

**RECONOCIMIENTO DE VOZ PARA UN SISTEMA DE INTERACCIÓN HUMANO
MÁQUINA ORIENTADO A PERSONAS CON LIMITACIONES MOTRICES**

DANIEL ALEXIS SANTANDER ARIZA
KEVIN STEVEN ROSERO RAMOS
MIGUEL STEVEN LIBREROS SEGURA

FUNDACIÓN UNIVERSITARIA CATÓLICA LUMEN GENTIUM
FACULTAD DE INGENIERÍA
INGENIERÍA DE SISTEMAS
SANTIAGO DE CALI
2020

**RECONOCIMIENTO DE VOZ PARA UN SISTEMA DE INTERACCIÓN HUMANO
MÁQUINA ORIENTADO A PERSONAS CON LIMITACIONES MOTRICES**

DANIEL ALEXIS SANTANDER ARIZA
KEVIN STEVEN ROSERO RAMOS
MIGUEL STEVEN LIBREROS SEGURA

Trabajo de grado para optar al título de
Ingeniero de Sistemas

Directores

JOSÉ HERNANDO MOSQUERA DE LA CRUZ, M.Sc., Ing.
CARLOS DIEGO FERRÍN BOLAÑOS, M.Sc., Ing.

FUNDACIÓN UNIVERSITARIA CATÓLICA LUMEN GENTIUM
FACULTAD DE INGENIERÍA
INGENIERÍA DE SISTEMAS
SANTIAGO DE CALI
2020

Nota de aceptación:

Aprobado por el Comité de Grado en cumplimiento de los requisitos exigidos por la Fundación Universitaria Lumen Gentium para optar al título de Ingeniero de Sistemas

Jurado

Jurado

Santiago de Cali, Mayo 27 de 2020

DEDICATORIA

Daniel Alexis Santander Ariza

En memoria de mi abuela Sofía Sánchez de Ariza, por su amor y apoyo al momento de iniciar mi carrera universitaria. A mis padres Fabio Santander y Amparo Ariza, por el trabajo y sacrificio en todos estos años, gracias a ustedes he podido culminar una parte importante de mi vida. Pero ante todo a Dios por darme la voluntad y las fuerzas para llegar hasta aquí.

Miguel Steven Libreros Segura

A mis padres Miguel Libreros y Sandra Segura que me han brindado su paciencia y con su esfuerzo me han permitido llegar a cumplir hoy un sueño más, gracias por inculcar en mí el ejemplo de la perseverancia y valentía, de siempre mirar hacia adelante sin importar el obstáculo que se interponga.

Kevin Stiven Rosero Ramos

Para mis padres, quienes me impulsaron a empezar este camino y estuvieron día a día ahí para apoyarme y levantarme el ánimo cuando más lo necesité. A los distintos profesores que tuve a lo largo de mi camino académico pues su paciencia y dedicación lograron pasmar lo que actualmente soy y lo que puedo hacer. A mis compañeros de clases que siempre estuvimos unidos para sobrepasar los diferentes obstáculos que nos encontramos y de alguna manera por ser un segundo apoyo cuando algo parecía no tener solución.

AGRADECIMIENTOS

Los estudiantes agradecen a la Fundación Universitaria Católica Lumen Gentium Unicatólica, al programa de Ingeniería de Sistemas y a los directores José Hernando Mosquera De La Cruz, M.Sc., Ing. y Carlos Diego Ferrín Bolaños, M.Sc., Ing., por su constante dedicación y apoyo en todas las etapas de desarrollo para culminar este trabajo de grado, de igual forma agradecen al proyecto de investigación interna llamado Asistente virtual para personas tetraplégicas controlado por voz y gestos faciales (Primera Fase) de la IV Convocatoria, Resolución 389 del 26 de septiembre de 2019.

