

**UNIVERSIDAD POLITÉCNICA SALESIANA**  
**SEDE CUENCA**

**CARRERA DE INGENIERÍA ELECTRÓNICA**

TRABAJO DE TITULACIÓN PREVIO  
A LA OBTENCIÓN DEL TÍTULO DE  
INGENIERO ELECTRÓNICO

**PROYECTO TÉCNICO:**

“DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA PARA  
LA INTERACCIÓN AVATAR - PERSONA”

**AUTOR:**

RENATO ALEJANDRO CASTRO SERRANO

**TUTOR:**

ING. PAÚL ANDRÉS CHASI PESÁNTEZ

CUENCA – ECUADOR

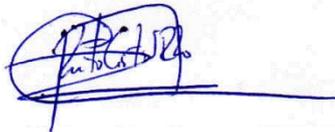
2018

## CESIÓN DE DERECHOS DE AUTOR

Yo, Renato Alejandro Castro Serrano con documento de identificación N° 0104775846 manifiesto mi voluntad y cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autor del trabajo de titulación: **“DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA PARA LA INTERACCIÓN AVATAR - PERSONA”**, mismo que ha sido desarrollado para optar por el título de: Ingeniero Electrónico, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En aplicación a lo determinado en la Ley de Propiedad Intelectual, en mi condición de autor me reservo los derechos morales de la obra antes citada. En concordancia, suscribo este documento en el momento que hago entrega del trabajo final en formato impreso y digital a la Biblioteca de la Universidad Politécnica Salesiana.

Cuenca, 3 de agosto del 2018



Renato Alejandro Castro Serrano

CI: 0104775846

AUTOR

## **CERTIFICACIÓN**

Yo declaro que bajo mi **tutoría** fue desarrollado el trabajo de titulación: “**DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA PARA LA INTERACCIÓN AVATAR - PERSONA**”, realizado por Renato Alejandro Castro Serrano, obteniendo el **Proyecto Técnico** que cumple con todos los requisitos estipulados por la Universidad Politécnica Salesiana.

**Cuenca, 3 de agosto del 2018**

A handwritten signature in blue ink, appearing to be 'P. Chasi Pesántez', written over a horizontal line.

Ing. Paúl Andrés Chasi Pesántez

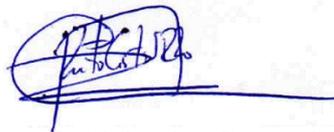
CI: 0103652095

**TUTOR DEL TRABAJO DE TITULACIÓN**

## DECLARATORIA DE RESPONSABILIDAD

Yo, Renato Alejandro Castro Serrano con número de cédula CI. 0104775846, autor del trabajo de titulación: “**DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA PARA LA INTERACCIÓN AVATAR - PERSONA**” certifico que el total contenido del **Proyecto Técnico**, es de mi exclusiva responsabilidad y autoría

Cuenca, 3 de agosto del 2018



Renato Alejandro Castro Serrano

CI: 0104775846

AUTOR

# AGRADECIMIENTOS

Quiero expresar mi más sincero agradecimiento a la Universidad Politécnica Salesiana y su personal docente, por toda la instrucción y las maravillosas experiencias que he adquirido durante mi estadía en este Centro de Estudios.

Además, agradezco de manera especial al Ing. Paúl Chasi, tutor de este trabajo de investigación, y al Grupo de Investigación en Telecomunicaciones y Telemática (GITEL), de manera especial al Ing. Fernando Guerrero y al Dr. Ing. Jack Bravo Torres, quienes me brindaron su confianza para la el desarrollo de este proyecto.

*Renato Alejandro Castro Serrano*

# **DEDICATORIAS**

Dedico este trabajo a mi familia, amigos, compañeros y en general a todas las personas que alguna vez cruzaron por mi vida y me dijeron que yo era capaz de dar mucho más. La culminación de este trabajo y de esta etapa universitaria es la mejor manera de agradecerles por esas palabras.

**Renato Alejandro Castro Serrano**

# ÍNDICE GENERAL

Agradecimientos .....	I
Dedicatorias.....	II
Índice General .....	III
Índice de Figuras .....	V
Índice de Tablas .....	V
Resumen.....	VI
Introducción .....	VII
Antecedentes .....	IX
Justificación (Importancia y Alcances).....	X
Objetivos .....	XI
Objetivo General .....	XI
Objetivos Específicos.....	XI
Capítulo 1: Estado del Arte .....	1
1.1 DESARROLLO Y EVOLUCIÓN DE LOS MÉTODOS DE INTERACCIÓN HUMANO-MAQUINA .....	1
1.2 USO DE AGENTES VIRTUALES INTELIGENTES PARA PROCESOS DE COMUNICACIÓN SOCIAL. ....	5
1.3 CREACIÓN DE AVATARS MEDIANTE SOFTWARE LIBRE .....	7
Capítulo 2: Marco Metodológico .....	10
2.1. DISEÑO DEL SISTEMA .....	10
2.2. SELECCIÓN DE AVATARS.....	11
2.3. ANIMACIÓN DEL AVATAR .....	12
2.2.1 CREAR EL ESQUELETO .....	14
2.2.2 COMANDOS BÁSICOS PARA LA EDICIÓN DEL ESQUELETO. ....	16
2.2.3 AÑADIR NOMBRES A LA ARMADURA .....	17
2.2.4 EMPARENTAR LA ARMADURA CON LA MALLA.....	18
2.2.5 CREACIÓN DE MODIFICADORES DE ARMADURA .....	19
2.3 PROGRAMACIÓN DE LA INTERACCIÓN.....	20
2.3.1 PROGRAMACIÓN EN AIML.....	22
2.3.2 RECONOCIMIENTO Y GENERACIÓN DE VOZ .....	23

2.3.3	INTERACCIÓN ENTRE BLENDER Y PYTHON .....	25
Capítulo 3:	Análisis de Resultados .....	26
3.1	PRUEBAS INTERNAS .....	26
3.2	PRUEBAS EXTERNAS .....	28
Capítulo 4:	Conclusiones y Recomendaciones .....	34
	TRABAJO FUTURO.....	36
	Referencias Bibliográficas .....	37
	Apéndices.....	40
	Apéndice A: Encuesta aplicada a los participantes que realizaron la interacción con el proyecto.....	40

# ÍNDICE DE FIGURAS

Fig. 1 Esquema general del Sistema para la interacción Avatar Persona .....	11
Fig. 2 Diseño de Avatar elegido: Cartoon Guy.....	12
Fig. 3 Etapas de la animación del Avatar en Blender .....	12
Fig. 4 Objetos desagregados en el Panel de objetos.....	13
Fig. 5 Panel de propiedades .....	13
Fig. 6 Objetos dependientes de uno principal en el panel de objetos .....	14
Fig. 7 Añadir armadura mediante el menú de adición .....	14
Fig. 8 Avatar con la armadura colocada sin ajustar .....	15
Fig. 9 Avatar con la armadura ajustada.....	15
Fig. 10 Rotación de los huesos.....	16
Fig. 11 Modificación del tamaño de un conjunto de huesos.....	17
Fig. 12 Modificación del tamaño de un conjunto de huesos.....	17
Fig. 13 a) huesos de las manos, b) huesos de los pies.....	18
Fig. 14 Armadura y malla se mueven conjuntamente.....	19
Fig. 15 Armadura y malla se mueven conjuntamente.....	19
Fig. 16 Agregar modificadores a los huesos .....	20
Fig. 17 Diagrama de flujo de la interacción total.....	21
Fig. 18 Diagrama de bloques del sistema.....	22
Fig. 19 Diagrama de reconocimiento de voz .....	24
Fig. 20 Resultados del funcionamiento general .....	29
Fig. 21 Resultados sobre la comprensión de palabras.....	31
Fig. 22 Resultados sobre comprensión de palabras .....	31
Fig. 23 Resultados sobre movimientos del Avatar .....	32
Fig. 24 Resultados sobre la interacción en general .....	33

# ÍNDICE DE TABLAS

Tabla 1: Resultados de las pruebas internas.....	27
--	----

# RESUMEN

En este proyecto técnico se propone el diseño y desarrollo de un sistema que permita la interacción entre una persona y un agente virtual (Avatar). Para desarrollar este proyecto se parte de un recorrido histórico sobre la evolución de los métodos de interacción humano – máquina, además se realiza un estudio teórico acerca de otros sistemas de interacción con agentes virtuales, inteligentes y no inteligentes, y sus diferentes aplicaciones. Seguidamente se evalúa el software de animación y el lenguaje de programación con sus respectivas librerías, para posteriormente utilizarlos en el diseño del sistema que permita una interacción básica, que posea cierto grado de inteligencia y que sea desarrollado en software libre. Se explica el procedimiento aplicado para la animación del Avatar en Blender, desde su descarga de una base de datos gratuita, hasta sus respectivas modificaciones que permitirán la animación del movimiento de todas las partes de la cara y cuerpo. Así mismo, se explica en detalle el método utilizado para realizar la interacción y para dotar al sistema de un cierto grado de inteligencia, que le permite aprender algunos datos del usuario; aspectos que fueron desarrollados en lenguaje Python con algunas librerías específicas. A continuación, se describe el procedimiento y las alternativas para la integración del software de animación y el lenguaje de programación elegidos, de forma que el sistema pueda trabajar con el registro y generación de voz, respuestas mediante inteligencia básica y con el agente virtual, el mismo que debe tener movimientos acordes a la interacción en sí. Finalmente se presentan los resultados según las pruebas realizadas, tanto por parte del investigador como por terceras personas, quienes experimentaron con el sistema y expresaron su criterio mediante una encuesta, en la que se evalúan aspectos generales y específicos del rendimiento del sistema. En estos resultados se evidencia que el proyecto desarrollado cumple con las características requeridas en el diseño.

# INTRODUCCIÓN

El presente proyecto consiste en el desarrollo de una aplicación que permite la interacción de una persona con un avatar mediante conversación en tiempo real, de acuerdo con una base de respuestas preestablecidas. El avatar seleccionado está desarrollado en el software Blender y fue obtenido de un repositorio digital donde varios diseñadores comparten sus trabajos para que estos puedan ser utilizados bajo la licencia Creative Commons. Aunque el avatar puede ser animado desde el propio Blender, generando movimientos preestablecidos a manera de un episodio de dibujos animados; en el presente trabajo se desarrolló la animación a partir de Python, con el objetivo de desarrollar la interacción en tiempo real.

Utilizando Python en conjunto con Blender, fue posible mover las articulaciones y músculos del avatar mediante scripts que a su vez se desarrollaron en función de la interacción que se buscaba con las personas. Es decir, el avatar se mueve conforme se desarrolla la interacción con la persona, utilizando movimientos preestablecidos pero que se activan de manera lógica dando la impresión de que se está desarrollado una interacción real.

La aplicación consta de un Avatar que cuenta con cierto nivel de “inteligencia” programada mediante el lenguaje AIML (Artificial Intelligence Mark-up Language). Esto permite que el avatar pueda aprender algunos datos previamente establecidos, como es el nombre de su interlocutor, lugar de residencia, intereses, etc.

Comenzamos el presente documento realizando una introducción al proyecto, describiendo de manera general el funcionamiento del mismo y el esquema capitular en que fue organizada la memoria bibliográfica.

A continuación, en el primer capítulo se desarrolla un estudio sobre el estado del arte, enfocado a la evolución de los métodos de interacción humano-máquina, donde se puede apreciar que a través del tiempo se ha desarrollado nuevas formas de comunicación con características más naturales y menos invasivas. Se analiza también el uso de avatares en proyectos de comunicación e interacción social, que previamente han sido desarrollados en diversos campos de aplicación como son: psicología, educación, y turismo. Este estudio previo permitió analizar el vasto campo de

aplicación al que se puede llevar el sistema desarrollado. Posteriormente, en este mismo capítulo, se analizan las herramientas utilizadas en el desarrollo de la aplicación como son: el software de animación Blender y el lenguaje de programación Python con el respectivo conjunto de librerías adicionales.

Luego, en el capítulo dos se realiza una explicación detallada del procedimiento realizado para llevar a cabo el desarrollo y la implementación del sistema. Comprende una descripción de las distintas etapas en las que el proyecto fue realizado, el proceso de animación de los Avatares y los métodos utilizados para realizar la interacción,

El capítulo tres comprende el detalle de las pruebas realizadas y los procesos para la verificación del funcionamiento del sistema, donde se describe su forma de aplicación y se presenta los correspondientes resultados obtenidos.

Finalmente, en el capítulo cuatro, se presenta las conclusiones a las que se ha llegado tras finalizar el desarrollo del proyecto, realizar las pruebas y obtener los resultados. Además, plantea las recomendaciones necesarias para futuras aplicaciones de este proyecto en nuevos procesos de investigación relacionados.

# ANTECEDENTES

Debido al vertiginoso crecimiento de las tecnologías de la información y comunicaciones (TIC), los métodos de interacción tradicionales han sido enriquecidos con el uso de innovadoras herramientas que nos permiten comunicarnos desde cualquier parte del mundo en cuestión de segundos. Por esta razón, las investigaciones e innovaciones que se desarrollan en la actualidad se han enfocado en aprovechar todas las ventajas posibles de las TICs en diferentes aplicaciones.

