en el proyecto}$$

Riesgo	Probabilidad	Impacto (semanas)	Nivel de riesgo	Clasificación
Falta de información en estado del Arte	60%	1	0,6	Bajo
Mala definición de requisitos	70%	2	1,4	Moderado
Errores en la planificación y estimación de costes	75%	2	1,5	Moderado
Baja viabilidad del proyecto	10%	4	0,4	Bajo
Estimación de riesgos pobre	45%	1	0,45	Bajo

Tabla 10: Análisis del riesgo en la fase de inicio del proyecto

Riesgo	Probabilidad	Impacto (semanas)	Nivel de riesgo	Clasificación
Errores en los casos de uso	50%	1	0,5	Bajo
Planificación demasiado superficial	35%	4	1,4	Moderado

Tabla 11: Análisis del riesgo en la fase de planificación

Riesgo	Probabilidad	Impacto (semanas)	Nivel de riesgo	Clasificación
Falta de personal debido a baja	20%	1	0,2	Bajo
Falta de personal debido a abandono	5%	10	0,5	Bajo
Retrasos en el tiempo planificado	90%	2	1,8	Alto
Mala cooperación	30%	4	1,2	Moderado
Falta de seguimiento del proyecto	10%	5	0,5	Bajo

Tabla 12: Análisis de riesgo en la fase de ejecución

Riesgo	Probabilidad	Impacto (semanas)	Nivel de riesgo	Clasificación
Errores desconocidos	20%	3	0,6	Bajo
Producto que no cumple requisitos	5%	10	0,5	Bajo
Documentación final de mala calidad	50%	2	1	Alto

Tabla 13: Análisis del riesgo en la fase de cierre

3.4.2 Amortiguación y contingencia

La amortiguación y la contingencia son dos estrategias utilizadas en la gestión de riesgos para abordar los posibles eventos o situaciones adversas que podrían afectar un proyecto. Ambas estrategias tienen como objetivo reducir o evitar los impactos negativos de los riesgos identificados.

La amortiguación de riesgos se refiere a la implementación de acciones preventivas y proactivas para reducir la probabilidad de que ocurran los riesgos o minimizar su impacto en caso de que se materialicen. La mitigación se lleva a cabo antes de que ocurra el riesgo, a través de medidas como la planificación adecuada, la asignación de recursos, la implementación de controles de calidad, la capacitación del personal, entre otros. La idea principal es prevenir o reducir la ocurrencia y el impacto de los riesgos, abordando las causas subyacentes y tomando medidas para evitar su manifestación.

La contingencia de riesgos, por otro lado, implica la preparación de planes alternativos para hacer frente a los riesgos en caso de que se materialicen. Se trata de tener planes de acción y recursos adicionales disponibles para responder rápidamente ante situaciones adversas. La contingencia se enfoca en la respuesta y la recuperación después de que ocurre el riesgo, buscando minimizar los daños y restaurar el proyecto a su estado normal lo más rápido posible. Esto puede incluir la implementación de planes de contingencia, el establecimiento de procedimientos de respuesta, la asignación de responsabilidades claras y la disponibilidad de recursos de respaldo.

La mitigación y la contingencia son estrategias complementarias en la gestión de riesgos. La mitigación se centra en la prevención y la reducción de los riesgos antes de que ocurran, mientras que la contingencia se enfoca en la preparación y respuesta ante los riesgos una vez que se materializan.

En este caso, no se ha considerado necesario implementar medidas de mitigación, ya que no se identificaron riesgos que requirieran de acciones preventivas específicas. Sin embargo, se ha previsto la creación de un plan de contingencia para los riesgos clasificados como altos, de manera que se cuente con soluciones rápidas y efectivas en caso de que estos riesgos se materialicen. La preparación de un plan de contingencia permitirá una respuesta ágil y organizada frente a situaciones adversas, minimizando así los impactos negativos y facilitando la recuperación del proyecto.

Retrasos en el tiempo planificado:

- Reasignación de recursos: Evaluar la disponibilidad y las capacidades del equipo de trabajo y considerar la posibilidad de reasignar recursos adicionales a las tareas críticas o que se encuentren con retraso. Esto ayudará a acelerar el progreso y reducir los tiempos de finalización.
- Realizar una revisión exhaustiva del cronograma actual y ajustarlo para reflejar los nuevos plazos y prioridades. Identificar las tareas más críticas y tomar medidas para acelerar su ejecución. Reasignar recursos, redistribuir las cargas de trabajo y establecer hitos y fechas límite realistas.

- Implementación de técnicas de gestión del tiempo: Utilizar técnicas de gestión del tiempo, como la compresión del cronograma, el enfoque en tareas críticas y el establecimiento de prioridades claras. Utilizar herramientas de gestión de proyectos que permitan el seguimiento y la coordinación efectiva de las actividades.
- Incremento de la comunicación y coordinación: Mejorar la comunicación y la coordinación entre los miembros del equipo, los interesados y otras partes involucradas en el proyecto. Esto ayudará a identificar y abordar los obstáculos y retrasos de manera más eficiente, así como a mantener a todos informados sobre los cambios en el cronograma.

Documentación final de mala calidad:

- Evaluación y análisis de la brecha: Realizar una evaluación exhaustiva de la documentación final para identificar las áreas o aspectos que no cumplen con las expectativas. Analizar las discrepancias entre lo entregado y lo esperado, identificando las razones detrás de estas diferencias.
- Comunicación abierta y transparente: Establecer una comunicación clara y abierta con los interesados y partes involucradas.
- Identificación de soluciones: Explorar opciones para mejorar la documentación y cumplir con las expectativas establecidas. Esto puede implicar la revisión y edición de la documentación existente, la adición de secciones faltantes o la realización de investigaciones o análisis adicionales.
- Asignación de responsabilidades y recursos: Definir claramente las responsabilidades y asignar los recursos necesarios para implementar las mejoras en la documentación (requiere más gastos).
- Revisión y aprobación adicional: Realizar revisiones y evaluaciones adicionales de la documentación mejorada para garantizar que ahora cumpla con las expectativas establecidas. Obtener la aprobación de los interesados y partes involucradas antes de finalizar y entregar la versión corregida de la documentación.

3.5 Viabilidad

Este proyecto de desarrollo de una aplicación de realidad virtual y aumentada con funcionalidad de Speech to Text para personas con diversidad funcional auditiva se puede calificar como viable basándonos en:

Demanda y necesidad: Existe una creciente demanda en el mercado de soluciones que mejoren la inclusión y accesibilidad de las personas con diversidad funcional. La falta de mecanismos de inclusión en la realidad virtual actualmente limita el acceso y la participación de este grupo de usuarios. Este proyecto aborda directamente esta necesidad, ofreciendo un paquete de Unity con una solución innovadora que permitiría a otros desarrolladores crear una aplicación con los recursos que permite a personas con

diversidad funcional auditiva disfrutar plenamente de las experiencias de realidad virtual y aumentada.

Ventajas competitivas: Esta aplicación se destaca por su capacidad de cambiar el radio de escucha, identificar al hablante y proporcionar transcripciones en tiempo real. Estas funcionalidades ofrecen una experiencia más accesible e interactiva para las personas con diversidad funcional auditiva, lo que la diferencia de otras soluciones existentes. Además, el uso de Unity y Vuforia como herramientas de desarrollo brinda una gran compatibilidad y flexibilidad, lo que facilita la creación de experiencias inmersivas y de alta calidad.

Análisis financiero: El análisis de costos demuestra que el proyecto es financieramente viable. Los gastos directos e indirectos, incluyendo los costos de desarrollo de software y hardware, se han tenido en cuenta y se ha estimado un retorno de inversión positivo. Además, se ha identificado un potencial de ingresos a través de modelos de negocio como ventas de la aplicación, suscripciones u otras fuentes de ingresos relacionadas con la realidad virtual y aumentada como se ha visto en otros ejemplos.

Capítulo 4

Análisis

4.1 Realidad Virtual

- 4.1.1 Casos de uso
- 4.1.2 Diagramas de secuencia
- 4.1.3 Diagrama de actividad general
- 4.1.4 Diagramas de estado

4.2 Realidad Aumentada

- 4.2.1 Casos de uso
- 4.2.2 Diagramas de secuencia
- 4.2.3 Diagrama de actividad general
- 4.2.4 Diagramas de estado

4. Análisis

4.1 Realidad Virtual

4.1.1 Casos de uso

Los casos de uso son una técnica utilizada en el análisis y diseño de sistemas y software para describir interacciones específicas entre un sistema y sus actores. Un caso de uso representa una función o proceso específico que el sistema debe ser capaz de realizar para satisfacer los requisitos de los usuarios.

En pocas palabras, un caso de uso describe una situación en la que un usuario interactúa con el sistema para lograr un objetivo particular. Define las acciones que realiza el usuario y cómo el sistema responde a esas acciones. Los casos de uso se utilizan para capturar los requisitos funcionales del sistema y proporcionan una visión detallada de cómo se utilizará el sistema en la práctica.

4.1.1.1 Diagrama de casos de uso

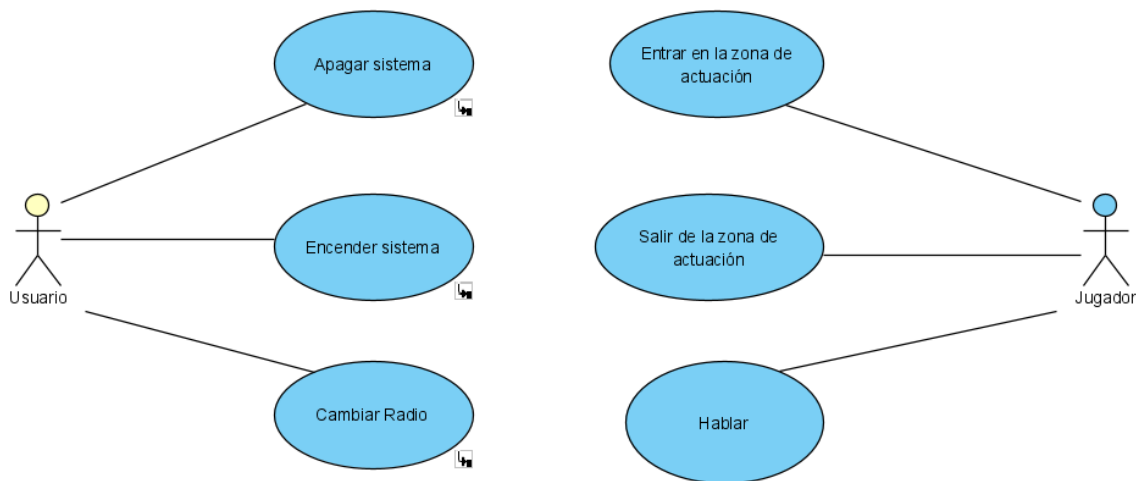


Diagrama 2: Diagrama de casos de uso del proyecto (VR)

4.1.1.2 Descripción de casos de uso

Caso de uso	Encender sistema
Actores	Usuario
Propósito	Encender el sistema para que este pendiente de eventos
Resumen	El caso inicia el sistema de manera que si hay alguien dentro activaría el speechtexter, si no es así continuara esperando la entrada de otro jugador a su rango de escucha.
Tipo	Primario y esencial
Precondiciones	Que el sistema se encuentre apagado
Postcondiciones	Se debe mostrar el estado de encendido, indicando si hay alguien en el rango (por lo tanto, el speechtexter está en funcionamiento) o no.