CONTENIDO

	Pág.
INTRODUCCIÓN	18
1. DESCRIPCIÓN DEL PROYECTO	19
1.1 DEFINICIÓN DEL PROBLEMA	19
1.2 ANTECEDENTES DEL PROBLEMA	19
1.3 FORMULACIÓN DEL PROBLEMA	21
2. JUSTIFICACIÓN	22
3. OBJETIVOS	23
3.1 OBJETIVO GENERAL	23
3.2 OBJETIVOS ESPECÍFICOS	23
4. MARCO REFERENCIAL	24
4.1 ESTADO DEL ARTE	24
4.1.1 Referencias académicas	24
4.1.2 Referencias comerciales	30
4.2 MARCO TEÓRICO	31
4.2.1 Interacción Humano-Máquina	31
4.2.2 Reconocimiento de Voz	32
4.2.3 Síntesis de voz	32
4.2.4 Sistema Operativo	33
4.2.5 Acceso a la información para personas con limitaciones motrices.	33
4.3 MARCO CONCEPTUAL	34
4.3.1 Captura y pre-proceso	34
4.3.2 Extracción de características	34
4.3.3 Clasificación y decisión	34
4.4 MARCO CONTEXTUAL	35
4.5 MARCO LEGAL	36

5. METODOLOGÍA	37
5.1 MÉTODO Y TIPO DE ESTUDIO	37
5.2 TÉCNICA E INSTRUMENTOS	37
5.3 ARQUITECTURA DEL SOFTWARE IMPLEMENTADO	38
5.3.1 Reconocimiento de voz	39
5.3.2 Detección de comandos	40
5.3.3 Control del sistema operativo	41
5.3.4 Síntesis de voz	42
6. PRUEBAS Y RESULTADOS	44
6.1 PROTOCOLO DE PRUEBAS	44
6.1.1 SUJETOS DE PRUEBA	44
6.1.2 RECONOCIMIENTO DE COMANDOS	46
6.1.3 RECONOCIMIENTO DE DICTADOS	47
6.1.4 INTERFAZ FINAL CON WHATSAPP DESKTOP	48
6.1.5 EXPERIENCIA DE USUARIO	50
7. ALCANCES Y LIMITACIONES	52
7.1 ALCANCES	52
7.2 LIMITACIONES	52
8. CONCLUSIONES	53
9. RECOMENDACIONES	54
BIBLIOGRAFÍA	55
ANEXOS	60

LISTA DE FIGURAS

	Pág.
Figura 1. Etapa de control del marco referencial 4.1.1.2.	26
Figura 2. Arquitectura interna de un sistema de reconocimiento de voz.....	32
Figura 3. Arquitectura interna de un sintetizador de voz.....	33
Figura 4. Diagrama de bloques general del sistema.....	38
Figura 5. Ejemplo de implementación Librería <i>SpeechRecognition</i>	39
Figura 6. Ejemplo de implementación de la función <i>in ()</i>	40
Figura 7. Lista de teclas soportadas por la Librería <i>pyautogui</i>	41
Figura 8. Ejemplos de uso de la Librería <i>PyAudio</i>	42
Figura 9. Ejemplo de implementación de la Librería <i>Pytsx3</i>	42

LISTA DE TABLAS

	Pág.
Tabla 1. Lista de comandos y acciones.....	40
Tabla 2. Perfil usuarios de prueba	45
Tabla 3. Resultado porcentualmente sobre el grado de familiaridad de los usuarios de prueba.	45
Tabla 4. Resultados de la prueba de comandos.....	46
Tabla 5. Resultados de la prueba de reconocimiento de voz.	48
Tabla 6. Protocolo de pruebas final.	49
Tabla 7. Resultados protocolo de pruebas final.....	49
Tabla 8. Resultados de la experiencia del usuario.....	51

LISTA DE ANEXOS

	Pág.
Anexo A. REPOSITORIO CODIGO FUENTE.....	60
Anexo B. CONSENTIMIENTO INFORMADO PARA PARTICIPACIÓN EN INVESTIGACIÓN.....	60

GLOSARIO

API'S: interfaz de programación de aplicaciones, conjunto de definiciones y protocolos que se utiliza para desarrollar e integrar el software de las aplicaciones.