Los métodos de interacción humano – máquina (HMI) que se desarrollan con las ventajas de las TICs han dado grandes resultados, tales como aquellos que han sido capaces de identificar las emociones de grupos de personas, interactuar para brindar información, enseñar nuevos idiomas y evaluar aprendizajes, entre otros. Las TICs han permitido que sea posible una interacción cada vez más fluida y realista, con el apoyo de ramas como la Inteligencia Artificial (IA), la computación, la robótica entre otros [1],[2],[3].

En este contexto, la Universidad Politécnica Salesiana propone el proyecto AVATAR, que consiste en agentes virtuales dotados de IA que buscarán generar interacción con otros seres humanos para fines específicos. Así, el proyecto: **“Diseño e Implementación de un sistema para la interacción Avatar - Persona”**, será el punto de partida para el desarrollo de AVATAR, ya que sus siguientes etapas tomarán como base el sistema desarrollado en este proyecto. Cabe recalcar que la temática desarrollada puede considerarse como tema innovador, pese a que aborda un tema muy común como lo es la interacción humano – máquina, este no ha sido enfocado, dentro de la universidad, hacia la creación de agentes virtuales.

Este proyecto técnico toma como bases los estudios previos sobre agentes virtuales e interacción humano – máquina, realizados en diversas instituciones alrededor del mundo, y tratará de concentrar las características principales de funcionamiento para desarrollar de un sistema de interacción básica, que será el punto inicial desde donde partirá el proyecto AVATAR y que, en lo posterior, permita toda clase de mejoras en virtud de alcanzar los objetivos a lo largo de las distintas etapas del proyecto de investigación mencionado.

## **JUSTIFICACIÓN (IMPORTANCIA Y ALCANCES)**

El Grupo de Investigación en Telecomunicaciones y Telemática (GITEL) de la Universidad Politécnica Salesiana, se encuentra desarrollando el proyecto **AVATAR Autism: Virtual Agent To Agument the Relationships in Children**, el mismo que está enfocado en ser una ayuda para el tratamiento de niños con trastorno del espectro autista (TEA) enfocado en mejorar las relaciones sociales de los niños.

Este proyecto fue dividido en varias etapas, siendo la primera de ellas la de desarrollar un sistema para la interacción Avatar – Persona, mismo que es el objeto de este proyecto técnico.

El objetivo inmediato de este trabajo, dentro del proyecto AVATAR, es disponer de un sistema propio, personalizado, realizado en software libre y con capacidades de mejora, que sea capaz de interactuar mediante la voz con un ser humano, a modo de una conversación. El sistema debe presentar un Agente Virtual que realice algunos movimientos básicos (como el movimiento bucal a la hora de hablar), y que sea capaz de responder ciertas preguntas básicas realizadas por el usuario, así como aprender algunos datos del mismo, como pueden ser su nombre, lugar de estudios y otros.

No se profundiza en algoritmos de Inteligencia Artificial, ya que esta característica puede ser implementada en las siguientes etapas, lo cual servirá para otros trabajos de titulación para los alumnos de la Universidad.

# **OBJETIVOS**

## **OBJETIVO GENERAL**

- Diseñar e implementar un sistema de interacción avatar – persona.

## **OBJETIVOS ESPECÍFICOS**

- Realizar un estado del arte referente a los conceptos teóricos sobre el uso de Avatars en la interacción con personas.
- Integrar un software de animación y un lenguaje de programación que permita implementar cierto grado de inteligencia en el avatar.
- Desarrollar una interfaz que permita la interacción entre el avatar y la persona en tiempo real.

# **CAPÍTULO 1: ESTADO DEL ARTE**

En este apartado se repasa la historia de los métodos de interacción humano-máquina. En este recorrido nos encontramos con grandes cambios y avances en cuanto a los mecanismos de interacción, que a nuestros tiempos no se limitan a un teclado y una pantalla sino que existen muchas formas de comunicación que aparentan más naturalidad en la interacción.

Estos nuevos mecanismos de interacción han servido además para desarrollar diversas aplicaciones, las mismas que también tendrán su espacio en este apartado, ya que conocerlas nos permite enfocar más en el objetivo planteado para este proyecto.

## **1.1 DESARROLLO Y EVOLUCIÓN DE LOS MÉTODOS DE INTERACCIÓN HUMANO-MAQUINA**

El desarrollo tecnológico ha conseguido que la inteligencia artificial vaya más allá de lograr que las máquinas sean capaces de tomar decisiones en función de las variables de su entorno. Actualmente se busca que las máquinas puedan interactuar con nosotros de diferentes maneras. Nos encontramos frente a un escenario en el que las películas de ciencia ficción de años pasados se vuelven realidad con robots que incluso imitan sentimientos humanos.

En este contexto, los métodos de interacción hombre-máquina, se han convertido en un tema de investigación de gran interés para muchos investigadores. Podemos citar por ejemplo a Koval [4] quien define la integración hombre-máquina

(HMI) como la pérdida de fronteras entre el humano y la máquina, donde las características del uno y del otro puedan ser interpretadas en conjunto.

A lo largo del tiempo, se ha buscado que la interacción entre el ser humano y un ordenador sea fácil y agradable, permitiendo reducir en lo posible la incomodidad que pueda significarle, al mismo tiempo de aumentar la eficiencia y rendimiento de su trabajo. Desde la década de los 60 se han realizado investigaciones en este campo y, para 1980, el diseño de métodos sofisticados de HMI se había convertido en una disciplina consolidada en la rama de la computación y la ingeniería. Jaquero [5] establece que las interfaces deben cumplir con las características de:

**Adaptabilidad:** Se da cuando el usuario es quien realiza la adaptación sobre la interfaz de la máquina para ajustarse a su agrado y característica personal. Por ejemplo, cuando el usuario puede personalizar la apariencia de la imagen en la pantalla del computador.

**Adaptividad:** Se presenta cuando el sistema realiza las acciones necesarias para adaptarse al usuario. Por ejemplo, un sistema que permita la reducción de errores por parte del usuario (como un auto-corrector en un procesador de texto), lo que aumenta la eficiencia de trabajo de este usuario.

La interacción con un ordenador puede ser verbal (oral) y no verbal. Cuando hablamos de interacción verbal por medio de la voz, se busca que entre el ser humano y el computador exista una comunicación que se aproxime a una conversación natural, donde participan aspectos como la sintaxis, semántica y pragmática que se desarrollan durante el diálogo. El objetivo primordial en una interacción verbal es que exista una comprensión mutua entre los interlocutores, de tal forma que ambos puedan entender el significado de lo que el otro trata de expresar. Sin embargo, en las conversaciones entre personas, es muy frecuente que uno de los participantes cambie, inserte o elimine palabras dentro de un contexto, sin alterar el sentido de la frase. Esto se vuelve un inconveniente cuando uno de los participantes es una máquina y tiene que interpretar lo expresado por la persona con quien interactúa, siendo que las máquinas por lo general entenderán un lenguaje formal, el mismo que no siempre coincidirá con el lenguaje natural con el que se expresará el usuario. De ahí, la importancia de que un sistema para interacción humano – máquina deba tener un analizador morfológico, para encontrar la raíz de las palabras y asociarlas con sinónimos o palabras similares,

y un analizador semántico, que se encarga de revisar el orden de las palabras y encontrar el significado correcto de las frases recibidas, para disminuir las dificultades que estos aspectos puedan conllevar [6], [7], [8].

La comunicación por medio de la voz ha evolucionado con el transcurso del tiempo, ya que antes se la consideraba solo como una opción que debe utilizarse en casos específicos, como los detallados por Cohen [9]:

- Cuando los ojos y manos del usuario están ocupados,
- Cuando se tiene limitaciones en el teclado y/o monitor,
- Cuando el usuario presenta una o más discapacidades
- Cuando la pronunciación es un tema importante
- Cuando la interacción con lenguaje natural es preferida

Sin embargo, en la actualidad es muy común encontrar personas que interactúan mediante la voz con un computador, ya sea por diversión, comodidad o por encontrarse frente a aplicaciones que requieren este tipo de interacción. Incluso, el desarrollo tecnológico actual ha superado la barrera de la comunicación verbal, convergiendo hacia la comunicación no verbal, donde la máquina no requiere de instrucciones verbales sino que busca obtener la información necesaria y establecer una interacción por métodos propios y adaptativos del entorno.

En el contexto de la comunicación no verbal, desde el año 1991, se ha propuesto el tema de la computación ubicua, un modelo que plantea que la interacción entre el usuario y el computador debe ser más natural y menos intrusiva, es decir, que el computador obtenga una retroalimentación sin que el usuario descuide su actividad, sino solamente mediante el reconocimiento y análisis de las emociones del usuario. Estos sistemas, llamados sistemas con conciencia emocional, se basan en expresiones faciales para estimar la emoción que se ha generado en el usuario [10].

Investigaciones más recientes se centran en la interacción directa entre el cerebro humano y la máquina, buscando que las órdenes sean dadas directamente mediante el pensamiento del usuario, evitando así el uso de cualquier interfaz física de comunicación. Como lo detalla Minguez [11], los sistemas en desarrollo se basan en la actividad del sistema nervioso interpretado como señales eléctricas, las cuales son receptadas por un sensor, procesadas y decodificadas para entender la intención del

usuario y realizar una determinada acción de respuesta. Las investigaciones en esta rama se enfocan en mejorar los sensores y algoritmos de procesamiento de señales, teniendo como meta diseñar sistemas no invasivos que permitan esta forma de comunicación. Las aplicaciones están orientadas desde fines sociales, como manejar una silla de ruedas solo con el pensamiento, hasta fines recreativos, como es la rama de los videojuegos.

Las interfaces cerebro-máquina también se han enfocado en la sustitución sensorial, que consiste en un sistema externo de sensores que captan la información del entorno para ser procesada por el cerebro humano. Esto permite que las personas con discapacidad visual o auditiva puedan tener percepciones de su entorno mediante sensores que sustituyen a los órganos correspondientes a sus sentidos perdidos. Los estudios en esta rama se enfocan en investigar y desarrollar la capacidad del cerebro para adaptarse a la información sensorial recibida externamente, permitiendo que la información percibida forme parte de procesos cognitivos completos [12].

Otra forma de HMI se presenta en la robótica, ciencia que ha avanzado hasta otorgar inteligencia artificial a robots que pueden realizar tareas específicas y hasta conversar con seres humanos. La investigación de Thomason [3] presenta un modelo de robot repartidor, el cual recibe órdenes por voz. Además, tiene un módulo de generación de voz que le permite, además de realizar la acción, dar una retroalimentación o indicación al usuario. El sistema fue probado por varios grupos de usuarios, quienes debían interactuar en conversaciones, así como dar órdenes aleatorias y verificar su cumplimiento. El resultado fue muy positivo, ya que hubo un mayor índice de usuarios que afirmaron que el robot les entendió, frente a quienes se sintieron frustrados durante la interacción.

Ligado también a la robótica, se encuentra el trabajo realizado por Meena y Jokinen [13], el mismo que utiliza el robot NAO, programado para interactuar con el usuario mediante gestos. Los usuarios que probaron este sistema, lo evaluaron en distintos niveles: Interface, Expresividad, Responsividad, Usabilidad y rendimiento general. En cada caso debían, previamente, responder un cuestionario sobre cuales son sus expectativas acerca del sistema y, finalmente, evaluar la experiencia.

## **1.2 USO DE AGENTES VIRTUALES INTELIGENTES PARA PROCESOS DE COMUNICACIÓN SOCIAL.**

El uso de avatars con fines de comunicación ha permitido el desarrollo de una diversidad de aplicaciones en distintas ramas de la investigación.

En el campo de la psicología, Watanabe y Okubo [14] proponen un modelo de agente virtual (VA) que representa al hablante, reproduciendo características como: gestos, cabeceo, parpadeo, expresiones faciales, paralenguaje y respiración. Mediante esta aplicación, dos hablantes remotos puedan interactuar en un mismo entorno virtual de comunicación, lo que permitió analizar los aspectos psicológicos y conductuales de los interlocutores mediante el comportamiento de sus Avatar.

En el estudio realizado por Moore y Cheng [15] se plantea el uso de avatars en un entorno virtual colaborativo (CVE) como soporte para mejorar ciertos impedimentos que padecen los niños con trastorno del espectro autista. Sin embargo este proyecto quedó en vías de aplicación.

En el ámbito educativo se han desarrollado varios proyectos que involucran el uso de avatars con diferentes propósitos. Macedonia y Groher [2] desarrollaron un estudio para demostrar que el uso de avatars puede ayudar a mejorar el aprendizaje verbal, mediante la ejecución de gestos y acciones. Tras realizar diversos experimentos se encontró que los Avatar pueden convertirse en entrenadores para aprendizaje en seres humanos. Las pruebas fueron realizadas para niños y adultos con agentes virtuales sin inteligencia ni retroalimentación, por lo que se propone que al agregarles inteligencia artificial (IA) se obtendría resultados más eficaces.