Tabla 14: Caso de uso de "Encender sistema" (VR)

Flujo de eventos principal	Acción del actor	Respuesta del sistema
	1. El caso de uso comienza cuando el usuario pulsa el botón de “power” de la interfaz.	2. El sistema cambia el estado a encendido y lo muestra por pantalla.

Tabla 15: Flujo de eventos principal de “Encender sistema” (VR)

Caso de uso	Apagar sistema
Actores	Usuario
Propósito	Apagar el sistema para que no esté pendiente de eventos y no muestre el radio de actuación
Resumen	El caso apaga el sistema y no muestra radio ni transcribe la voz (aunque haya jugadores dentro de radio)
Tipo	Primario y esencial
Precondiciones	Que el sistema se encuentre encendido
Postcondiciones	Se debe mostrar el estado de apagado y no mostrar el radio de actuación (ya que este no existe)

Tabla 16: Caso de uso de “Apagar sistema” (VR)

Flujo de eventos principal	Acción del actor	Respuesta del sistema
	1. El caso de uso comienza cuando el usuario pulsa el botón de “power” de la interfaz.	2. El sistema cambia el estado a apagado y lo muestra por pantalla.

Tabla 17: Flujo de eventos principal de “Apagar sistema” (VR)

Caso de uso	Cambiar radio
Actores	Usuario
Propósito	Cambiar el radio de acción del sistema speechtexter (aumentar o reducir)
Resumen	Cambia el radio de acción del sistema al utilizar una slider en el interfaz, esta slider define tanto el tamaño del radio visual como el detector de hablantes.
Tipo	Primario y esencial
Precondiciones	
Postcondiciones	El radio no puede ser 0 ya que sería lo mismo que pulsar el botón de apagar

Tabla 18: Caso de uso de “Cambiar radio” (VR)

Flujo de eventos principal	Acción del actor	Respuesta del sistema
	1. El caso de uso comienza cuando el usuario utiliza la slider del interfaz	2. El sistema cambia el radio de actuación y el del elemento visual que lo muestra. 3. El sistema detecta si al cambiar el radio un nuevo jugador se encuentra en la zona de acción.

Tabla 19: Flujo principal de "Cambiar radio" (VR)

Caso de uso	Entrar en la zona de actuación
Actores	Jugador
Propósito	Entrar en la zona que permite la transcripción de voz
Resumen	Mediante un “collider” esférico se detecta a los jugadores que colisionen con él, asignando un id que permita identificarlo como hablante en el caso de que hable. Gracias a este proceso el speechtexter no gasta recursos si no hay nadie cerca.
Tipo	Secundario y esencial
Precondiciones	El sistema debe de estar encendido
Postcondiciones	Si es el primer jugador que entra en la zona debe de señalizarse que el speechtexter está a la espera

Tabla 20: Caso de uso de "Entrar en la zona de actuación" (VR)

Flujo de eventos principal	Acción del actor	Respuesta del sistema
	1. El caso de uso comienza cuando el sistema detecta un jugador entrar al “collider”.	2. El sistema guarda la información del jugador y le asigna un id. 3. El sistema escribe en la interfaz principal “Escuchando...”
	3. Si este jugador habla se le selecciona como el hablante y se le marca con un elemento visual.	

Tabla 21: Flujo principal de "Entrar en la zona de actuación" (VR)

Caso de uso	Salir de la zona de actuación
Actores	Jugador
Propósito	Salir de la zona que permite la transcripción de voz
Resumen	Mediante un “collider” esférico se detecta a los jugadores que abandonan la zona de transcripción que rodea al usuario, al abandonar este radio el identificador del jugador se borra y se deja de tener en cuenta en el sistema
Tipo	Secundario y esencial
Precondiciones	El sistema debe de estar encendido
Postcondiciones	Si el último jugador de la zona de actuación abandona esta el sistema debe señalizar que el speechtexter no está a la espera debido a que no hay posibles jugadores con los que interactuar

Tabla 22: Caso de uso de "Salir de la zona de actuación" (VR)

Flujo de eventos principal	Acción del actor	Respuesta del sistema
	1. El caso de uso comienza cuando el sistema detecta un jugador abandonar la zona de actuación.	2. El sistema elimina la información del jugador. 3. El sistema escribe en la interfaz principal “No hay nadie cerca”.

Tabla 23: Flujo principal de "Salir de la zona de actuación" (VR)

Caso de uso	Hablar
Actores	Jugador
Propósito	Comunicarse con el usuario (un jugador también)
Resumen	Al hablar el speechtexter que se encontraba a la espera de una entrada de audio, inicia los scripts relacionados con el speech to text y convierte el audio en una variable de texto que se mostrara en la interfaz principal.
Tipo	Principal y esencial
Precondiciones	El sistema debe de estar encendido y el jugador encontrarse en el radio de actuación
Postcondiciones	Señalizar debidamente al hablante

Tabla 24: Caso de uso de "Hablar" (VR)

Flujo de eventos principal	Acción del actor	Respuesta del sistema
	1. El caso de uso comienza cuando el sistema detecta un jugador hablar.	2. El sistema activa los scripts correspondientes. 3. El sistema convierte la entrada de audio en una variable de texto. 4. El sistema muestra la frase en el interfaz principal del usuario.

Tabla 25: Flujo principal de "Hablar" (VR)

4.1.2 Diagramas de secuencia

Después de haber identificado y comprendido los casos de uso del proyecto, los diagramas de secuencia proporcionan una representación visual clara de las interacciones entre los actores y el sistema en un escenario específico.

Estos diagramas muestran la secuencia de acciones y mensajes que se intercambian entre los objetos del sistema a medida que se ejecuta un caso de uso. Proporcionan una vista detallada de cómo se comunican los objetos y cómo se realizan las operaciones en el sistema en respuesta a las acciones de los actores.

El objetivo principal de los diagramas de secuencia es definir cómo se lleva a cabo una funcionalidad específica, detallando el orden de las acciones, las invocaciones de métodos, los eventos y las respuestas correspondientes.

4.1.2.1 Diagrama de secuencia de “Encender sistema”

Se enciende el sistema y muestra el estado actual que depende de si se encuentra un jugador dentro de la zona de acción o no (el usuario es un jugador, pero no es un posible hablante por lo que no aplica como jugador real para estos casos de uso).

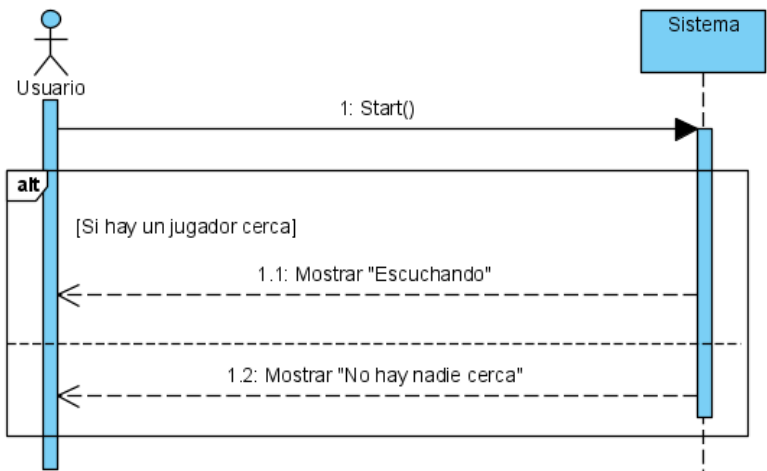


Diagrama 3:Diagrama de secuencia de "Encender sistema" (VR)

4.1.2.2 Diagrama de “Apagar sistema”

Apaga el sistema y entra en un estado en el cual el único evento que puede haber es encenderlo de nuevo, se reemplaza el texto de la interfaz principal donde podría estar transcrita la conversación, indicado el estado de escucha o la falta de jugadores en el radio.

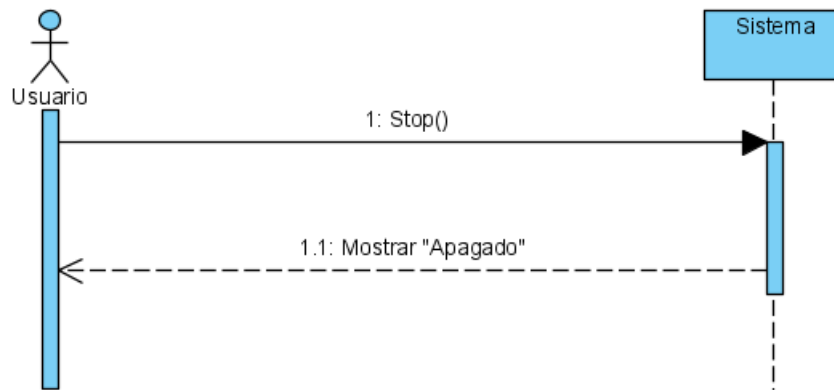


Diagrama 4: Diagrama de secuencia de "Apagar sistema" (VR)

4.1.2.3 Diagrama de “Cambiar radio”

El radio siempre se muestra con un bucle que genera una onda circular en el suelo que comienza en el centro (el usuario) y cuando alcanza el radio establecido con la slider, al cambiar el radio se ajusta este valor y se aplica en los objetos necesarios.

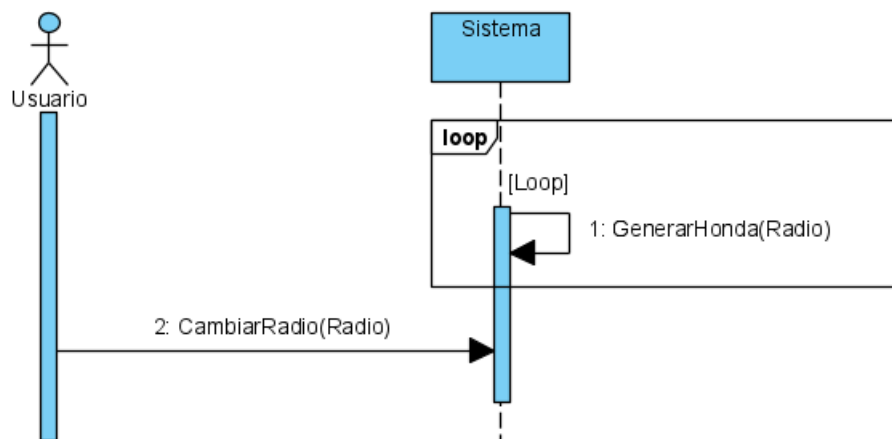


Diagrama 5: Diagrama de secuencia de “Cambiar radio” (VR)

4.1.2.4 Diagrama de “Entrar en la zona de actuación”

Al entrar un jugador en contacto con el “collider” (ya sea porque se acerque el usuario o el jugador) este es registrado para identificarlo y utilizarlo en otros métodos y si es el primer jugador en entrar a la zona de acción cambiará el texto de la interfaz de “No hay nadie cerca” a “Escuchando...”