Los Avatar pueden incluso ser utilizados para detectar cambios emocionales en las personas, basándose en el análisis conductual, analizando expresiones verbales, faciales y corporales. Wang y Geigel [1] diseñaron un sistema que utiliza VA para detectar las emociones de un grupo de personas mediante cámaras y micrófonos, teniendo como meta mejorar el estado de ánimo de todos mediante la música. El VA reproduce un género musical determinado y analiza los cambios en el estado de ánimo del grupo de personas para saber si este fue de su agrado, según el resultado puede decidir si debe o no cambiar de género musical. Un siguiente paso de este proyecto sería que el VA realice un diagnóstico previo de las personas para poder predecir,

mediante algoritmos de aprendizaje de máquina, el género musical que puede ser del agrado del colectivo.

Otros modelos se han planteado de comunicación entre un humano y un computador con fines investigativos. Entre los más destacados tenemos al proyecto ALICE, (Artificial Linguistic Internet Computer Entity), que consiste en un *chatbot*, es decir, un robot de conversación. Realiza una interacción con el usuario mediante entrada y salida de texto escrito, de tal forma que el usuario escribe una pregunta y el chatbot le otorga una respuesta adecuada, para ello cuenta con una gran base de datos de preguntas y respuestas preestablecidas. Con este proyecto se dio a conocer el Lenguaje de etiquetado para Inteligencia Artificial (AIML) [16].

En este mismo campo, otro proyecto de gran interés es el agente virtual MAXINE, que cuenta con un sistema de análisis emocional, mediante el análisis del lenguaje verbal y no verbal. Así mismo, puede simular expresiones emotivas mediante movimientos corporales [17].

Los proyectos ALICE y MAXINE han servido como plataformas de desarrollo para diversos proyectos. Baldassarri [18] diseñó un sistema que era capaz de, según las palabras utilizadas, reconocer la emoción de la persona y darle una respuesta apropiada. El diseño, que es bastante sencillo, fue realizado mediante la fusión de dos proyectos de IA previamente diseñados, MAXINE y ALICE, con la mejora de insertar el componente emocional.

El uso de agentes virtuales es una ayuda para una mejor interacción, desde la perspectiva del humano. Según diversos experimentos realizados, cuando el agente virtual tiene aspecto humanoide, se aumenta la familiaridad durante la interacción [14]. La condición de Avatar personalizado también ayuda a que el agente pueda adaptarse a la individualidad de cada persona, según sus propias necesidades y características. Esto es de gran utilidad, especialmente, cuando el avatar tiene por fundamento la enseñanza [15].

Los avatars creados con fines de interacción pueden tener diversidad de usos incluso con fines terapéuticos. El proyecto AVATAR, presentado por los docentes de la Universidad Politécnica Salesiana, Fernando Guerrero y Jack Bravo [19] propone el uso de agentes virtuales como soporte al tratamiento de niños con TEA, en virtud de

mejorar su capacidad de interacción social. Este tratamiento estaría vinculado al tratamiento habitual especializado y debe ser monitoreado por un terapeuta experto.

El proyecto se ha planteado realizar avatars semipersonalizados, para que los niños puedan sentirse en confianza y no tener miedo de interactuar con ellos. Entre las etapas propuestas se encuentran el diseño de Avatars que reproduzcan la acción del usuario, como si fueran marionetas virtuales, y otros que, dotados de IA, puedan proponer y evaluar actividades complementarias al tratamiento de estos niños [19],[20],[21].

El proyecto AVATAR toma como base los proyectos realizados en Invirtua [22], donde se utiliza un Avatar para realizar una comunicación indirecta con el paciente, pero detrás del Avatar se encuentra un terapeuta, en tiempo real, proponiendo actividades y evaluando el desarrollo del niño. La diferencia que plantea AVATAR es que el sistema sea autónomo, con actividades automatizadas y personalizadas, evitando así la necesidad de que el terapeuta tenga presencia real durante la interacción, de esta forma se puede aumentar la cantidad de actividades que el Avatar puede ofrecer al paciente en función de sus necesidades [19].

### **1.3 CREACIÓN DE AVATARS MEDIANTE SOFTWARE LIBRE**

El uso de herramientas de software libre es de gran importancia, ya que en la actualidad existen diversos programas que permiten su uso libre y presentan diversas herramientas y grandes ventajas, con lo que es posible la creación de aplicaciones sin incurrir en grandes gastos económicos. Es el caso del software de animación Blender y el lenguaje de programación Python, que son las herramientas que se utilizarán en el desarrollo de este proyecto.

Blender es un software de animación multiplataforma, gratuito y de código abierto, que cuenta con las principales características necesarias para animación, simulación, edición de video, diseño de juegos entre otras. Permite realizar desde pequeños proyectos hasta trabajos muy avanzados, es de uso libre y se puede utilizar para cualquier propósito gracias a que cuenta con Licencia Pública General (GPL) [23].

Además, existen repositorios digitales que cuentan con diseños realizados por profesionales y aficionados, que son publicados y compartidos bajo licencias gratuitas.

Es este el caso de la página Blend Swap, que es una comunidad que cuenta con miles de diseños en 3D en formatos compatibles con Blender, realizados por artistas y publicados bajo la licencia Creative Commons, que es el formato de licencias más extendido y utilizado para código abierto, donde los autores pueden gestionar su licencia y conservar los derechos de autor [24],[25].

Python es un lenguaje de programación de alto nivel, lo que significa que su estructura es muy fácil de entender para el programador, además de presentar la característica de *portabilidad*, que permite que un programa pueda ser utilizado en diversos ordenadores sin una mayor modificación de su estructura [26],[27].

Una de las principales ventajas de Blender es que permite configurar herramientas de modo avanzado mediante scripts de Python, lo cual le permite el trabajo conjunto de cualquier aplicación realizada en este lenguaje de programación y una animación realizada en Blender.

Mediante el uso de Blender y Python se pretende realizar un sistema básico de interacción entre un ser humano y un Avatar. Esta aplicación está proyectada para que a futuro, aprovechando las ventajas del lenguaje Python, pueda incorporar mejores funciones que permitan, además de interactuar mediante la recepción de instrucciones, generar interacción siendo el Avatar el proponente de preguntas y temas de conversación, lo cual puede servir como complemento terapéutico al tratamiento de niños con trastorno del espectro autista, entre otras múltiples aplicaciones que se pueden desarrollar.

El lenguaje Python tiene la ventaja de contar con un amplio repositorio de librerías totalmente libres que permiten el desarrollo de múltiples aplicaciones. La librería **speech recognition**, integra otras librerías de manejo de sonido y dispositivos y permite realizar reconocimiento de voz a partir de una entrada de audio. Para ello, utiliza varios sistemas, algunos de los cuales requieren librerías adicionales. El principal método que utilizan es la conexión con el sistema de reconocimiento de voz Google Cloud Speech API, que no requiere paquetes adicionales y es muy eficaz [28].

Adicionalmente, Python tiene soporte para trabajar directamente con el lenguaje AIML y todos sus componentes, el mismo que se desarrolló por el proyecto ALICE, permitiendo la interacción con una base de datos y un “bot” encargado de

aprender datos del usuario y encontrar una respuesta apropiada a una pregunta o frase. Para ello, Python cuenta con una librería de nombre **aiml** que posee todos los elementos necesarios [16],[29].

Otra ventaja del lenguaje Python es que, mediante la librería **pywin32** permite la interacción con el Win32 API, el mismo que permite trabajar con varias de las características de Windows, como son las Interfaces de usuario multilenguaje (MUI), acceso a las tarjetas de audio y a los ingresos de teclado, entre otras opciones [30],[31].

## **CAPÍTULO 2: MARCO METODOLÓGICO**

En este apartado se describe el proceso de desarrollo del sistema para la interacción Avatar – persona en tiempo real, desde su diseño hasta una descripción de las diferentes etapas que se realizaron para su desarrollo, las mismas que incluyen desde la selección del Avatar de una base de datos con licencia libre, su modificación, la programación de la interacción y el vínculo entre el lenguaje de programación y el software de animación

### **2.1. DISEÑO DEL SISTEMA**

El sistema fue diseñado para que, mediante una entrada del usuario por voz, se realice una interacción básica con un Avatar. Para ello, la voz debe ser captada, reconocida y transformada a texto.

La respuesta del Avatar se realizará en base a la pregunta o frase ingresada, con un sistema básico de IA que debe elegir la respuesta más adecuada y aprender ciertos datos como el nombre del usuario. Una vez elegida la respuesta, esta debe devolverse al usuario mediante audio y, a su vez, el Avatar debe generar un movimiento, dando la impresión de estar comunicándose como una persona.

Para realizar el sistema se optó por el uso de software libre, siendo elegido Python como lenguaje de programación y Blender como software de animación. En la figura 1 se observa un esquema general del proyecto con sus respectivos elementos y etapas de funcionamiento.

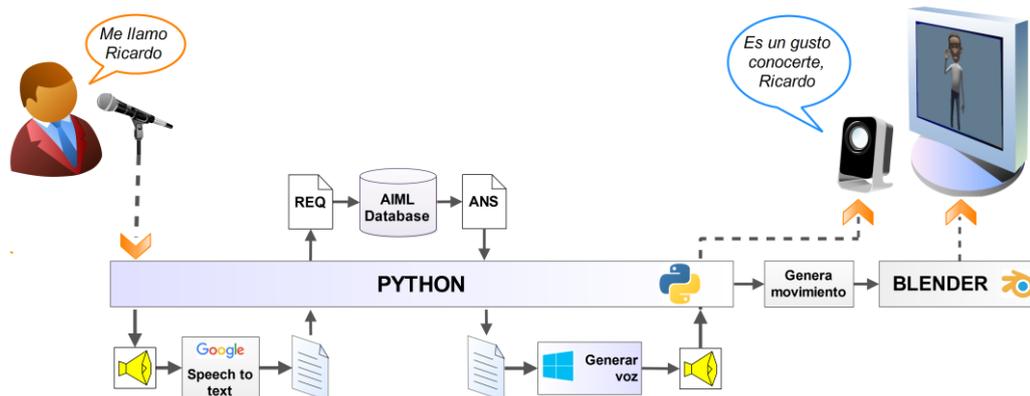


Fig. 1 Esquema general del Sistema para la interacción Avatar Persona  
**Fuente:** Autor

## 2.2. SELECCIÓN DE AVATARS

Los avatars utilizados son de uso libre, bajo la licencia Creative Commons, diseñados específicamente para Blender, Para la selección de los avatars se consideró como criterio principal que estos tengan características humanas o humanoides, con el objeto de que la interacción sea más amigable, además de poder utilizar un esqueleto básico uniforme que no requiera mayores modificaciones entre diferentes diseños. En la primera selección, cuyo único objetivo era seleccionar avatars humanoides, se recogió más de cincuenta avatars con diferentes formas de robot, animales y seres humanos. Posteriormente se realizó una segunda selección, considerando algunos detalles, tales como la presencia de cuatro o cinco dedos en sus manos, rasgos faciales explícitos, proporciones adecuadas, etc., obteniéndose finalmente quince avatars para el desarrollo del proyecto.

Una tercera y definitiva selección se realizó considerando las propiedades técnicas de los avatars, donde se analiza la factibilidad de animación, tanto para movimientos corporales como para expresiones faciales. Finalmente, se decidió trabajar con el modelo llamado *Cartoon Guy*, ya que este presenta un buen diseño y tiene todos los requerimientos que permiten su correcta animación. En la figura 2 se presenta la imagen de este avatar, con el que se va a desarrollar la interacción.

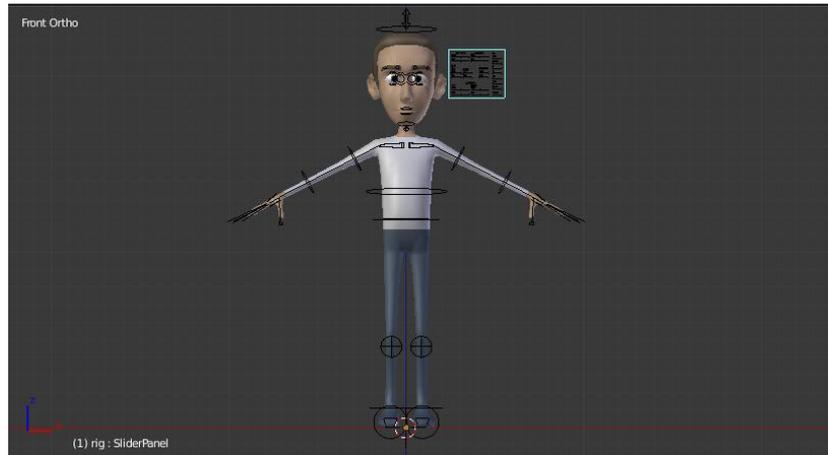


Fig. 2 Diseño de Avatar elegido: Cartoon Guy  
**Fuente:** Autor

### 2.3.ANIMACIÓN DEL AVATAR

A continuación, se explica de forma detallada todo el proceso para la generación de animación en un AVATAR. Este proceso consta de cuatro etapas, como se presenta en el esquema mostrado en la figura 3. Cada una de estas etapas se describen en este apartado.