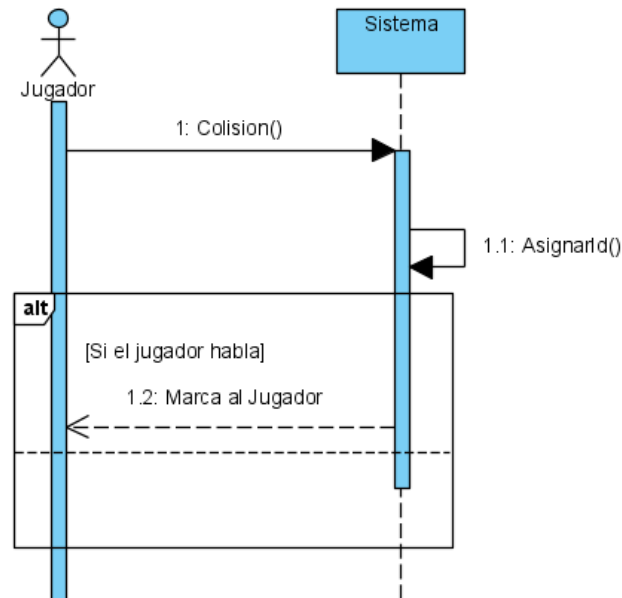


Diagrama 6: Diagrama de secuencia de "Entrar en la zona de acción" (VR)

4.1.2.5 Diagrama de “Salir de la zona de acción”

Al salir del radio de acción se elimina la información del jugador y si la zona se queda sin jugadores se muestra que no hay nadie lo suficientemente cerca como para detectarlo, para que así el usuario pueda o moverse o aumentar el radio de escucha.

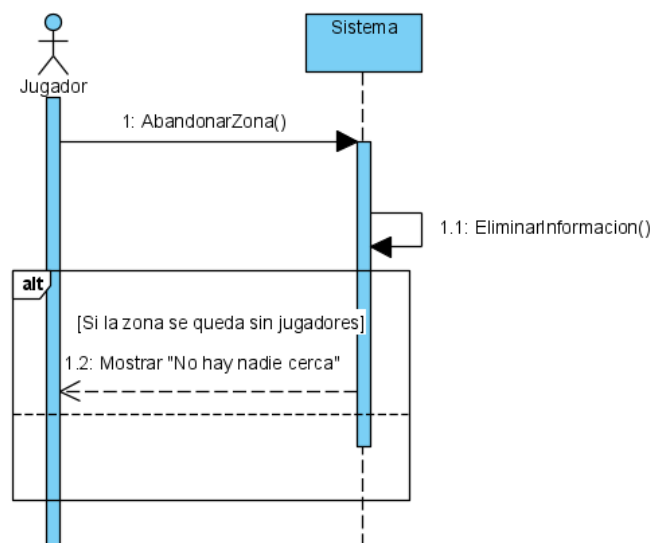


Diagrama 7: Diagrama de secuencia de "Salir de la zona de acción" (VR)

4.1.2.6 Diagrama de “Hablar”

El sistema está en constante búsqueda de frases (en español), cuando el jugador habla este lo detecta y transcribe la información del audio de entrada en texto.

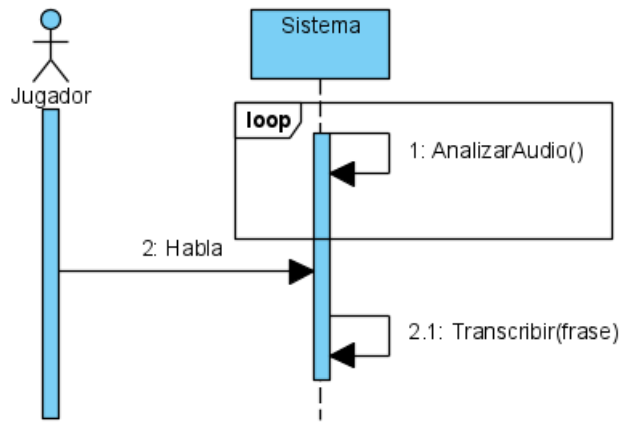


Diagrama 8: Diagrama de secuencia de "Hablar" (VR)

4.1.3 Diagrama de actividad general

En el diagrama de actividad se representa el flujo de acciones que ocurren en el sistema al iniciar el programa. La interacción comienza al pulsar el botón de encendido, lo que da inicio al funcionamiento del sistema. Durante todo el proceso, el radio de escucha se muestra constantemente en pantalla. Si el sistema no detecta la presencia de ningún jugador, permanecerá en espera y mostrará el mensaje "No hay nadie cerca". Sin embargo, cuando detecta a un jugador, activará la función de speechtexter y mostrará en pantalla el mensaje "Escuchando". Cuando un jugador habla, se le señalará visualmente y se transcribirá su mensaje en tiempo real. Es importante destacar que, en caso de que todos los jugadores abandonen la zona o el usuario se aleje, el speechtexter se apagará y el sistema esperará a encontrar un nuevo jugador, mostrando nuevamente el mensaje "No hay nadie cerca". En cualquier momento, es posible pulsar el botón de apagado para detener el sistema.

Cabe mencionar que la acción de cambiar el radio de escucha no se ha incluido en el diagrama de actividad, ya que esta función es independiente del resto de las acciones. Puede realizarse en cualquier momento, incluso con el sistema apagado.

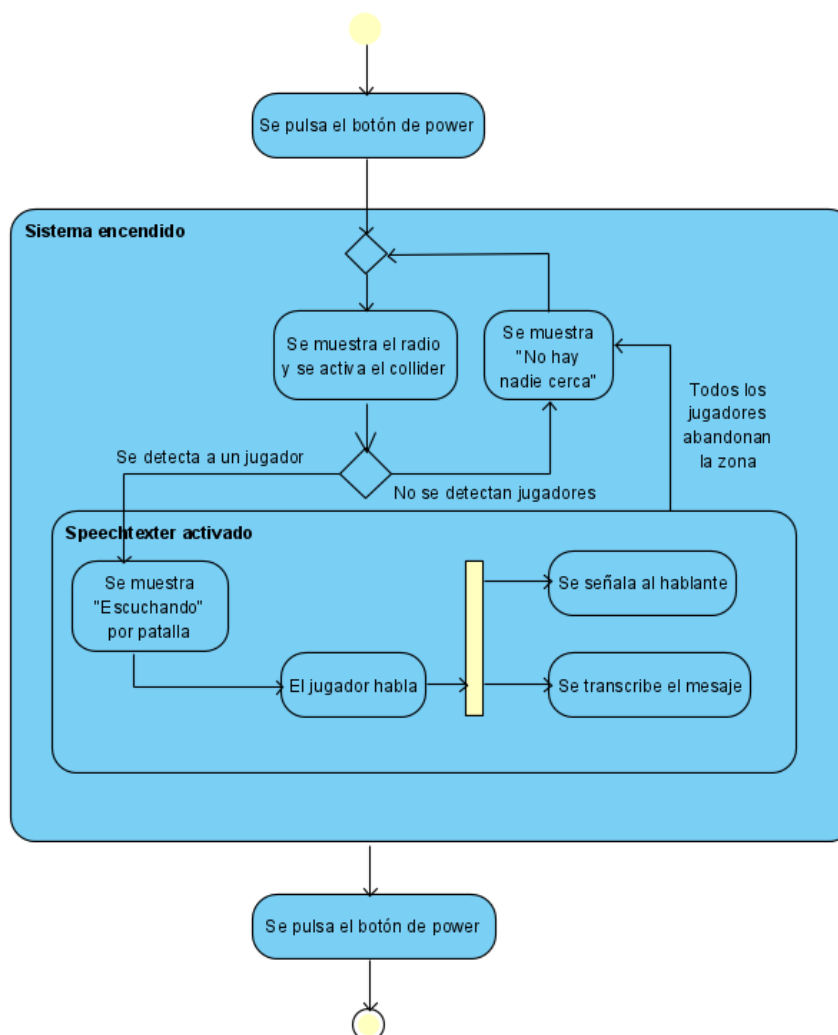


Diagrama 9: Diagrama de actividad de la aplicación (VR)

4.1.4 Diagrama de estado

En la aplicación, el sistema puede encontrarse en tres estados diferentes:

- Apagado: En este estado, el sistema no muestra el radio de escucha, no realiza transcripciones ni busca jugadores. Todas las funcionalidades están desactivadas menos la que permite variar el radio de efecto.
- Encendido: Cuando el sistema se encuentra en este estado, se muestra el radio de escucha y está en espera de detectar nuevos jugadores.
- Escuchando: Este estado indica que el sistema ha encontrado un jugador y se encuentra en el proceso de escucha y transcripción de su mensaje. Durante este estado, el sistema está encendido y realizando la tarea de capturar y transcribir la voz del jugador.

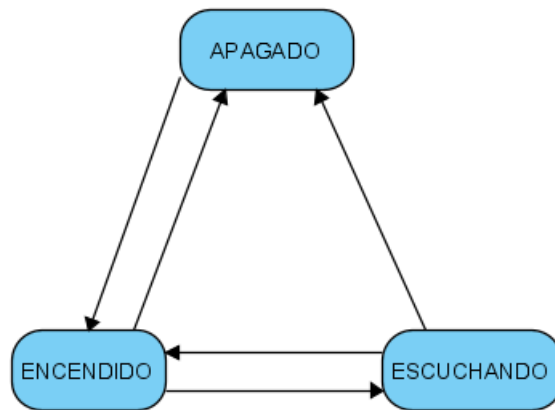


Diagrama 10: Diagrama de estados del sistema (VR)

4.2 Realidad Aumentada

Debido a que la versión de la aplicación para realidad aumentada (smartphones) es igual a la de la realidad virtual, pero con menos requisitos y funcionalidades, los siguientes diagramas serán similares y un poco más escuetos. Esto se debe a que en la realidad aumentada no es necesario la detección de personas ni el uso de un limitador espacial para el uso del speech to text (la zona de acción en la aplicación de realidad virtual), en este caso solo se limita a transcribir la entrada de audio siempre que el sistema esté encendido.