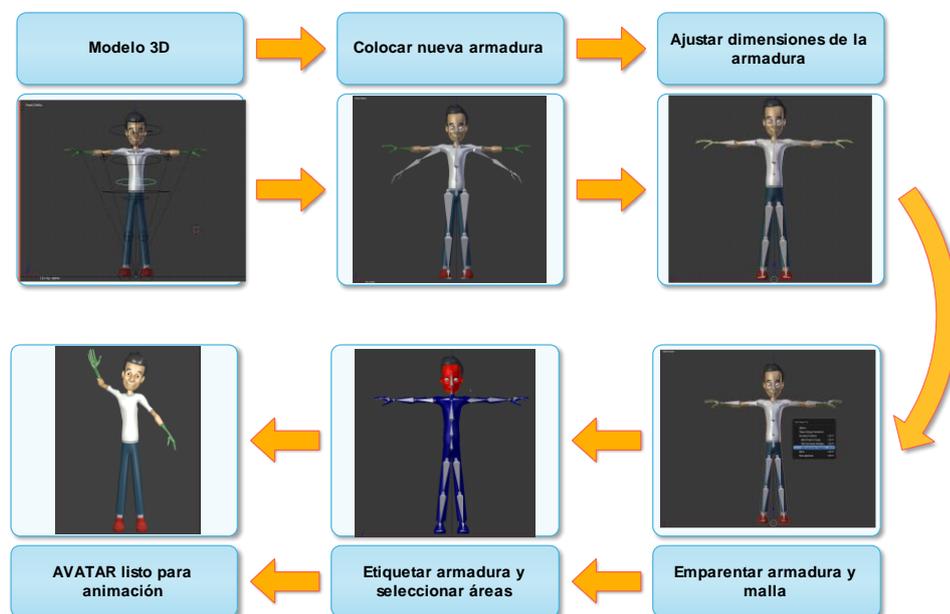


Fig. 3 Etapas de la animación del Avatar en Blender  
**Fuente:** [21]

### 2.2.1. EMPARENTAR OBJETOS

En el *Panel de Objetos*, mostrado en la figura 4, se observan los diferentes objetos que conforman el Avatar entre ellos tenemos: Cabello, Cuerpo, Dent\_Inferior, Dent\_Superior, Lengua, Ojo\_Derecho, Ojo\_Izquierdo.

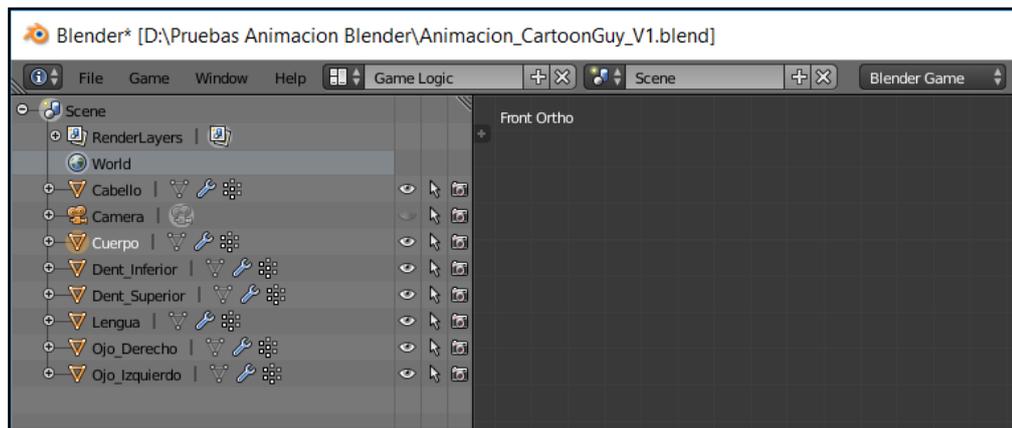


Fig. 4 Objetos desagregados en el Panel de objetos

Fuente: Autor

Lo principal en la primera parte del proceso es vincular las partes separadas al Cuerpo para tener un solo objeto de tal manera que el movimiento, la rotación y el escalado del cuerpo principal afecte a todas las partes al mismo tiempo. Para ello, se debe seleccionar cada una de las partes independientes y luego en el *Panel de Propiedades* en la pestaña de Objetos se selecciona *Parent*, y en el menú desplegado se elige la opción *Cuerpo*, tal como se expone en la figura 5.

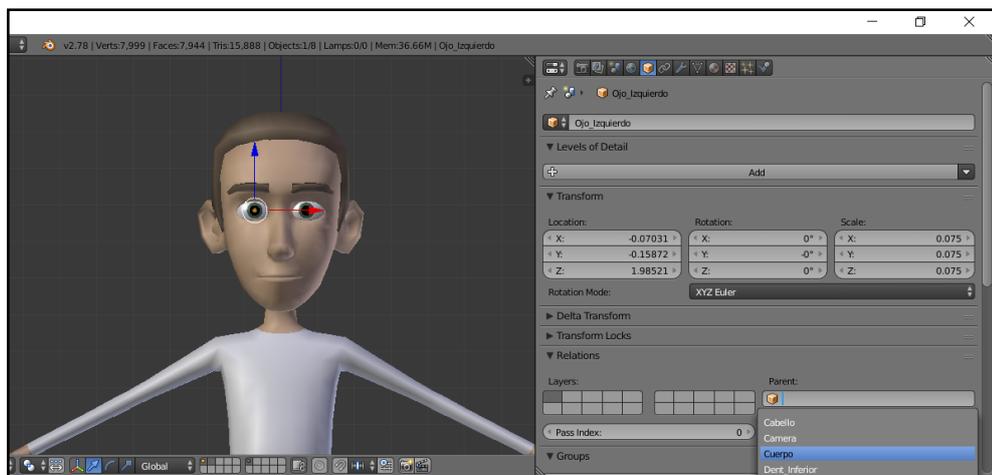
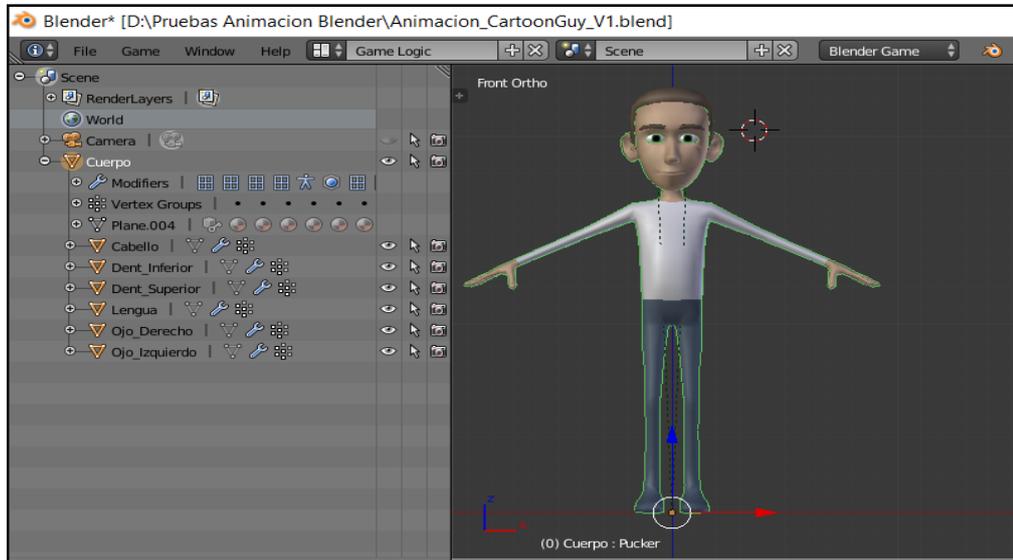


Fig. 5 Panel de propiedades

Fuente: Autor

Una vez realizado, se puede observar en el **Panel de Objetos** que las partes como Cabello, Dent\_Inferior, Dent\_Superior, etc., ahora dependen del cuerpo principal, como se evidencia en la figura 6.



**Fig. 6** Objetos dependientes de uno principal en el panel de objetos  
**Fuente:** Autor

### 2.2.1 CREAR EL ESQUELETO

Mediante el comando **Shift+A** se despliega el menú Add, para añadir objetos. En el mismo, se elige la opción **Armature -> Human(Meta-Rig)**, como se presenta en la figura 7.



**Fig. 7** Añadir armadura mediante el menú de adición  
**Fuente:** Autor

Una vez añadida la armadura (esqueleto), quedará dispuesto sobre el avatar como se presenta en la figura 8:

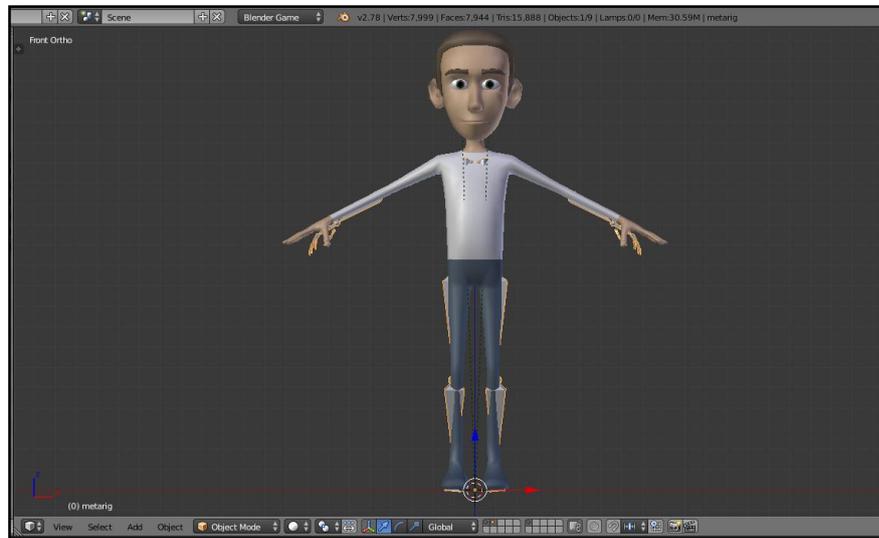


Fig. 8 Avatar con la armadura colocada sin ajustar  
**Fuente:** Autor

Desde el **Panel de Propiedades**, en la pestaña **Armature**, se activa la opción **X-Ray**, para poder visualizar el esqueleto completo y poder acoplarlo al cuerpo del avatar. Al pulsar la tecla **S** nos permite realizar los primeros ajustes del tamaño de la armadura al tamaño del cuerpo, de tal forma que el avatar quedará como se presenta en la figura 9.

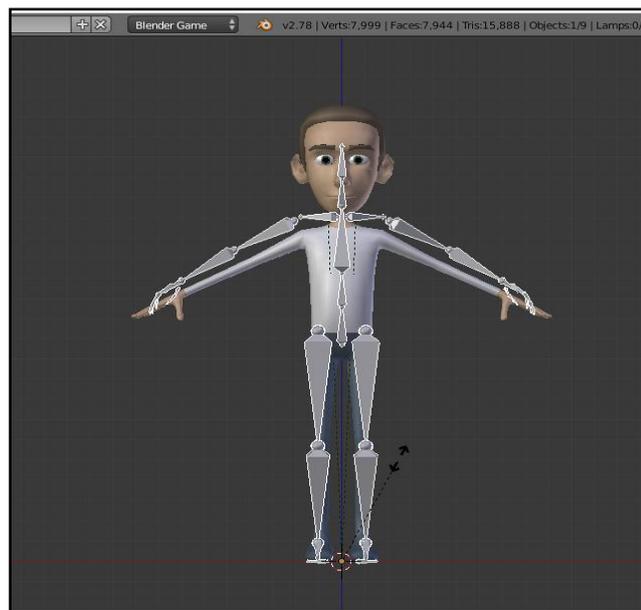


Fig. 9 Avatar con la armadura ajustada  
**Fuente:** Autor

Para acoplar el esqueleto al cuerpo con mayor detalle, se cambia la vista a **Modo Edición** y desde el Panel lateral izquierdo se activa la opción **X-Axis Mirror** para que al modificar la ubicación de los brazos y manos del lado izquierdo se refleje en el lado derecho o viceversa.

### 2.2.2 COMANDOS BÁSICOS PARA LA EDICIÓN DEL ESQUELETO

El software Blender presenta varios accesos directos por teclado a comandos para realizar modificaciones en la armadura. En primer lugar, con la tecla **C** se marca los huesos que se desea modificar y con la tecla **R** se modifica la rotación de todos los huesos seleccionados, como se observa en la figura 10.

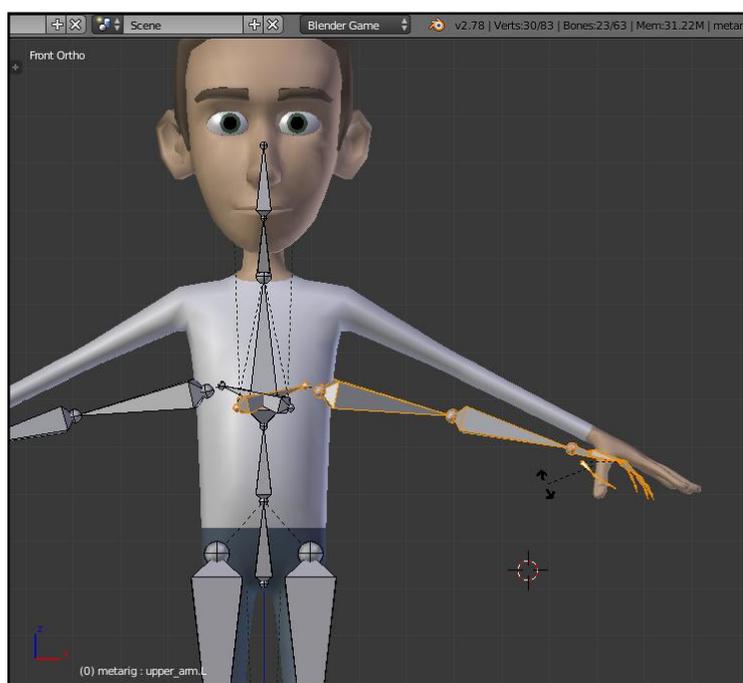


Fig. 10 Rotación de los huesos  
**Fuente:** Autor

Lo más importante ahora será acoplar el tamaño de los distintos huesos al tamaño que tenga el cuerpo. Para ello, se debe presionar la tecla **S** y con ello es posible modificar el tamaño del hueso. En caso de requerir mover varios huesos a la vez, previamente se puede utilizar la tecla **G** para seleccionar varios huesos y posteriormente cambiarles de tamaño en conjunto, como se presenta en la figura 11.



Una vez realizado todos los procesos de creación y modificación de la armadura, se deberá tener el modelo con el correcto tamaño de los huesos en todos los detalles del avatar. En la figura 13 se presenta un ejemplo de como quedan los huesos en las manos y los pies.

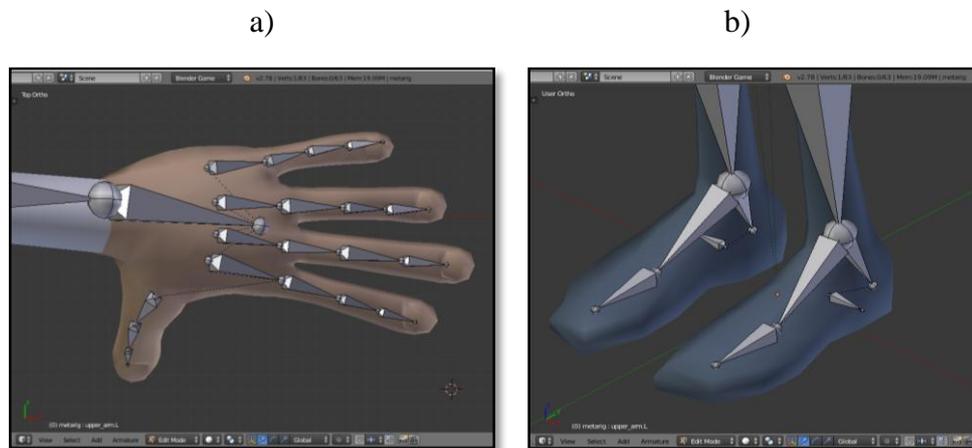


Fig. 13 a) huesos de las manos, b) huesos de los pies

Fuente: Autor

#### 2.2.4 EMPARENTAR LA ARMADURA CON LA MALLA

Cuando se realice la animación, la instrucción dada será únicamente dirigida hacia la armadura, pero la acción final debe mover completamente el avatar. Al insertar la armadura, esta es un ente separado de la malla del avatar, por lo que se requiere realizar el procedimiento para emparentar los dos objetos, de tal forma que el movimiento de los huesos afecte a todo el avatar.

Para realizar el emparentado, la ventana debe estar en el modo Objeto. Después se debe seleccionar en primer lugar la malla y, oprimiendo Shift en el teclado, la armadura, de tal manera que ambas estén seleccionadas al mismo tiempo. Luego, mediante el comando **Control+P**, se despliega un cuadro de dialogo, en el mismo se debe seleccionar la opción **Set Parent To** y posteriormente **With Automatic Weights**. Con esta opción, la malla y la armadura quedan emparentadas y pueden moverse de forma conjunta. Para comprobar que este correcto, se cambia la vista al Modo Pose, y al seleccionar y mover cualquier hueso, debe moverse conjuntamente con la malla, como se observa en la figura 14.



Fig. 14 Armadura y malla se mueven conjuntamente  
Fuente: Autor

## 2.2.5 CREACIÓN DE MODIFICADORES DE ARMADURA

Los modificadores serán los objetos que permitan realizar las diferentes acciones sobre la armadura, que se reflejan en el movimiento final del avatar. Cada uno de los huesos deberá tener un modificador con el mismo nombre que el hueso, para que la acción se realice correctamente.

Para la creación de los modificadores, se debe acudir al Panel de Propiedades, donde consta la opción Modificadores. En el menú que se despliega, se debe seleccionar la opción **Deform Armature**, como se evidencia en la figura 15.

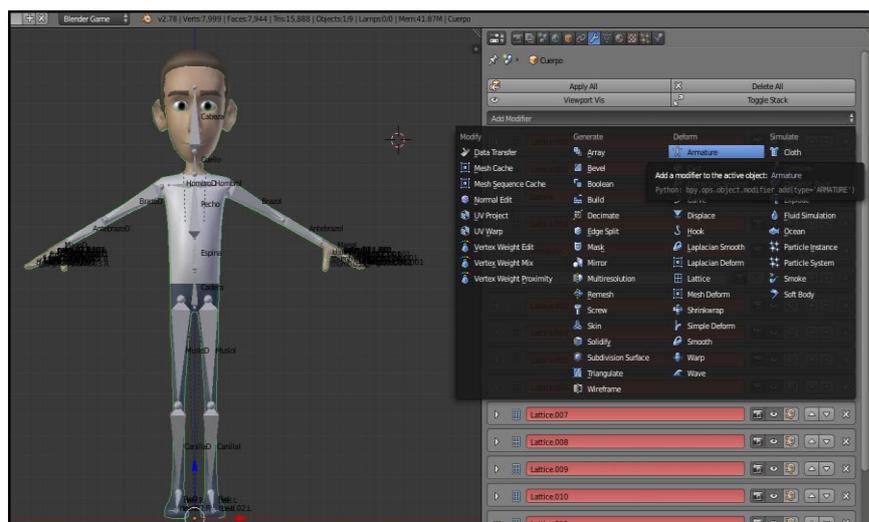
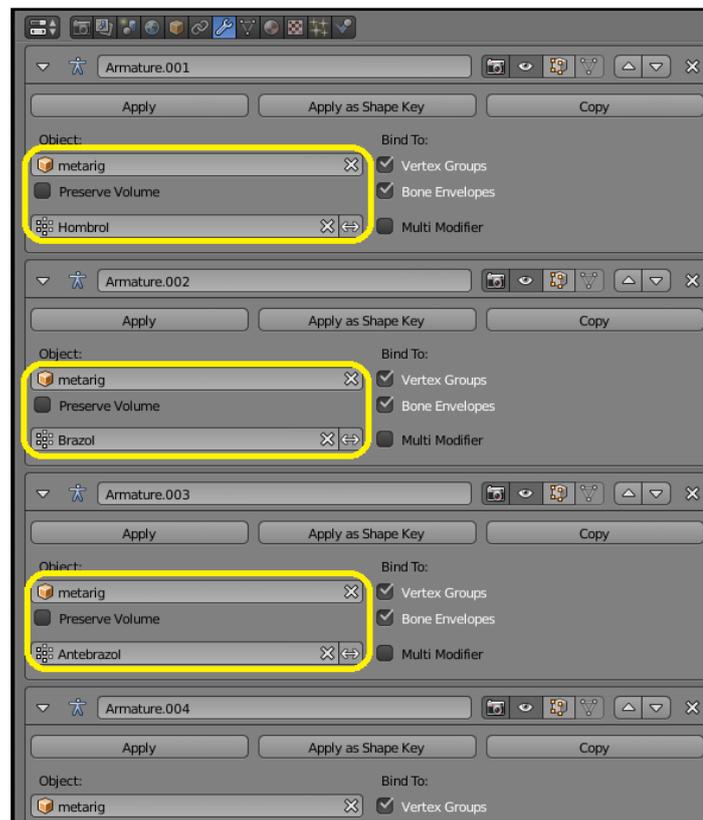


Fig. 15 Armadura y malla se mueven conjuntamente  
Fuente: Autor

Dentro del mismo Panel se visualiza un recuadro en donde se debe seleccionar el objeto (Object) y el *Grupo de Vértices*, esto permitirá que cada hueso mueva el grupo de vértices que se encuentra definido y tiene el mismo nombre que el hueso. Se debe eliminar aquellos grupos de vértices que tengan nombres diferentes y se procede a insertar varios modificadores en cada hueso, como se presenta en la figura 16.



**Fig. 16 Agregar modificadores a los huesos**  
Fuente: Autor

Una vez que la armadura cuenta con todos los modificadores necesarios, está lista para ser animada mediante comandos de Python o manualmente en el timeline de Blender.

### 2.3 PROGRAMACIÓN DE LA INTERACCIÓN

Esta etapa se la desarrolló de forma independiente al desarrollo del Avatar. El proceso total de la interacción se evidencia en el flujograma presentado en la figura 17.

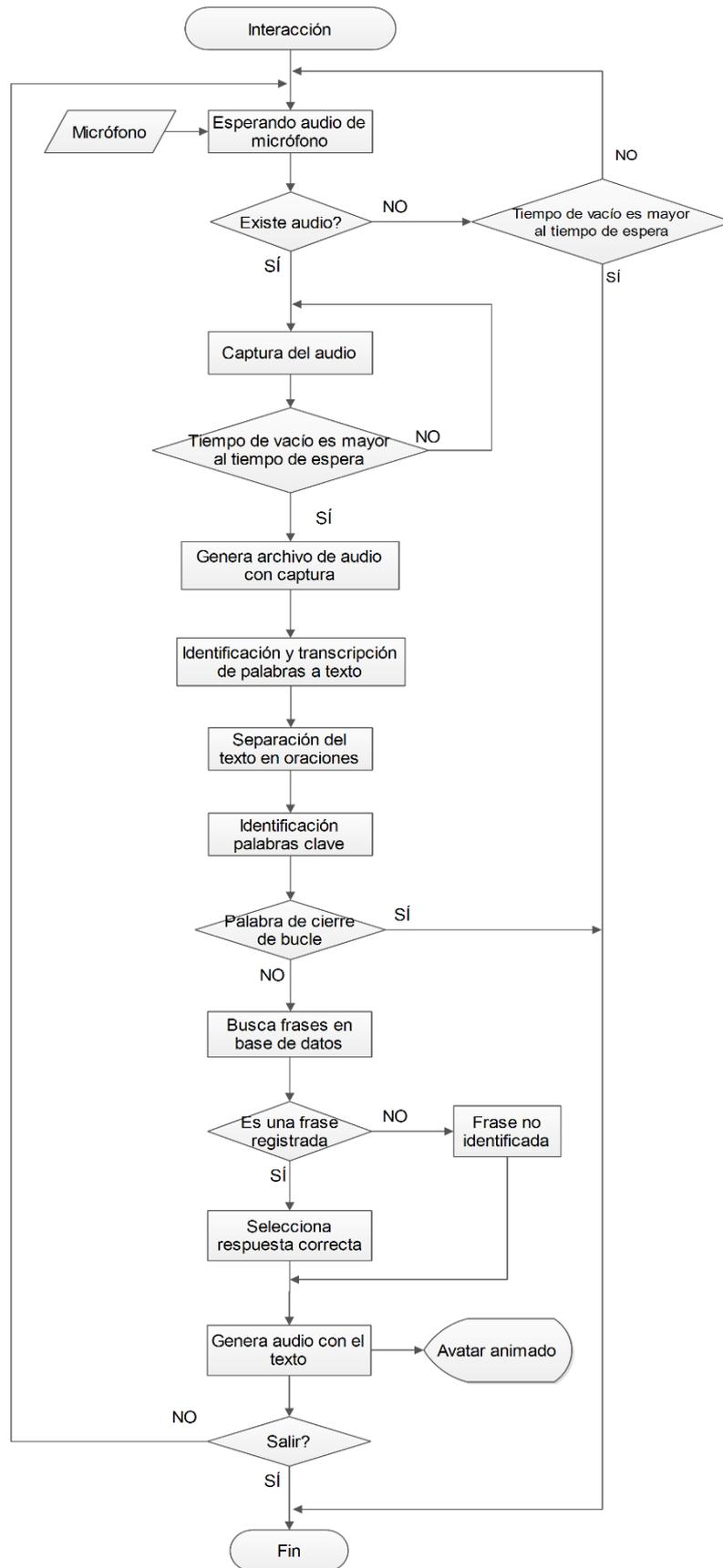


Fig. 17 Diagrama de flujo de la interacción total  
Fuente: Autor

Para el desarrollo y las pruebas de esta etapa se utilizó el compilador Jupyter Notebook. Esta etapa se divide en varias sub etapas, como se expone en el diagrama de bloques presentado en la figura 18.