4.2.1 Casos de uso

4.2.1.1 Diagrama de casos de uso

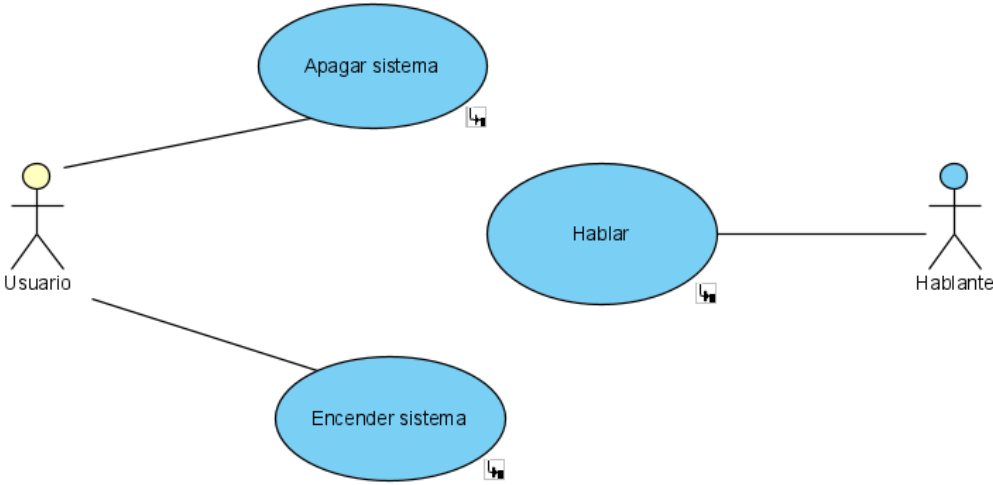


Diagrama 11: Diagrama de casos de uso (AR)

4.2.1.2 Descripción de casos de uso

Caso de uso	Encender sistema
Actores	Usuario
Propósito	Encender el sistema para que este pendiente de eventos
Resumen	El caso inicia el sistema de manera que si en la entrada de audio se detecta una frase en español será procesada y escrita.
Tipo	Primario y esencial
Precondiciones	Que el sistema se encuentre apagado
Postcondiciones	Se debe mostrar el estado de encendido

Tabla 26: Caso de uso de "Encender sistema" (AR)

Flujo de eventos principal	Acción del actor	Respuesta del sistema
	1. El caso de uso comienza cuando el usuario pulsa el botón de “power” de la interfaz.	2. El sistema cambia el estado a encendido y lo muestra por pantalla.

Tabla 27: Flujo principal de "Encender sistema" (AR)

Caso de uso	Apagar sistema
Actores	Usuario
Propósito	Apagar el sistema para que no esté pendiente de eventos
Resumen	El caso apaga el sistema y no transcribe la voz
Tipo	Primario y esencial
Precondiciones	Que el sistema se encuentre encendido
Postcondiciones	Se debe mostrar el estado de apagado

Tabla 28: Caso de uso de "Apagar sistema" (AR)

Flujo de eventos principal	Acción del actor	Respuesta del sistema
	1. El caso de uso comienza cuando el usuario pulsa el botón de “power” de la interfaz.	2. El sistema cambia el estado a apagado y lo muestra por pantalla.

Tabla 29: Flujo principal de "Apagar sistema" (AR)

Caso de uso	Hablar
Actores	Hablante
Propósito	Comunicarse con el usuario
Resumen	Al hablar el speechtexter que se encontraba a la espera de una entrada de audio, inicia los scripts relacionados con el speech to text y convierte el audio en una variable de texto que se mostrara en la interfaz principal.
Tipo	Principal y esencial
Precondiciones	El sistema debe de estar encendido
Postcondiciones	

Tabla 30: Caso de uso de "Hablar"

Flujo de eventos principal	Acción del actor	Respuesta del sistema
	1. El caso de uso comienza cuando el sistema recibe una entrada de audio	2. El sistema activa los scripts correspondientes. 3. El sistema convierte la entrada de audio en una variable de texto. 4. El sistema muestra la frase en el interfaz principal del usuario.

Tabla 31: Flujo principal de "Hablar" (AR)

4.2.2 Diagramas de secuencia

4.1.2.1 Diagrama de secuencia de "Encender sistema"

Se enciende el sistema y muestra el estado actual

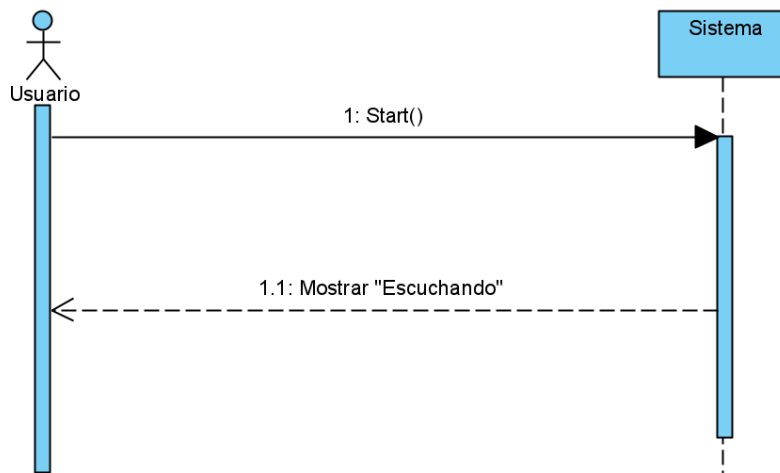


Diagrama 12: Diagrama de secuencia de "Encender sistema" (AR)

4.1.2.2 Diagrama de secuencia de "Apagar sistema"

Apaga el sistema y lo muestra en la interfaz, no ocurrirá ningún cambio excepto por la imagen la cámara hasta su nueva activación.

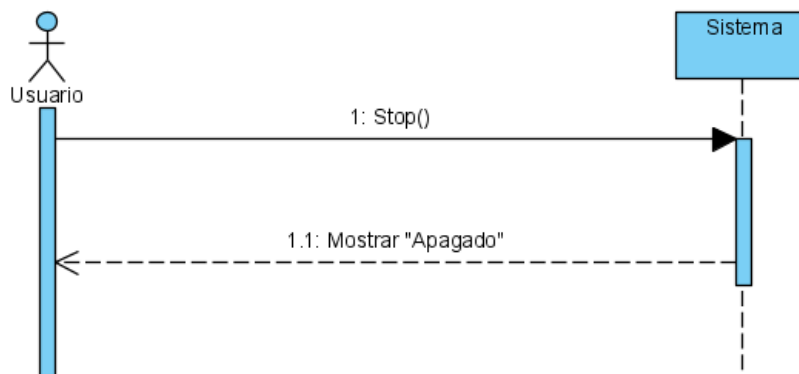


Diagrama 13: Diagrama de secuencia de "Apagar sistema" (AR)

4.1.2.4 Diagrama de secuencia de “Hablar”

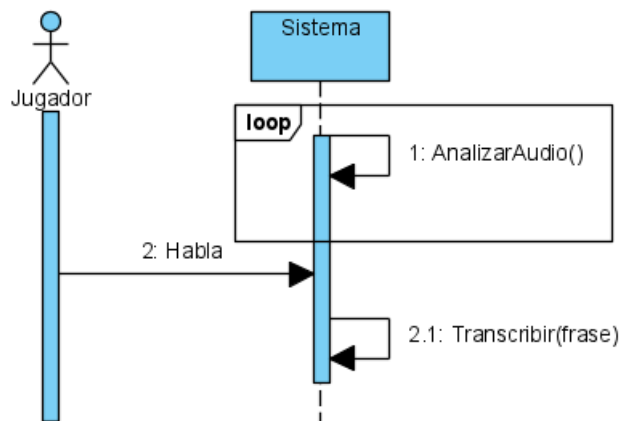


Diagrama 14: Diagrama de secuencia de "Hablar" (AR)

4.1.3 Diagrama de actividad general

Podemos observar que el diagrama de actividad general es mucho más simple el de realidad virtual ya que no hay condiciones, al encenderlo transcribe lo q escucha y al apagarlo no.

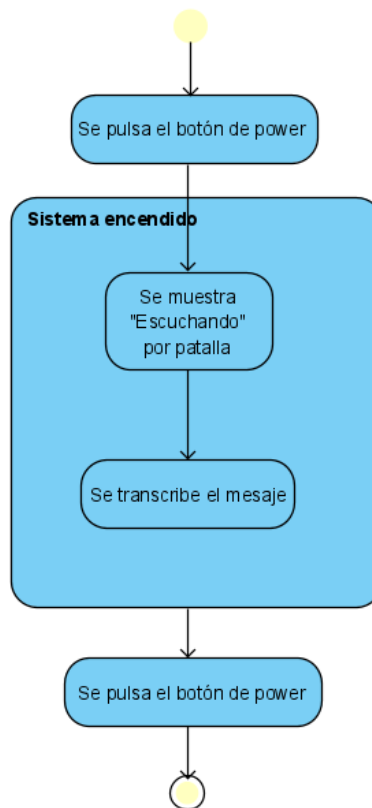


Diagrama 15: Diagrama de actividad general (AR)

4.2.4 Diagrama de estado

En el caso de la realidad aumentada el sistema solo tiene dos estados posibles lo que lo vuelve muy simple:

- Apagado: En este estado no realiza transcripciones y pone “Apagado”. Lo único que se puede observar es la imagen de la cámara
- Encendido: Cuando el sistema se encuentra en este estado está preparado para transcribir el audio de entrada poniendo en la interfaz “Escuchando” hasta que se transcriba la voz del hablante.

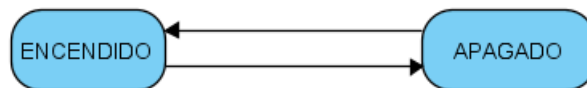


Tabla 32: Diagrama de estados (AR)

Capítulo 5

Diseño

5.1 Interfaz

5.2 Otros elementos

5. Diseño

El diseño de un proyecto es una etapa crucial prestando atención al diseño para garantizar una interfaz intuitiva y atractiva, así como otros elementos importantes para una experiencia inmersiva en la realidad virtual y aumentada.

En primer lugar, se explorará el diseño de la interfaz, que desempeña un papel fundamental en la comunicación entre los usuarios y la aplicación. Se buscará una interfaz clara y fácil de usar, que permita una interacción fluida y eficiente. Se prestará especial atención a la disposición de los elementos, la elección de colores y la legibilidad del texto.

Además de la interfaz, se analizarán otros aspectos del diseño que contribuyen a la inmersión y la comprensión de la experiencia. Uno de ellos es el diseño del escenario en Unity, que crea el entorno virtual en el que se desarrolla la aplicación.

Otro elemento importante del diseño es la visualización del radio, que muestra el alcance y la extensión del sistema de Speech to Text. Se buscará una representación gráfica clara del radio de escucha para que los usuarios puedan comprender y controlar el alcance de la interacción. También se decidirá el diseño de la marca del hablante, que tiene como objetivo facilitar la identificación y el seguimiento de los personajes que hablan en la aplicación.

5.1 Interfaz

Realidad Virtual

La creación del diseño de la interfaz principal involucra tomar decisiones estratégicas sobre su posición y apariencia. Después de considerar varias opciones, se descartó la idea de utilizar subtítulos fijos en la pantalla debido a las posibles molestias y dificultades de lectura que podrían causar en el usuario de las gafas de realidad virtual. Es importante tener en cuenta que los objetos ubicados lejos del centro de visión pueden ser más difíciles de percibir con las gafas puestas.

Otra opción considerada fue generar una interfaz similar a un bocadillo de cómic, que aparecería sobre los jugadores que están hablando, mostrando el texto transcrito. Sin embargo, se determinó que mover constantemente el bocadillo junto con la cámara para visualizarlo al apartar la mirada limitaría la experiencia del usuario y resultaría invasiva.

Tras evaluar estas opciones, se concluyó que la mejor solución era crear una interfaz 3D posicionada como una pequeña mesa en la parte delantera inferior de la cámara. Esta ubicación permite un fácil acceso e interacción con los elementos de la interfaz, al tiempo que se evita una visualización constante cuando el usuario no esté interesado en la transcripción. Además, esta posición no resulta invasiva para la vista y se adapta bien a entornos de comunicación y reuniones, matiz de extrema importancia, siendo la interacción con otros el objetivo principal del metaverso.