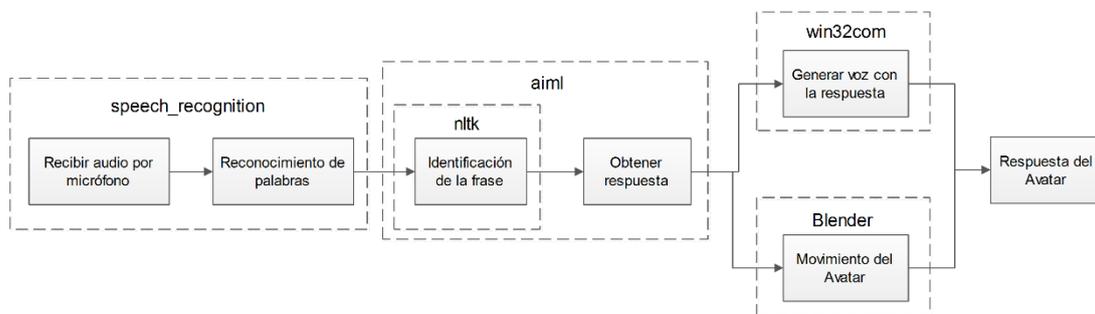


Fig. 18 Diagrama de bloques del sistema  
Fuente: Autor

Como primer paso, se instalaron en la carpeta raíz de Python3 las librerías necesarias: aiml, speech\_recognition, nltk, win32com y speech\_recognition. Cada una de las librerías cumple una función específica en las sub etapas de la interacción, tal como se presenta en la figura..

### 2.3.1 PROGRAMACIÓN EN AIML

La librería aiml, que no solo permite trabajar con el lenguaje de etiquetado del mismo nombre, sino que además genera una especie de cerebro robótico, encargado de almacenar los datos obtenidos de la respuesta de la interacción para ser utilizado después. En otras palabras, esta librería es la encargada de que el Avatar “aprenda” durante la interacción, generando así un cierto grado de inteligencia.

Para utilizar la librería aiml es necesario generar un script de inicialización con la extensión .xml. Este archivo, además de inicializar el proceso de aprendizaje en el script principal, llama a uno o varios archivos con extensión .aiml, los mismos que almacenan las respuestas predeterminadas para que se lleve a cabo la interacción. Estos archivos pueden ser tan extensos como sea necesario, para que la conversación pueda abarcar mayor cantidad de temáticas. El lenguaje AIML se lo maneja mediante las diferentes etiquetas que cumplen funciones específicas

Dentro de la etiqueta `<category>` se establece toda una categoría de interacción, la misma que incluye la pregunta, respuesta u opciones de respuesta y opciones de desambiguación.

La etiqueta `<pattern>` sirve para identificar la pregunta o *frase patrón* que se recibe como entrada del usuario, la misma que devolverá una respuesta etiquetada como `<template>`. Para que la frase patrón pueda ser reconocida debe ingresada exactamente. Sin embargo, se puede utilizar la *estrella* (\*) para completar datos adicionales que pueden variar. Además, los datos ingresados como estrella pueden ser extraídos y almacenados. Al utilizar la etiqueta `<star>` se obtiene el dato ingresado, el mismo que puede almacenarse en una variable mediante la etiqueta `<set>` y recuperado más adelante mediante la etiqueta `<get>`. De esta forma el sistema aprende los datos del usuario, otorgando al sistema un cierto grado de inteligencia.

Otra etiqueta utilizada es `<srail>`, esta tiene el objeto de eliminar ambigüedades con frases de uso común que puedan tener el mismo significado que otras. Por ejemplo, se otorgan respuestas para el patrón “HOLA”, y se escriben etiquetas de desambiguación para que otras frases como “BUENOS DÍAS” estén relacionadas al mismo patrón y devuelvan respuestas similares.

El tema de desambiguación es uno de los más importantes del sistema, ya que se debe considerar la mayor cantidad de opciones posibles de frases o formas de expresión que el usuario podría utilizar. Esto equivale a realizar un análisis semántico, ya que la interacción estará programada sobre un patrón genérico, mientras que la desambiguación se encargará de detectar otras frases de significado similar y reducirlas al patrón base.

### **2.3.2 RECONOCIMIENTO Y GENERACIÓN DE VOZ**

La interacción Avatar – persona se la realiza por medio de la voz. Para ello es necesario que el sistema reciba la voz del usuario, la interprete y pueda devolver la respuesta adecuada, como si se tratase de una conversación. Para el desarrollo del software de reconocimiento de voz se utilizaron las librerías **speech\_recognition** y **wave**, La configuración básica de esta librería permite seleccionar varias opciones de

sistemas prediseñados y se optó por elegir el sistema de reconocimiento de voz de Google.

El hardware utilizado es un micrófono sencillo para computador, el cual no ofrece grandes ventajas en cuanto a calidad y reducción de ruido. Por ello, fue necesario el ajuste de algunos parámetros internos de la librería `speech_recognition`, que permitió mejorar la adquisición del audio y hacer que el sistema sea más robusto a la presencia de ruidos externos.

El proceso de reconocimiento de voz se presenta en el diagrama de la figura 19:

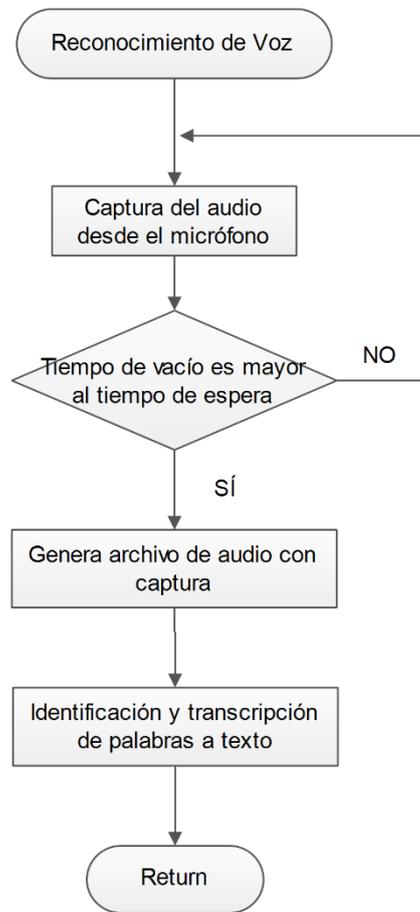


Fig. 19 Diagrama de reconocimiento de voz  
Fuente: Autor

La voz ingresa y es detectada en formato *audiodata*, el mismo que nos permite descomponer sus distintas propiedades del formato wav y extraer solo el vector de amplitud de potencia de cada muestra del audio. Esto permite detectar el inicio del ingreso de voz, cuando la potencia máxima, determinada por un umbral definido, es

detectada, se comienza la captura del audio que será utilizado para la interpretación. Cuando se tiene el audio capturado en audiodata, el sistema de reconocimiento vocal de Google se encarga de hacer la respectiva traducción a texto.

Para la respuesta del Avatar al usuario, se desarrolla un sistema que genere el sonido de las palabras para que el usuario lo entienda sin dificultad. Para ello, se utilizó la librería **win32com**, que, entre sus diversas funciones, posee un sistema de generación vocal, que permite seleccionar entre las voces instaladas en el sistema operativo.

### **2.3.3 INTERACCIÓN ENTRE BLENDER Y PYTHON**

Blender es un software que posee internamente un compilador de Python. Las librerías seleccionadas para desarrollar la interacción están disponibles en su totalidad para la versión de Python 3.6, que es la más común y reciente de este lenguaje de programación. Sin embargo, la versión más reciente y estable de Blender es la 2.79b, misma que posee internamente una versión inferior de Python, la 3.5, la misma que no es muy común y no posee gran diversidad de librerías para su uso, en especial las necesarias para el proyecto. Por ello, no fue posible integrar el algoritmo de interacción directamente en el compilador propio de Blender.

Como alternativa para el desarrollo de este proyecto, se utilizó un método de interacción externa mediante simulación de teclado. Para realizar este método, mediante el algoritmo generado en Python, se simulan entradas por teclado en un momento determinado. Mientras, en Blender, se instalaron sensores de teclado, que esperan que una tecla determinada sea pulsada para poder realizar una determinada acción, que en este caso se trata de los movimientos del avatar

## **CAPÍTULO 3: ANÁLISIS DE RESULTADOS**

Para evaluar el correcto funcionamiento del sistema, se realizaron pruebas internas y externas. Las pruebas internas consistieron en pruebas realizadas al sistema por parte del investigador, para determinar de forma preliminar tasas de error en el funcionamiento básico, así como tiempos de respuesta del sistema. Las pruebas externas, en cambio, son experimentos de interacción del Avatar con terceras personas, completamente ajenas al desarrollo del sistema, para obtener criterios cuantitativos sin sesgo sobre el funcionamiento basado en su experiencia. Para estas pruebas se determinó una serie de pasos a seguir por las personas seleccionadas y, después de experimentar la interacción con el Avatar, respondieron a una encuesta entre las que se encontraban preguntas sobre el funcionamiento general y otras sobre la calidad del sistema evaluado.

### **3.1 PRUEBAS INTERNAS**

Estas pruebas se realizaron en base a repeticiones de acciones para comprobar el porcentaje de aciertos y errores que tendría el sistema, además de medir el tiempo promedio de respuesta a la interacción. En la Tabla 1 se presenta los resultados de estas pruebas realizadas por parte del investigador.

**TABLA 1: RESULTADOS DE LAS PRUEBAS INTERNAS**

Fuente: Autor

<b>PRUEBA</b>	<b>INTENTOS</b>	<b>% ACIERTOS</b>	<b>% ERROR</b>
Respuesta al saludo con la voz	50	96%	4%
Respuesta al saludo con movimiento	50	82%	18%
Aprendizaje	50	88%	12%
Respuesta a preguntas generales	50	84%	16%
<b>PRUEBA</b>	<b>INTENTOS</b>	<b>PROMEDIO</b>	
Tiempo de respuesta	50	2.068 segundos	

Las dos primeras pruebas consistieron en la respuesta al saludo por voz y movimiento respectivamente, con lo que se evalúa el registro de audio y la interacción Blender-Python realizada mediante simulación de ingreso por teclado. La respuesta al saludo con voz tuvo un 94% de efectividad, mientras que la respuesta con movimiento tuvo un 82%.

En la tercera prueba se le indicó al Avatar algunos datos que debía aprender y repetir, para verificar el funcionamiento del sistema de IA básico mediante lenguaje AIML. El resultado es que el 88% de datos los aprendió correctamente, teniendo un error del 12%.

La última prueba fue sobre la interacción y respuesta ante preguntas generales, la misma que se realizó para tener un criterio general del rendimiento del sistema ante una conversación más completa. Para ello, se realizaron preguntas al azar, sobre los contenidos previamente establecidos en la base de datos AIML, y de la misma manera se respondió cuando el Avatar sugería algún tema de conversación. Ante esta prueba, el porcentaje de acierto del sistema fue del 84%, mientras que en el 16% las frases no fueron comprendidas de manera correcta.

Adicionalmente, se tomó una muestra de 50 pruebas aleatorias y se calculó el tiempo de respuesta del sistema, el resultado fue que el sistema tiene un retardo de aproximadamente 2 segundos entre la recepción de la instrucción y la respuesta otorgada.

### 3.2 PRUEBAS EXTERNAS

Para tener una evaluación imparcial sobre el rendimiento del sistema, se decidió que el mismo debe ser evaluado por un grupo de personas que no estén involucradas en el desarrollo. Con este propósito, se realizó un experimento de interacción con 33 personas, a quienes se pidió que realicen preguntas y traten de mantener una conversación con el Avatar. Se pretendió evaluar algunos aspectos puntuales, como la respuesta al saludo, el aprendizaje mediante IA básica y otros aspectos más generales que serían susceptibles a la opinión de los participantes. Para ello, se dio a todos los participantes las siguientes instrucciones previas, que debían cumplir y esperar la respuesta del avatar:

1. Saludar al Avatar
2. Presentarse con su nombre
3. Realizar cualquier pregunta y tratar de mantener la interacción

La primera instrucción procuraba evaluar la respuesta al saludo del avatar, la misma que involucra movimientos específicos y respuestas aleatorias apropiadas. La segunda instrucción sirvió para analizar la capacidad de aprendizaje del Avatar, ya que este debía responder al participante con una frase que incluía repetir su nombre. La tercera instrucción quedaba abierta para que los usuarios realicen cualquier pregunta y continúen con la interacción de forma libre, esto con el fin de evaluar a nivel general la capacidad de interacción según el criterio de los participantes.

Tomando como ejemplo lo realizado por Thomason [10] y Meena y Jokinen [11], se utilizó una encuesta para recoger las experiencias de los participantes durante el la interacción de prueba. Las preguntas se plantearon con el fin de que los participantes evalúen si el sistema cumplía los requerimientos básicos y si tuvieron dificultades para lograr una interacción, estas preguntas están disponibles en el Apéndice A.

Las primeras cinco preguntas evalúan el comportamiento básico del Avatar, de tal forma que se puede determinar si el sistema es o no funcional. Las respuestas a estas preguntas se las presenta en la figura 20.

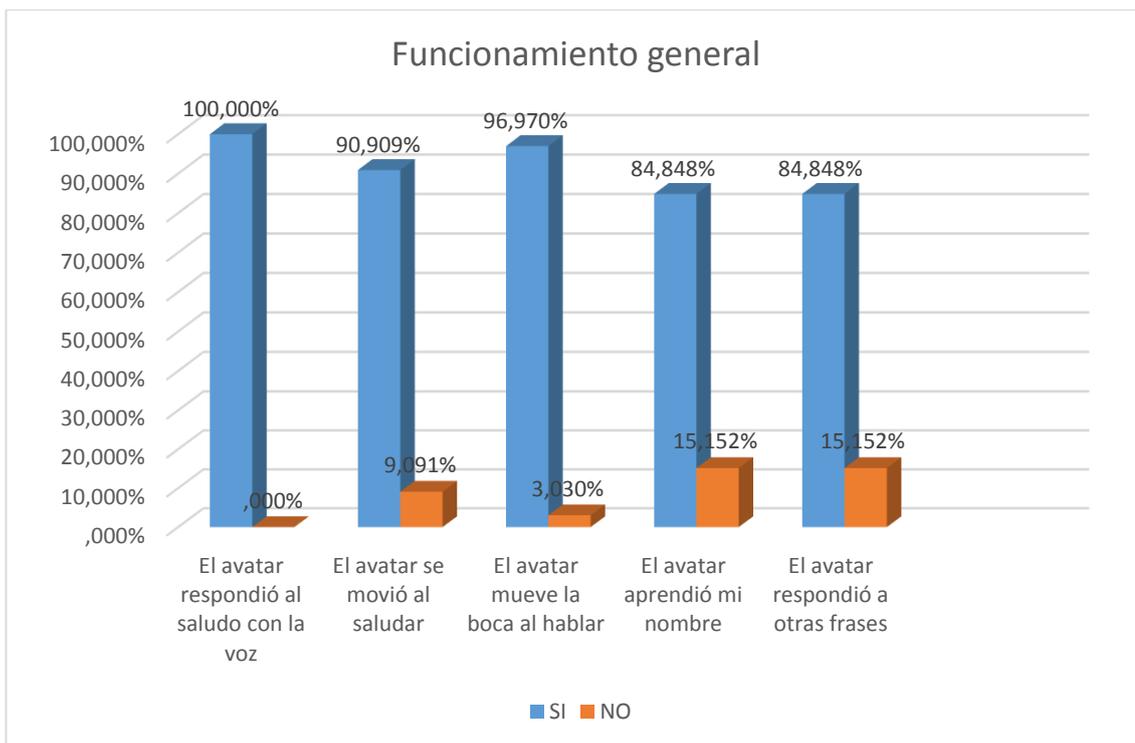


Fig. 20 Resultados del funcionamiento general  
Fuente: Autor

En la pregunta 1: **El avatar respondió al saludo con la voz?**, el total de participantes expresaron que sí recibieron respuesta del Avatar ante el saludo inicial. En la pregunta 2: **El avatar se movió al saludar?**, el 90,91% de los evaluados indicaron que el avatar sí se movió haciendo ademán de saludo, mientras que el 10,09% indicaron que el Avatar no se movió durante la respuesta al saludo. Con estas dos preguntas se pudo determinar que la respuesta al saludo por voz funciona a la perfección, mientras que la respuesta con movimiento es muy buena presenta alguna dificultad.

En la pregunta 3: **El avatar mueve la boca al hablar?**, el 96.97% de participantes manifestaron que el Avatar sí mueve la boca al hablar, mientras que para el 3.03% el Avatar no realizó el movimiento bucal durante la interacción. Se deduce entonces que el movimiento bucal funciona casi en su totalidad, presentando falencias mínimas.

En la pregunta 4: **El avatar aprendió mi nombre?**, el 84.85% de encuestados respondió que el Avatar sí aprendió su nombre, mientras que el restante 15.15%

respondió que no lo hizo. Con esta pregunta se verificó el funcionamiento básico del sistema de IA mediante lenguaje AIML, obteniéndose un resultado altamente positivo.

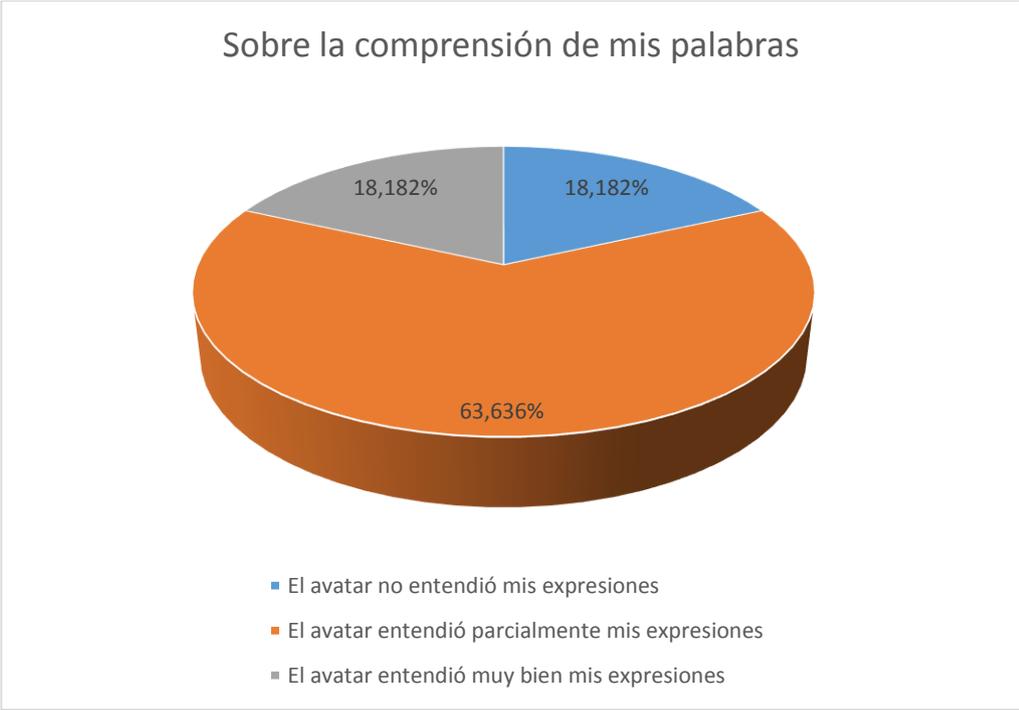
En la pregunta 5: **El avatar respondió a otras frases?**, el 84.85% de participantes indicaron que el Avatar sí responde otras frases, pero el 15.15% indicaron que esto no sucedió. Esto debido a que se dio libertad a los participantes de realizar cualquier pregunta, sin darles mayor orientación acerca de los temas que fueron considerados en la base de datos AIML. El experimento se realizó de esta forma justamente para poder determinar la capacidad de interacción del Avatar ante preguntas aleatorias típicas de una interacción básica, ante lo cual se obtuvo un resultado muy aceptable.

Luego de analizar los resultados obtenidos en el primer grupo de preguntas, se puede deducir que el sistema tiene un rendimiento del 90% respecto a su funcionamiento básico, presentando falencias que representan un 10%, las mismas que son susceptibles de mejoras en futuras investigaciones.

El siguiente grupo de preguntas se enfocó en evaluar de forma más específica algunos aspectos sobre el rendimiento del sistema desde la perspectiva de un usuario común. En la primera pregunta se consultó si el Avatar comprendió las palabras expresadas por los participantes. Los resultados se expresan en la figura 21.

Como se evidencia en la figura, solo el 18% de los encuestados manifiesta que el Avatar no comprendió las palabras expresadas, lo que implica que el 82% expresó que el Avatar sí entendió, aunque la mayoría, que representa un 63%, ha manifestado que solo entendió parcialmente.

En la figura 22 presentamos los resultados de la segunda pregunta, en la que se consultó a los participantes si tuvieron o no que repetir más de una vez las instrucciones, frases o preguntas al Avatar. Como se evidencia, el menor porcentaje de encuestados, que representan el 12%, afirmaron que siempre tuvieron que repetir varias veces las palabras, mientras que el porcentaje mayoritario del 72% expresaron que solo pocas veces tuvieron que repetir las palabras. Un 12% de encuestados manifestaron que el Avatar siempre entendió al expresarse una sola vez.



**Fig. 21 Resultados sobre la comprensión de palabras**  
Fuente: Autor



**Fig. 22 Resultados sobre comprensión de palabras**  
Fuente: Autor

La tercera pregunta, cuyo resultado se presenta en la figura 23, se consultó la opinión de los participantes sobre los movimientos que realiza en Avatar durante la interacción. El 96,97% de encuestados opinaron que los movimientos fueron muy discretos, mientras que una minoría representada por el 3.03% opinaron que el Avatar presentó movimientos excesivos que podían distraer de la conversación. Ningún participante opinó que el Avatar no tuvo movimiento en el transcurso de la interacción.

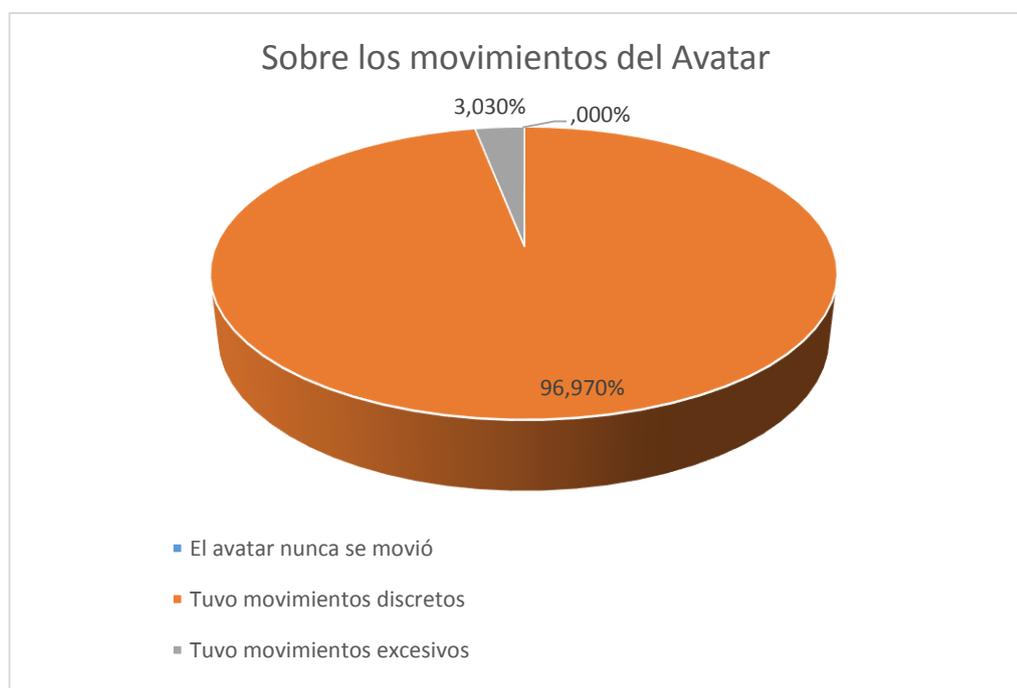


Fig. 23 Resultados sobre movimientos del Avatar  
Fuente: Autor

La última pregunta tuvo como objetivo que los participantes opinen si el sistema desarrollado está en la capacidad de mantener una conversación fluida, sus resultados se aprecian en la figura 24.

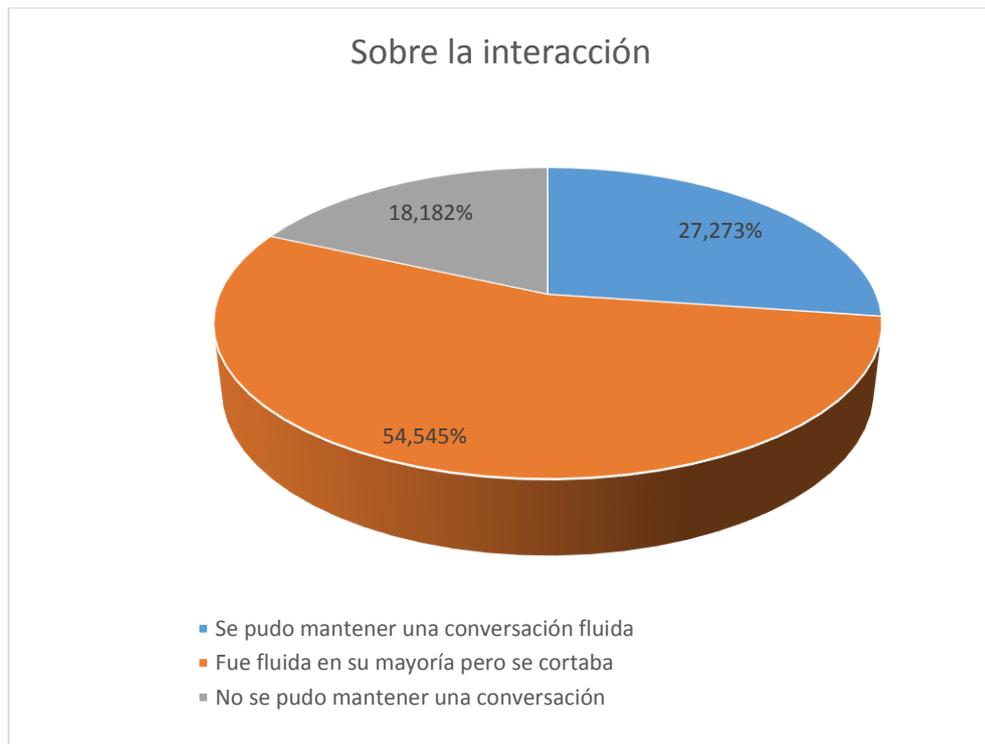


Fig. 24 Resultados sobre la interacción en general  
Fuente: Autor

Ante esta pregunta, el 54% de participantes opinaron que sí se pudo mantener una conversación, pero en su momento sufría interrupciones. Esto debido a que tuvieron que repetir algunas veces las expresiones o a que el Avatar no contenía en la base de datos AIML los temas que proponían los participantes. Un 27% de evaluados manifestaron que sí se puede conversar fluidamente con el Avatar, sin ninguna interrupción. Mientras que, el porcentaje menor que representa al 18% de evaluados, manifestaron que no se puede mantener la interacción fluida con el Avatar.