Ilustración 4: Escena de la película “Ironman” (escritorio holográfico)

Para el diseño de la interfaz, se ha seguido la idea de una mesa virtual que evoca la estética de una mesa de mezclas o una escena futurista, similar a las representadas en diversas películas contemporáneas (ver ilustración 4) pero, al contrario de estos casos, destacando por su simplicidad, consiguiendo así una interacción intuitiva y sencilla. Se ha creado un mockup que refleja esta visión de la interfaz:



Ilustración 5: Mockup de la interfaz principal (VR)

El diseño de la interfaz se ha organizado de la siguiente manera: el elemento más a la izquierda es un cuadro de texto donde se mostrará la transcripción en tiempo real. A continuación, se encuentra una slider que permite al usuario definir el radio de actuación del sistema. En la parte superior derecha se ubica un botón de encendido/apagado que controla el funcionamiento del sistema. Esta disposición facilita el acceso y la interacción con los elementos principales de la interfaz, pudiendo alcanzar los dos objetos interactivos con la misma mano (derecha) de manera cómoda, brindando una experiencia intuitiva y cómoda para el usuario.

Tras analizar diferentes ejemplos en el estado del arte, se ha observado que los subtítulos en tecnologías de realidad virtual y aumentada suelen utilizar un color blanco debido a su alta legibilidad. Esta elección se encuentra presente en muchas plataformas de “streaming” y se debe a que el color blanco es poco común en el entorno cotidiano, lo cual ayuda a destacar los subtítulos. Sin embargo, para garantizar la legibilidad en todo momento, se ha añadido un fondo con una tonalidad más oscura, siguiendo el

enfoque utilizado en los subtítulos de Vacation Simulator. Esta combinación de un texto en color blanco sobre un fondo oscuro asegura una óptima legibilidad, incluso en escenarios virtuales donde pueden existir diferentes niveles de iluminación.

Realidad Aumentada

La implementación de la interfaz para la aplicación de realidad aumentada sigue una dirección similar a la de la realidad virtual. Sin embargo, debido a la comodidad del uso de un smartphone, no es necesario preocuparse por la ubicación de los elementos cerca de los límites de la pantalla, ya que se puede aprovechar toda su superficie sin riesgo de resultar incómodo para el usuario. En esta ocasión, el botón de apagado se ha colocado en la parte superior derecha de la interfaz, cerca de la zona de botones del teléfono o de los gestos utilizados para la navegación, si el dispositivo no cuenta con botones físicos. Esta ubicación garantiza un acceso fácil y cómodo para el usuario, además de estar en un lugar familiar donde suele interactuar habitualmente. Se ha creado un mockup que refleja el interfaz planeado:



Ilustración 6: Mockup de la interfaz (AR)

5.2 Otros elementos

Visualización del radio

Para la visualización del radio se ha decidido crear una onda circular que se origina desde los pies del usuario hasta alcanzar el radio seleccionado en el que se puede interactuar con otros jugadores. Este diseño se asemeja a la visualización habitual de un sonar, elemento muy similar a este proyecto ya que consiste en representar en una pantalla elementos detectados mientras sonidos.

Tras un estudio se ha elegido asignar a dicha onda un color naranja, debido a que es un color cálido y llamativo que atrae la atención del espectador de manera efectiva. Tiene asociaciones con la energía, la creatividad y la confianza. En adición a esto es el color con el que se asocia en la ETSE (Escuela Técnica Superior de Ingeniería) a la

carrera que curso y pretendo superar con este proyecto, Ingeniería Multimedia, por lo que le confiere un significado especial.

Visualización del hablante

Para proporcionar una indicación clara del hablante en la aplicación, se ha creado un archivo PNG que se coloca en la cabeza del jugador (estando la misma en el centro). Este archivo se orienta continuamente hacia la cámara mediante un script, lo que garantiza que siempre se visualice de la misma forma, sin importar el ángulo de visión del usuario.



Ilustración 7: Marca de hablante

En consonancia con la estética y diseño del proyecto, se ha seleccionado una imagen compuesta por dos ondas, que recuerdan al símbolo asociado al aumento de volumen. Esta elección se ha realizado considerando la coherencia visual y la intención de comunicar la acción de hablar. Para mantener la cohesión estética, se ha utilizado el mismo color que la onda circular que se expande desde el suelo, creando así una conexión visual entre ambos elementos.

Escenario virtual

A pesar de que el escenario en sí no juega un papel crucial en la interacción del proyecto, se ha optado por aplicar una estética "low-poly" al entorno virtual. Esta elección estética se caracteriza por la utilización de formas y modelos con un número limitado de polígonos, lo que resulta en un aspecto simplificado y minimalista.

El escenario diseñado representa un parque virtual con elementos comunes como bancos, mesas y fuentes. La intención detrás de esta elección es crear un entorno agradable y familiar para los usuarios, sin destacar ni distraer su atención del objetivo principal: la interacción entre los jugadores a través del sistema de transcripción y comunicación.

La estética "low-poly" contribuye a una experiencia visual agradable y optimizada en términos de rendimiento, ya que los modelos simplificados requieren menos recursos de procesamiento. Además, al utilizar jugadores falsos en el entorno (ya que este proyecto en realidad es una simulación de interacción real), se proporciona una sensación de presencia y realismo, incluso si no hay interacciones directas con ellos.

Capítulo 6

Implementación y pruebas

6.1. Implementación

6.2. Pruebas funcionales

6.3. Pruebas de usabilidad

6. Implementación y pruebas

6.1. Implementación

En este apartado se hará un resumen de las tecnologías usadas, una descripción detallada de los escenarios creados en Unity y se explicará la implementación de los scripts o acciones más relevantes.

Tecnologías

En el proyecto se han utilizado varias tecnologías que han contribuido a su desarrollo y funcionalidad. A continuación, se enumeran y se explica brevemente cada una de ellas:

- Unity: Unity es un motor de desarrollo de videojuegos y aplicaciones interactivas. Se ha utilizado como la plataforma principal para la creación del proyecto, permitiendo la creación de escenas, modelos, animaciones y lógica de juego. Unity es compatible con múltiples plataformas, lo que facilita la portabilidad del proyecto.
- Realidad Virtual (VR): Se ha empleado la tecnología de realidad virtual para ofrecer una experiencia inmersiva a los usuarios. La realidad virtual permite sumergirse en un entorno simulado e interactuar con él utilizando dispositivos como las gafas de VR, tecnología que Unity ofrece con el XR Toolkit y establece acciones predefinidas como el movimiento mediante teletransportes.
- Realidad Aumentada (AR): La realidad aumentada se ha utilizado para combinar elementos virtuales con el entorno real, superponiendo información digital sobre el mundo físico. Esta tecnología permite interactuar con objetos virtuales mientras se mantiene una conexión con el entorno real, este apartado se atribuye en su totalidad a las funcionalidades de Vuforia y XR Toolkit.
- Speech-to-Text: La tecnología de reconocimiento de voz ha sido empleada para convertir el habla en texto. Esto permite transcribir las palabras habladas por los usuarios y utilizarlas en la aplicación, esta función ha sido proporcionada por j1mmyto9 de manera solidaria para uso libre a través de GitHub (Copyright (c) 2020 Tô Ngọc Khánh).
- Android: El proyecto ha sido desarrollado para plataformas móviles como Android. Se han utilizado las características y herramientas específicas dentro de las funciones de Speech-to-Text.
- C#: El lenguaje de programación C# se ha utilizado para desarrollar scripts y comportamientos interactivos en Unity. C# es un lenguaje potente y versátil que proporciona las herramientas necesarias para implementar la lógica y la interacción en el proyecto.
- Escenas y modelos 3D: Se han creado y utilizado escenas y modelos 3D para construir el entorno virtual del proyecto. Las escenas representan los diferentes

espacios virtuales en los que los usuarios pueden interactuar, mientras que los modelos 3D son objetos y personajes virtuales que se integran en el entorno.

Escena Realidad Virtual

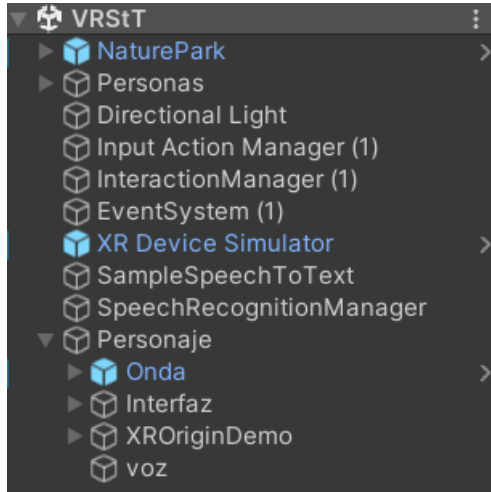


Ilustración 8: Elementos de la escena (VR)

En la escena de Unity podemos encontrar diferentes elementos:

- *NaturePark*: El conjunto de “prefabs” y objetos 3D que componen el escenario del proyecto.
- *Personas*: Conjunto de jugadores falsos que se encuentran distribuidos por la escena.
- *Directional Light*: Luz direccional de la escena (la cual tiene una tonalidad azul).
- *Input Action Manager*: Asocia la entrada de comandos de los mandos y las Oculus de manera que Unity los reconozca.
- *Interaction Manager*: Asocia las interacciones a los nuevos inputs.
- *Event system*: En este objeto se añade un módulo que permite la interacción con los GameObjects de tipo “UI”, en este caso la slider y el botón.
- *XR Device Simulator*: Este objeto permite hacer pruebas con el teclado en vez de conectar las Oculus, es necesario desactivarlo para que funcione el proyecto en Oculus, pero resulta útil en las primeras fases de desarrollo.
- *SampleSpeechToText*: Objeto con el módulo que permite la transcripción.
- *SpeechRecognitionManager*: Este objeto suele usarse para realizar acciones mediante comandos de voz, en este caso se utiliza solo para rescatar el resultado de la transcripción ya que está suscrito al evento “onResultCallback” (se explicará más adelante).
- *Personaje*: Este objeto es el conjunto de elementos que se desplazan con el personaje y por comodidad el objeto “voz” que es el que se encarga de dar las diferentes instrucciones en el proceso de transcripción según los eventos que ocurren en la simulación de interacción.

Escena Realidad Aumentada

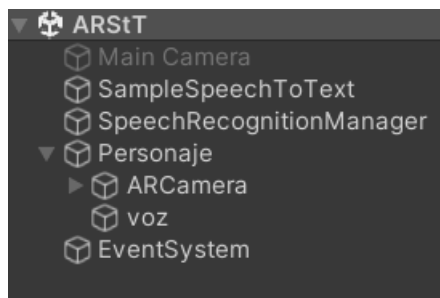


Ilustración 9: Elementos de la escena (AR)

En la escena de Unity podemos encontrar diferentes elementos:

- *SampleSpeechToText*: Objeto con el módulo que permite la transcripción.
- *SpeechRecognitionManager*: Este objeto suele usarse para realizar acciones mediante comandos de voz, en este caso se utiliza solo para rescatar el resultado de la transcripción ya que está suscrito al evento “onResultCallback”.
- *Event system*: En este objeto permite la interacción con la interfaz.
- *Personaje*: Este objeto contiene la cámara AR que muestra la imagen de la cámara del dispositivo y voz, que realiza las acciones de relacionadas con el speechtexter y lo relaciona con la interfaz, que es un “hijo” de la cámara.