# **CAPÍTULO 4: CONCLUSIONES Y RECOMENDACIONES**

Los métodos de interacción Humano – máquina han avanzado a lo largo del tiempo. Se desarrollan y mejoran sistemas para que esta interacción sea más fluida, tratando de que cada vez se aproxime más a una interacción real, para lo cual se investigan nuevos métodos de interpretación, algoritmos búsqueda de información e inteligencia artificial, además de proponer nuevas interfaces que faciliten la comunicación para el usuario.

La integración directa del lenguaje Python y el software Blender no pudo realizarse debido a la incompatibilidad entre versiones tanto del lenguaje y sus librerías como del software. Como alternativa se optó por aplicar una interacción externa mediante simulación de teclado, con la cual se logró el objetivo deseado, ya que Python generaba los comandos necesarios para que Blender realice las acciones deseadas para la interacción, de tal forma que el software y el lenguaje de programación trabajan en conjunto de manera síncrona en tiempo real.

El uso de lenguaje AIML permitió desarrollar una aplicación que posea un cierto grado de inteligencia, ya que el sistema desarrollado es capaz de responder algunas preguntas del usuario, proponer otros temas de conversación y aprender algunos datos del usuario como su nombre, lugar de nacimiento y otros. Sin embargo, el sistema presenta ciertas falencias debido a que utiliza y reconoce frases muy genéricas. Además, en muchos casos no es capaz de identificar entre una respuesta del usuario no almacenada y una palabra o frase sin comprensión, por lo que, para

cualquiera de los dos casos, optará por devolver una frase indicando que no entendió o a su vez cambiar de tema.

De las pruebas realizadas se concluye que el sistema desarrollado cumple con los requerimientos básicos del diseño, ya que se logró que el Avatar responda tanto con voz y movimientos a la interacción con personas, casi a la perfección. Sin embargo, se encontró ciertas dificultades en la recepción de la señal de voz y en la interpretación de mensajes complejos.

Se concluye además que el sistema tiene una respuesta aceptable, tanto en respuesta por voz, movimientos, aprendizaje e interacción a nivel general. Las pruebas realizadas a participantes ajenos al proyecto dejan entrever que el sistema debe ser mejorado, para obtener una interacción más fluida y más realista.

Como recomendación, una forma de mejorar este sistema es mediante la implementación de algoritmos más complejos de inteligencia artificial, como pueden ser las redes neuronales o el aprendizaje de máquina. Esto en pos de mejorar la interacción y el reconocimiento de frases, otorgando un mejor análisis semántico, que a su vez permitirá que el sistema devuelva respuestas más precisas ante las preguntas o requerimientos del usuario.

Se recomienda también realizar un estudio minucioso al sistema de captación de audio, con equipos más sofisticados y filtros que limpien la señal de audio que ingresa, de tal forma que el sistema pueda detectar de mejor manera el sonido y, por ende, reconocer con mayor exactitud las palabras y frases expresadas por el ser humano que utiliza el sistema.

## **TRABAJO FUTURO**

El presente trabajo consiste solamente en la primera etapa del proyecto AVATAR, que es desarrollado por el Grupo de Investigación en Telecomunicaciones y Telemática (GITEL) de la Universidad Politécnica Salesiana sede Cuenca. Este proyecto contempla etapas posteriores que tendrán como base lo realizado en esta investigación.

En una siguiente etapa, se prevé mejorar los movimientos del Avatar, mediante la captura de movimientos reales de seres humanos, utilizando una cámara y diversos puntos de detección.

La cumbre de este proyecto será que el Avatar interactúe con niños que presentan trastorno del espectro autista (TEA), generando conversaciones y ayudando en el desarrollo y seguimiento de tratamientos profesionales para mejorar la interacción social de estos niños. El sistema completo que cumpla con estos objetivos requerirá de un muy elevado nivel de programación en algoritmos de inteligencia artificial, reconocimiento de voz e imágenes y telecomunicaciones mediante TICs para el monitoreo respectivo por parte de los profesionales en el área.

Todos los ámbitos antes mencionados respecto al proyecto AVATAR son propicios para generar varios proyectos de investigación y trabajos de titulación, a realizarse por alumnos de la universidad y docentes e investigadores de GITEL.

## REFERENCIAS BIBLIOGRÁFICAS

- [1] Y. Wang, J. Geigel, y A. Herbert, “Reading Personality: Avatar vs. Human Faces”, en *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, pp. 479–484.
- [2] M. Macedonia, I. Groher, y F. Roithmayr, “Intelligent virtual agents as language trainers facilitate multilingualism”, *Front. Psychol.*, vol. 5, p. 295, 2014.
- [3] J. Thomason, S. Zhang, R. J. Mooney, y P. Stone, “Learning to Interpret Natural Language Commands through Human-Robot Dialog.”, presentado en IJCAI, 2015, pp. 1923–1929.
- [4] S. Koval, “Convergencias tecnológicas en la era de la integración hombre-máquina”, *Razón Palabra*, vol. 16, núm. 75, 2011.
- [5] V. M. L. Jaquero, “Interfaces de usuario adaptativas basadas en modelos y agentes software”, 2005.
- [6] G. W. Furnas, T. K. Landauer, L. M. Gomez, y S. T. Dumais, “The vocabulary problem in human-system communication”, *Commun. ACM*, vol. 30, núm. 11, pp. 964–971, 1987.
- [7] H. Oberquelle, I. Kupka, y S. Maass, “A view of human—machine communication and co-operation”, *Int. J. Man-Mach. Stud.*, vol. 19, núm. 4, pp. 309–333, 1983.
- [8] V. Vitsentiy, “Improvement of human-machine interaction with applications to information retrieval system”, presentado en Intelligent Systems, 2002. Proceedings. 2002 First International IEEE Symposium, 2002, vol. 1, pp. 363–368.

- [9] P. R. Cohen y S. L. Oviatt, “The role of voice input for human-machine communication”, *Proc. Natl. Acad. Sci.*, vol. 92, núm. 22, pp. 9921–9927, 1995.
- [10] P. Lago y C. J. Guarín, “An affective inference model based on facial expression analysis”, *IEEE Lat. Am. Trans.*, vol. 12, núm. 3, pp. 423–429, 2014.
- [11] J. Minguez, “Tecnología de interfaz cerebro-computador”, 2008.
- [12] P. Bach-y-Rita y S. W. Kercel, “Sensory substitution and the human-machine interface”, *Trends Cogn. Sci.*, vol. 7, núm. 12, pp. 541–546, dic. 2003.
- [13] R. Meena, K. Jokinen, y G. Wilcock, “Integration of gestures and speech in human-robot interaction”, presentado en Cognitive Infocommunications (CogInfoCom), 2012 IEEE 3rd International Conference on, 2012, pp. 673–678.
- [14] T. Watanabe y M. Okubo, “An embodied virtual communication system for human interaction sharing”, en *Systems, Man, and Cybernetics, 1999. IEEE SMC '99 Conference Proceedings. 1999 IEEE International Conference on*, 1999, vol. 5, pp. 1060–1065 vol.5.
- [15] D. Moore, Yufang Cheng, P. McGrath, y N. J. Powell, “Collaborative Virtual Environment Technology for People With Autism”, *Focus Autism Dev. Disabil.*, vol. 20, núm. 4, pp. 231–243, nov. 2005.
- [16] R. Wallace, “Artificial linguistic internet computer entity (ALICE)”, 2001.
- [17] S. Baldassarri y E. Cerezo, “Maxine: Embodied conversational agents for multimodal emotional communication”, *Comput. Graph.*, 2012.
- [18] S. Baldassarri, E. Cerezo, y D. Anaya, “Interacción emocional con actores virtuales a través de lenguaje natural”, *Interacción*, pp. 343–352, 2007.
- [19] L. F. Guerrero-Vásquez, J. F. Bravo-Torres, y M. López-Nores, “AVATAR ”autism: Virtual agents to augment relationships in children””, en *2017 IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, 2017, pp. 1–4.
- [20] L. F. Guerrero-Vásquez, D. X. Landy-Rivera, J. F. Bravo-Torres, M. López-Nores, y R. Castro-Serrano, “AVATAR: Contribution to Human-Computer interaction processes through the adaptation of semi-personalized virtual agents”, presentado en Congreso Bienal de IEEE Argentina ARGENCON 2018, 2018.
- [21] L. Guerrero, D. Landy-Rivera, G. Ávila, J. Bravo-Torres, y M. López-Nores, *Restricted Interest-Based Adaptation of Avatar for Interaction with Children with Autism Spectrum Disorder*. 2019.

- [22] “Interactive Avatars”, *Invirtua Animation For Autism*. [En línea]. Disponible en: <https://invirtua.com/>. [Consultado: 25-jul-2018].
- [23] B. Foundation, “About”, *blender.org*. .
- [24] B. S. LLC, “Blend Swap”, *Blend Swap*. [En línea]. Disponible en: <http://www.blendswap.com>. [Consultado: 26-jul-2018].
- [25] M. Caeiro Rodriguez, R. Miguez Vazquez, y A. Blanco Pesquera, “Management of Copyright Issues and Creative Commons Licenses in Edu-AREA”, *IEEE Lat. Am. Trans.*, vol. 13, núm. 11, pp. 3655–3660, nov. 2015.
- [26] A. Downey, J. Elkner, y C. Meyers, “Aprenda a Pensar como un Programador con Python”, 2002.
- [27] J. Knowlton, “Python. tr: Fernández Vélez, María Jesús (1 edición) 2009”.
- [28] “SpeechRecognition”, *PyPI*. [En línea]. Disponible en: <https://pypi.org/project/SpeechRecognition/>. [Consultado: 26-jul-2018].
- [29] “python-aiml”, *PyPI*. [En línea]. Disponible en: <https://pypi.org/project/python-aiml/>. [Consultado: 26-jul-2018].
- [30] “pywin32”, *PyPI*. [En línea]. Disponible en: <https://pypi.org/project/pywin32/>. [Consultado: 26-jul-2018].
- [31] M. Satran, “Windows API Index”. [En línea]. Disponible en: <https://docs.microsoft.com/en-us/windows/desktop/apiindex/windows-api-list>. [Consultado: 26-jul-2018].

## APÉNDICES

### APÉNDICE A: ENCUESTA APLICADA A LOS PARTICIPANTES QUE REALIZARON LA INTERACCIÓN CON EL PROYECTO



#### Funcionamiento general:

- |   |        |        |
|---|--------|--------|
| 1. El avatar respondió al saludo con la voz | Sí ( ) | No ( ) |
| 2. El avatar se movió al saludar            | Sí ( ) | No ( ) |
| 3. El avatar mueve la boca al hablar        | Sí ( ) | No ( ) |
| 4. El avatar aprendió mi nombre             | Sí ( ) | No ( ) |
| 5. El avatar respondió a otras frases       | Sí ( ) | No ( ) |

#### Funcionamiento específico:

Por favor seleccione la opción que más se aproxime a su experiencia durante la interacción con el Avatar:

1. Sobre la comprensión de mis palabras:
  - El avatar no entendió mis expresiones ( )
  - El avatar entendió parcialmente mis expresiones ( )
  - El avatar entendió muy bien mis expresiones ( )

2. Para que el avatar entienda mis palabras:
- Siempre tuve que repetirlas varias veces ( )
  - Algunas veces tuve que repetirlas varias veces ( )
  - Siempre entendió al decirlas una sola vez ( )
3. Sobre los movimientos del avatar:
- El avatar nunca se movió ( )
  - El avatar tuvo algunos movimientos ( )
4. Sobre los movimientos del avatar:
- Los movimientos fueron discretos ( )
  - Los movimientos fueron muy toscos y distraían de la conversación ( )
5. Sobre la interacción:
- Se pudo mantener una conversación fluida ( )
  - La conversación fue fluida en su mayoría, pero a veces se cortaba ( )
  - No se pudo mantener una conversación ( )