Pulsar el botón de encender/apagar

Al pulsar el botón de encender o apagar el sistema comprueba el estado actual en el script “iniVoz”, si este define el estado (estados mostrados en el capítulo 4, análisis, diagrama 13) actual como “Apagado” este pasará a constar como encendido, entrando a la espera de un evento de colisión con el “collider” que llamaría a la función “OnTriggerEnter”.

En el caso de estar en el estado “Encendido” o “Escuchando” este pasará a “Apagado” y cesará su funcionamiento.



Ilustración 10: Eventos del botón

Contacto con jugadores

Como se ha mencionado con anterioridad al entrar un jugador en el área de escucha del usuario se crea un evento que llama a la función “OnTriggerEnter” en el script “iniVoz”, este cambia el estado a “Escuchando” y guarda el GameObject del personaje y le asigna un id. (Al ser una simulación de un caso real online solo existe una entrada de audio por lo que el hablante se elige y no se detecta).

Por otro lado, si el jugador sale de la zona se llama a la función “OnTriggerExit” y se elimina la información del jugador (y se le desactiva la marca de hablante si la tenía activa). Si este era el último jugador en la zona el estado volverá de “Escuchando” a “Encendido”.

Cambiar radio y efecto de onda

El efecto de onda reside en un script que escala esta onda en un bucle infinito, siempre que el sistema esté encendido. Cada vez que el radio es cambiado el script cambia su variable de “radio” para ajustarse siempre al tamaño real.

Por otro lado, el cambio de radio se realiza desde la slider, cada vez que se interactúa con ella se llama a la función “cambiarRadio” que pasa de manera dinámica el valor de la slider (entre 1 y 9) y escala el “collider” encargado de la interacción.

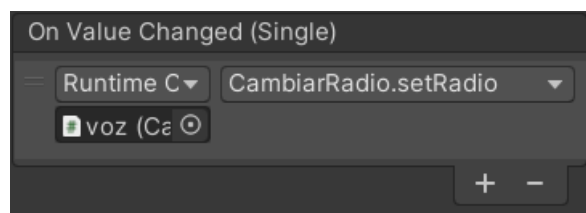


Ilustración 11: Eventos de la slider

Análisis del script de SpeechToText

Permite la integración de la funcionalidad de reconocimiento de voz (Speech-to-Text) en la aplicación. Este script proporciona métodos para configurar y controlar la grabación de audio, así como para recibir los resultados del reconocimiento de voz.

En la parte de la inicialización, se implementa el patrón Singleton para asegurarse de que solo haya una instancia de este script en todo el proyecto. Además, se define una variable booleana "isShowPopupAndroid" que determina si se debe mostrar o no una ventana emergente en dispositivos Android durante la grabación de voz. El script proporciona métodos como "Setting" para configurar el idioma del reconocimiento de voz, "StartRecording" para iniciar la grabación de audio y "StopRecording" para detener la grabación.

También se define un evento "onResultCallback" que se dispara cuando se obtienen los resultados del reconocimiento de voz. Este evento permite que otros scripts o componentes se suscriban y reciban los resultados del reconocimiento.

Posicionar interfaz

Para que la interfaz siempre se encuentre a disposición del usuario y que el sistema sea intuitivo esta interfaz hace uso de un script que comprueba en cada interacción el vector forward del usuario y lo coloca en esa posición (esta función no afecta a su altura de manera que la interfaz nunca tape la visión del usuario).



Ilustración 13: Sistema encendido sin hablante en el campo de visión



Ilustración 12: Sistema con el hablante en el campo de visión

Simulador de hablante (marca)

Justo bajo el botón de apagado hay un botón invisible que permite cambiar el hablante de manera que pueda ser representada de manera real una interacción en grupo con el usuario, este botón llama a la función `cambiarHablante()` y cambia la marca de posición entre los personajes que se encuentran dentro del radio en orden de entrada, siendo el último

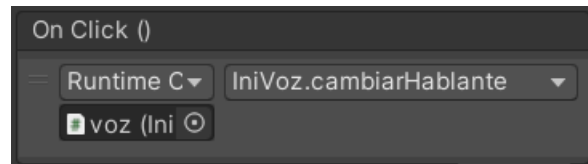


Ilustración 14: Botón de cambiar hablante

Señalar hablante fuera del rango de visión

Para que el usuario pueda encontrar con facilidad al hablante es necesario una indicación que muestre que el usuario se encuentra en esa situación y donde puede encontrar al hablante, para cumplir esto se ha generado un canvas en la cámara, cada marca tiene el script “No Mirando”, este script comprueba siempre que la marca esté activa si esta se ve, si no es así muestra una flecha apuntando hacia la dirección en la que se encuentra el hablante (el script tiene un regulador para la sensibilidad). Para conseguir la dirección se comprueba el ángulo que forma el vector “derecha” del usuario con el vector director que une al usuario y la marca y de igual forma con el vector “izquierda”, de esta manera se puede comprobar el ángulo más pequeño y comprobar el sentido de rotación idóneo.

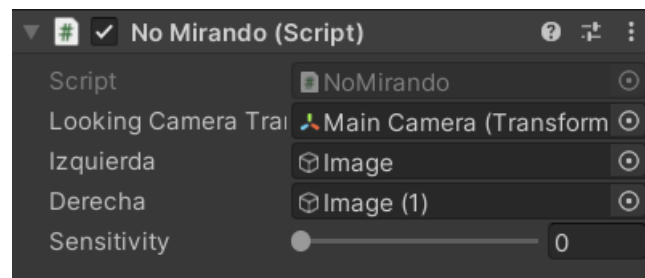


Ilustración 15: Script de "No Mirando"

6.2. Pruebas funcionales

El objetivo principal de las pruebas funcionales es asegurar que el sistema cumpla con los requisitos establecidos y que las funcionalidades principales se ejecuten de manera correcta. Estas pruebas se basan en los casos de uso definidos y también en otros procedimientos relevantes de la aplicación que pueden no haber sido abordados en secciones anteriores.

Realidad Virtual

Encender sistema



Ilustración 16: Sistema encendido sin jugadores en la zona de actuación

La función de encender el sistema se considera exitosa porque cumple con los siguientes criterios:

- Activación sin errores: La función de encendido del sistema se ejecuta sin generar errores o excepciones inesperadas.
- Comportamiento esperado: Al activar el sistema, se verifica que todas las funcionalidades relacionadas se inicialicen correctamente y estén listas para su uso.
- Indicadores visuales: Se observa visualmente que el sistema ha sido encendido, mediante el cambio de texto en la interfaz gráfica y la aparición de la onda circular.

Apagar sistema



Ilustración 17: Sistema apagado (VR)

La función de apagar el sistema se considera exitosa porque cumple con los siguientes criterios:

- Apagado sin errores: La función de apagado del sistema se ejecuta sin generar errores o excepciones inesperadas
- Indicadores visuales: Se observa visualmente que el sistema ha sido apagado, mediante el cambio de texto en la interfaz gráfica y la desaparición de la onda circular.

Cambiar radio



Ilustración 18: Sistema encendido con el hablante cambiado

La función de cambiar el radio se considera exitosa porque cumple con los siguientes criterios:

- Cambio efectivo: Al realizar el cambio en el radio, se verifica que la modificación surte efecto inmediatamente y se refleje en el comportamiento del sistema. Al aumentar o disminuir el valor, la zona de detección se ajusta en consecuencia, se puede comprobar gracias al cambio de estado del sistema comparando la ilustración 12 y la ilustración 14.
- Retroalimentación visual: Al cambiar el radio, se proporciona una retroalimentación visual para confirmar la acción realizada. Tanto en la interfaz gráfica, como en la onda que se expande bajo los pies del usuario.
- Persistencia del cambio: Se verifica que el cambio en el radio se mantiene incluso después de reiniciar el sistema.



Ilustración 19: Sistema encendido hablante marcado

Hablar



Ilustración 20: Sistema transcribiendo voz de supuesto jugador

La función de hablar, que se encarga de la transcripción del habla, se considera exitosa porque cumple con los siguientes criterios:

- Activación y detección del habla: La función de hablar se activa correctamente cuando se detecta la voz del usuario.

- Transcripción precisa: Durante la función de hablar, el sistema transcribe con precisión y en tiempo real las palabras habladas por el usuario.
- Retroalimentación visual: Durante el proceso de transcripción, se proporciona una retroalimentación visual al usuario para indicar que el sistema está escuchando.
- Marca al hablante: El jugador del que se transcribe la voz es marcado de manera correcta y eficaz.

Entrar en la zona de actuación

La función de entrar en la zona de actuación se considera exitosa porque cumple con los siguientes criterios:

- Cambio de estado: Cuando un jugador entra en la zona de actuación el estado cambia y el speechtexter se activa y espera una entrada de audio que transcribir, se puede observar comparando la ilustración 16 y 14.

Salir de la zona de actuación



Ilustración 21: Sistema al quedarse sin jugadores en la zona de actuación

La función de salir de la zona de actuación se considera exitosa porque cumple con los siguientes criterios:

- Cambio de estado: Cuando todos los jugadores abandonan la zona de actuación el estado cambia y el speechtexter deja de esperar un audio que transcribir, se puede observar comparando la ilustración 14 y 16.

Adaptación del tamaño de texto



Ilustración 22: Sistema transcribiendo una frase muy extensa

La función de "Adaptación del tamaño de texto" se refiere a la capacidad del sistema de ajustar dinámicamente el tamaño del texto según la longitud de la frase, se puede comprobar en la ilustración 17 que funciona de manera adecuada (esta función es propia de Unity y configurable de manera sencilla desde el módulo del cuadro de texto).

Realidad Aumentada

Encender sistema

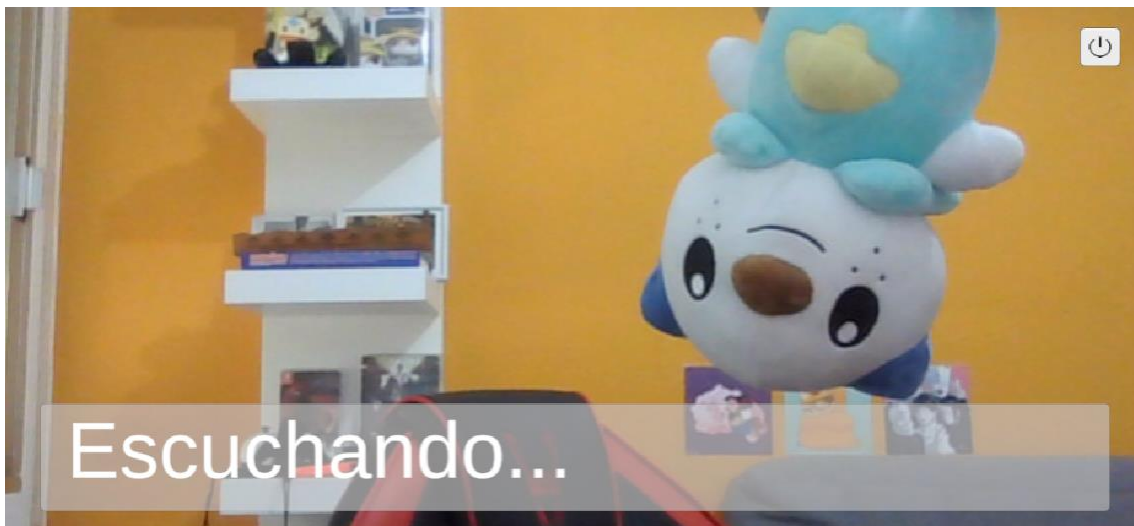


Ilustración 23: Sistema encendido (AR)

La función de encender el sistema se considera exitosa porque cumple con los siguientes criterios:

- Activación sin errores: La función de encendido del sistema se ejecuta sin generar errores o excepciones inesperadas.

- Comportamiento esperado: Al activar el sistema, se verifica que todas las funcionalidades relacionadas se inicialicen correctamente y estén listas para su uso.
- Indicadores visuales: Se observa visualmente que el sistema ha sido encendido, mediante el cambio de texto en la interfaz gráfica.

Apagar sistema

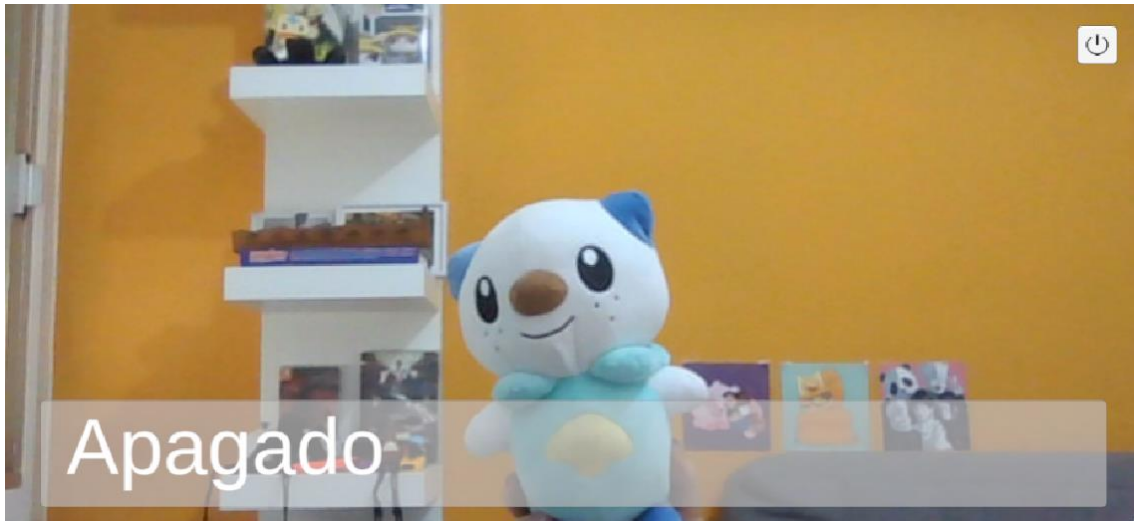


Ilustración 24: Sistema apagado (AR)

La función de apagar el sistema se considera exitosa porque cumple con los siguientes criterios:

- Apagado sin errores: La función de apagado del sistema se ejecuta sin generar errores o excepciones inesperadas.
- Indicadores visuales: Se observa visualmente que el sistema ha sido apagado, mediante el cambio de texto en la interfaz gráfica.

Hablar

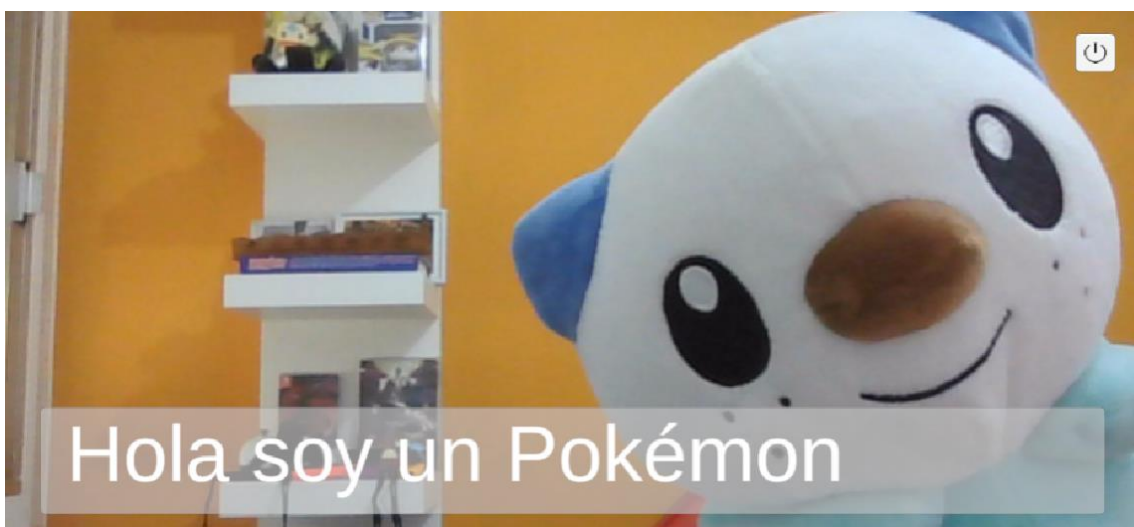


Ilustración 25: Sistema transcribiendo voz de hablante (AR)

La función de hablar, que se encarga de la transcripción del habla, se considera exitosa porque cumple con los siguientes criterios:

- Activación y detección del habla: La función de hablar se activa correctamente cuando se detecta la voz del usuario.
- Transcripción precisa: Durante la función de hablar, el sistema transcribe con precisión y en tiempo real las palabras habladas por el usuario.

6.3. Pruebas de usabilidad

Estas pruebas implican la selección de un grupo de usuarios que representan una variedad de perfiles y niveles de experiencia en aplicaciones VR y móviles. El objetivo principal es observar y analizar las diferentes dificultades y desafíos que pueden surgir durante la interacción con la aplicación.

En el caso específico de este proyecto, se llevó a cabo un estudio de usabilidad con la participación de 5 usuarios. Entre ellos, se incluyeron personas con diferentes niveles de experiencia en el uso de esta tecnología. Dos de los participantes tienen una experiencia limitada o nula en VR, mientras que otros tres tienen una experiencia promedio con el uso la realidad virtual. Además, se involucró otro usuario con un nivel avanzado de experiencia, caracterizado por un uso bastante constante de las gafas Oculus.

Estos perfiles de usuarios representan una muestra diversa que permite obtener información valiosa sobre los desafíos y obstáculos que pueden surgir durante la interacción con la aplicación. Al observar las experiencias de estos usuarios, identificaremos puntos de mejora en términos de usabilidad, fluidez de la interfaz, y cualquier otra área que pueda afectar la experiencia general del usuario.

El objetivo de las pruebas de usabilidad que puntúen (del 0 al 5) la experiencia para realizar diferentes tareas teniendo en cuenta lo intuitivo y sencillo que es, para optimizar la aplicación y garantizar una experiencia agradable y efectiva para todos los usuarios, independientemente de su nivel de experiencia previa. Con base en los resultados de estas pruebas, se pueden realizar ajustes y mejoras en la interfaz de usuario, el flujo de navegación y otros aspectos relevantes para lograr un producto final que satisfaga las necesidades y expectativas de los usuarios.

Realidad Virtual

Tarea	Exp. baja 1	Exp. baja 2	Exp. media 1	Exp. media 2	Exp. alta	Media
Encender sistema	4	4	5	5	5	4,6
Apagar sistema	5	5	5	5	5	5
Cambiar radio	2	3	4	5	5	3,8
Hablar	4	5	5	5	5	4,8

Tabla 33: Pruebas de usabilidad (VR)

Al hablar con los usuarios se ha determinado que las mayores dificultades al interactuar con la interfaz no son causadas por un mal diseño de interfaz si no las mecánicas propias de realidad virtual para interactuar con este (se interactúa con un rayo que sale del mando derecho) por lo que quien tiene mayor experiencia con esta tecnología no ha percibido ningún punto negativo.

Realidad Aumentada

Tarea	Exp. baja 1	Exp. baja 2	Exp. media 1	Exp. media 2	Exp. alta	Media
Encender sistema	5	5	5	5	5	5
Apagar sistema	5	5	5	5	5	5
Hablar	4	5	5	5	5	4,8

Tabla 34: Pruebas de usabilidad (AR)

La interacción con un teléfono móvil resulta extremadamente fácil para cualquier usuario, cabe destacar que esta aplicación solo consta con un elemento interactuable mediante la pantalla, un botón, por lo que es difícil que lleve a error.

Capítulo 7

Conclusiones

7.1. Revisión de costes

7.2. Conclusiones

7.3. Trabajo futuro

7. Conclusiones

7.1 Revisión de costes

Dado que este proyecto ha sido desarrollado únicamente por mí, no resulta relevante analizar los costos finales de personal o material. Sin embargo, es importante destacar que el proyecto ha experimentado algunos retrasos significativos en cuanto a la duración temporal. Se inició un mes más tarde de lo previsto debido a circunstancias externas, y también se tuvieron que enfrentar interrupciones ocasionadas por motivos de salud y una mudanza. En la siguiente tabla se presentan los tiempos finales aproximados, seguidos de un diagrama que muestra la realidad de la línea temporal del proyecto.

Aunque estos retrasos han impactado en la planificación inicial, se ha trabajado diligentemente para minimizar su impacto en la calidad y funcionalidad del proyecto. A pesar de los desafíos, se ha logrado cumplir con los objetivos establecidos y un producto satisfactorio.

ID	TAREAS	t
1	Introducción	19
1.1	Introducción	11
1.1.1	<i>Aumento del uso de XR</i>	6
1.1.2	<i>Inclusión y accesibilidad</i>	5
1.2	Motivación	2
1.3	Objetivos	5
1.4	Organización de la memoria	1
2	Estado del arte	20
2.1	Análisis de aplicaciones similares	9
2.2	Tecnologías	11
2.2.1	<i>Unity</i>	7
2.2.2	<i>Vuforia</i>	4
3	Plan de trabajo	59
3.1	Requisitos	20
3.2	Especificaciones	12
3.3	Costes	16
3.4	Riesgos	8
3.5	Viabilidad	8
4	Análisis	40
5	Diseño	35
6	Implementación y pruebas	444
6.1	Implementación	410
6.2	Pruebas funcionales	23
6.4	Pruebas de usabilidad	11
7	Conclusiones	26
7.1	Revisión de costes	10
7.2	Conclusiones	8
7.3	Trabajo futuro	8
	TOTAL	648

Tabla 35: Tiempo real de las tareas

En el seguimiento del proyecto, se ha utilizado la codificación de tareas en color verde para aquellas que se completaron más rápidamente de lo esperado, mientras que se ha empleado el color rojo para resaltar las tareas que experimentaron retrasos respecto a la planificación inicial. Es importante destacar que el "ahorro de tiempo" en el estado del arte se debe a la falta total de referencias, lo cual afectó considerablemente las fases de concertación de requisitos y análisis.

Afortunadamente, durante la etapa de implementación del proyecto, se pudo aprovechar la información recopilada en las fases anteriores y aprovechar las potentes herramientas ofrecidas por Unity, así como el apoyo de la comunidad de usuarios. Esto permitió agilizar significativamente el desarrollo del proyecto, especialmente en lo que respecta a los métodos de Speech-to-Text.

A pesar de que el tiempo total de las tareas se haya visto reducido debido a diversas interrupciones y a la imposibilidad de avanzar durante los lunes y martes debido a las prácticas de empresa, el período de tiempo en el que se ha desarrollado el proyecto ha sido más extenso, finalizando el día **6 de Julio de 2023**.

Estas interrupciones y limitaciones temporales han impactado en la duración global del proyecto, prolongándolo más allá de la planificación inicial. Sin embargo, se han tomado medidas para optimizar el tiempo disponible y avanzar en la medida de lo posible, lo cual ha permitido mantener un progreso constante a pesar de los obstáculos.

7.2 Conclusiones

En conclusión, el proyecto ha sido un desafío emocionante y gratificante que ha permitido desarrollar una aplicación de transcripción de voz en entornos de realidad virtual y realidad aumentada. A lo largo del proceso, se han abordado diversos aspectos técnicos y se ha aplicado un enfoque centrado en el usuario para garantizar una experiencia óptima.

A pesar de los desafíos y obstáculos encontrados, se ha logrado superarlos mediante una planificación cuidadosa, una gestión efectiva del tiempo y la búsqueda de soluciones creativas. El proyecto ha demostrado la importancia de la adaptabilidad y la flexibilidad en el desarrollo de software, especialmente cuando se trabaja en entornos emergentes como la realidad virtual y aumentada.

El resultado final es una aplicación funcional y efectiva que cumple con los objetivos establecidos. Se ha logrado implementar la transcripción de voz de manera precisa y se ha diseñado una interfaz intuitiva que mejora la comunicación entre los usuarios. Además, se han aplicado técnicas de diseño y usabilidad para garantizar una experiencia agradable y accesible.

Sin embargo, también se han identificado áreas de mejora y aprendizaje a lo largo del proyecto. La gestión del tiempo y la planificación inicial podrían haberse optimizado para evitar algunos retrasos y contratiempos. Asimismo, la investigación y el análisis previos podrían haberse ampliado para obtener una comprensión más completa de las tecnologías y enfoques existentes.

En general, el proyecto ha sido un éxito y ha brindado una base sólida para futuros desarrollos en el ámbito de la transcripción de voz en entornos de realidad virtual y realidad aumentada. Se han adquirido habilidades técnicas y de gestión de proyectos, y se ha demostrado la capacidad de adaptarse a los desafíos en un entorno en constante evolución.

El trabajo realizado en este proyecto sienta las bases para continuar explorando el potencial de la transcripción de voz en la comunicación inmersiva y puede servir como punto de partida para futuras investigaciones y aplicaciones en este campo en constante crecimiento.

7.3 Trabajo Futuro.

A pesar de haber alcanzado los objetivos principales del proyecto, existen diversas oportunidades de mejora y áreas de trabajo futuro que podrían expandir y enriquecer la aplicación de transcripción de voz en entornos de realidad virtual y realidad aumentada. A continuación, se detallan algunas de las posibles líneas de desarrollo y mejoras para futuras etapas del proyecto:

Creación de un modelo real con interacción de personas en línea

Actualmente, el proyecto se ha centrado en la simulación de interacciones entre jugadores virtuales. Un paso importante en el trabajo futuro sería permitir la interacción en tiempo real con personas reales, lo que requeriría desarrollar un sistema de comunicación en línea y establecer una conexión fluida entre los usuarios.

Implementación de selección de idioma

Para aumentar la versatilidad y el alcance de la aplicación, sería recomendable incluir una función que permita a los usuarios seleccionar el idioma al que desean que se realice la transcripción. Esto permitiría que la herramienta se usará en países con idiomas variados. Actualmente cambiar el idioma es sencillo desde una variable perteneciente a un “script” pero no es una función desarrollada para que el usuario use dinámicamente.

Funcionalidad de traducción automática

Una mejora adicional sería agregar una función de traducción automática que permita a los usuarios recibir transcripciones en tiempo real en su propio idioma, incluso si el hablante se expresa en otro idioma. Esto ampliaría significativamente las capacidades de comunicación y facilitaría la interacción entre usuarios de diferentes lenguajes.

Bibliografía

- Accesibilidad e inclusión: Dos aspectos clave para las personas con discapacidad.* (2018, diciembre 3). <https://blogs.worldbank.org/es/latinamerica/accesibilidad-e-inclusi-n-dos-aspectos-clave-para-las-personas-con-discapacidad>
- BOE-A-1978-31229 Constitución Española.* (s. f.). Recuperado 14 de julio de 2023, de <https://www.boe.es/buscar/act.php?id=BOE-A-1978-31229>
- BOE-A-2022-5140 Ley 6/2022, de 31 de marzo, de modificación del Texto Refundido de la Ley General de derechos de las personas con discapacidad y de su inclusión social, aprobado por el Real Decreto Legislativo 1/2013, de 29 de noviembre, para establecer y regular la accesibilidad cognitiva y sus condiciones de exigencia y aplicación.* (s. f.). Recuperado 14 de julio de 2023, de <https://www.boe.es/buscar/act.php?id=BOE-A-2022-5140>
- Catalunya, U. O. de. (s. f.). *Metaverso: Así condicionará nuestro futuro.* UOC (Universitat Oberta de Catalunya). Recuperado 14 de julio de 2023, de <https://www.uoc.edu/portal/es/news/actualitat/2022/235-metaverso-condicionara-futuro.html>
- Conexiones digitales en el metaverso / Meta.* (s. f.). Conexiones digitales en el metaverso. Recuperado 14 de julio de 2023, de <https://about.meta.com/es/metaverse/>
- «El metaverso puede ayudar en la humanización de internet». (2022, mayo 17). *Alfa y Omega.* <https://alfayomega.es/dia-de-internet-los-expertos-hablan-de-las-ventajas-y-de-los-riesgos-del-metaverso/>
- “En diez años dejaremos de usar móviles porque el Metaverso es el futuro de la humanidad”. (s. f.). Tulankide. Recuperado 14 de julio de 2023, de <https://www.tulankide.com/es/201cen-diez-anos-dejaremos-de-usar-moviles-porque-el-metaverso-es-el-futuro-de-la-humanidad201d>

Guía para un Metaverso inclusivo: Accesibilidad y realidad virtual. (2022, mayo 6). *Vicmix Reality*. <https://vicmixreality.com/2022/05/06/guia-para-un-metaverso-inclusivo-accesibilidad-y-realidad-virtual/>

Martínez Durá, J. J. (2021a). *Tema 4. Gestión del tiempo*. Universitat de València.

Martínez Durá, J. J. (2021b). *Tema 5. Gestión de costes*. Universitat de València.

Martínez Durá, J. J. (2021c). *Tema 6. Gestión de riesgos*. Universitat de València.

Martínez Plumé, J. (2021a). *Tema 3. Planificación y casos de uso*. Universitat de València.

Martínez Plumé, J. (2021b). *Tema 4. Análisis, parte dinámica*. Universitat de València.

Microsoft Windows Text-to-Speech for Unity – The Industrious Squirrel. (2015, julio 2).
<https://chadweisschaar.com/blog/2015/07/02/microsoft-speech-for-unity/>

Montoto, O. C. (s. f.). *Accesibilidad en XR, metaverso accesible e inclusivo. Guía de accesibilidad en entornos de Realidad Virtual (VR)*. Recuperado 14 de julio de 2023, de <https://olgacarreras.blogspot.com/2022/05/accesibilidad-en-xr-metaverso-accesible.html>

Nations, U. (s. f.). *La Declaración Universal de Derechos Humanos / Naciones Unidas*.
United Nations; United Nations. Recuperado 11 de julio de 2023, de <https://www.un.org/es/about-us/universal-declaration-of-human-rights>

Speech Recognition Grammar Specification Version 1.0. (s. f.). Recuperado 14 de julio de 2023, de <https://www.w3.org/TR/speech-grammar/>

Technologies, U. (s. f.-a). *Dirección y Distancia de Un Objeto a Otro—Unity Manual*.
Recuperado 21 de julio de 2023, de <https://docs.unity3d.com/es/2019.4/Manual/DirectionDistanceFromOneObjectToAnother.html>

Technologies, U. (s. f.-b). *Rotación y Orientación en Unity—Unity Manual*. Recuperado 21 de julio de 2023, de <https://docs.unity3d.com/es/2018.4/Manual/QuaternionAndEulerRotationsInUnity.html>

Technologies, U. (s. f.-c). *Unity - Scripting API:*

Android.AndroidAssetPacks.DownloadAssetPackAsync. Recuperado 14 de julio de 2023, de

<https://docs.unity3d.com/ScriptReference/Android.AndroidAssetPacks.DownloadAssetPackAsync.html>

Technologies, U. (s. f.-d). *Unity - Scripting API: DictationRecognizer*. Recuperado 14 de julio de 2023, de

<https://docs.unity3d.com/ScriptReference/Windows.Speech.DictationRecognizer.html>

Technologies, U. (s. f.-e). *Unity - Scripting API: GrammarRecognizer*. Recuperado 14 de julio de 2023, de

<https://docs.unity3d.com/ScriptReference/Windows.Speech.GrammarRecognizer.html>

Technologies, U. (s. f.-f). *Unity - Scripting API: KeywordRecognizer*. Recuperado 14 de julio de 2023, de

<https://docs.unity3d.com/ScriptReference/Windows.Speech.KeywordRecognizer.html>

Technologies, U. (s. f.-g). *Unity - Scripting API: Transform*. Recuperado 21 de julio de 2023, de <https://docs.unity3d.com/ScriptReference/Transform.html>

To, J. (2023). *Speech And Text in Unity iOS and Unity Android* [C#].

<https://github.com/j1mmyto9/speech-and-text-unity-ios-android> (Obra original

publicada en 2016)

Unity—Scripting API: Vector2. (s. f.). Recuperado 21 de julio de 2023, de

<https://docs.unity3d.com/es/530/ScriptReference/Vector2.html>

Unity—Scripting API: Vector3. (s. f.). Recuperado 21 de julio de 2023, de

<https://docs.unity3d.com/es/530/ScriptReference/Vector3.html>

XR Interaction Toolkit Examples—Version 2.3.2. (2023). [C#]. Unity Technologies.

<https://github.com/Unity-Technologies/XR-Interaction-Toolkit-Examples> (Obra original publicada en 2019)

ysb. (s. f.). *Accesibilidad y herramientas accesibles para todos—Web ONCE* [Page].

Recuperado 14 de julio de 2023, de <https://www.once.es/dejanos-ayudarte/accesibilidad>, <https://www.once.es/dejanos-ayudarte/accesibilidad>