APLICACIONES: programa informático diseñado como herramienta para permitir a un usuario realizar uno o diversos tipos de tareas.

ASISTENTE VIRTUAL: programa informático capaz de reconocer el lenguaje natural utilizado por el usuario permitiendo establecer una conversación para responder preguntas, hacer recomendaciones o realizar acciones solicitadas.

AUTOMATIZACIÓN: es un sistema donde se transfieren tareas, realizadas habitualmente por operadores humanos a un conjunto de elementos tecnológicos.

COEFICIENTE CEPSTRALES: coeficientes para la representación del habla basados en la percepción auditiva humana. Estos surgen de la necesidad, en el área del reconocimiento de audio automático.

COMANDOS DE VOZ: sirven para controlar las aplicaciones o sistemas sin necesidad de utilizar el teclado o la pantalla.

CUADRIPLÉJIA/TETRAPLÉJIA: signo clínico por el que se produce parálisis total o parcial de brazos y piernas causada por un daño en la médula espinal, específicamente en alguna de las vértebras cervicales.

DOMÓTICA: conjunto de técnicas orientadas a automatizar una vivienda, que integran la tecnología en los sistemas de seguridad, gestión energética, bienestar o comunicaciones.

ESPECTRO DE RUIDO: repartición de la energía sonora en función de cada una de estas, permite establecer si el ruido contiene frecuencias bajas (graves), medias o altas (agudas).

FONEMAS: descripción teórica de los sonidos de la lengua.

INTERACCIÓN HUMANO-MÁQUINA: estudio dedicado a diseñar, evaluar e implementar sistemas informáticos interactivos para el uso humano, y a estudiar los fenómenos relacionados más significativos.

INTERFAZ DE VOZ: dispositivo que permite la interacción humana con ordenadores a través de una plataforma de voz/habla para iniciar procesos o servicios automatizados.

LESIÓN RAQUIMEDULAR: daño a la médula espinal que puede abarcar simultáneamente las meninges, los vasos sanguíneos y el tejido nervioso.

LIMITACIONES MOTRICES: deficiencia que provoca en el individuo que la padece alguna disfunción en el aparato locomotor.

NAVEGADOR: programa que permite navegar por internet u otra red informática de comunicaciones.

PROTOTIPO: primer ejemplar que se fabrica de una figura, un invento u otra cosa, y que sirve de modelo para fabricar otras iguales, o molde original con el que se fabrica.

PYTHON: se trata de un lenguaje de programación multiparadigma, ya que soporta orientación a objetos, programación imperativa y, en menor medida, programación funcional.

RECONOCIMIENTO DE VOZ: herramienta computacional capaz de procesar la señal de voz emitida por el ser humano y reconocer la información contenida en ésta, convirtiéndola en texto o emitiendo órdenes que actúan sobre un proceso.

SÍNTESIS DE VOZ: producción artificial del habla, convierte el lenguaje de texto normal en habla.

SISTEMA DE CÓMPUTO: conjunto de elementos electrónicos que interactúan entre sí, (Hardware) para procesar y almacenar información de acuerdo con una serie de instrucciones.

SISTEMA OPERATIVO: programa o software básico de un ordenador; las funciones básicas del Sistema Operativo son administrar los recursos del ordenador, coordinar el hardware y organizar los archivos y directorios de su sistema.

SOFTWARE: conjunto de programas y rutinas que permiten a la computadora realizar determinadas tareas.

TRAQUEOTOMÍA: operación quirúrgica que consiste en practicar una abertura en la tráquea para comunicarla con el exterior.

UNIDAD DE CONTROL: dirige la secuencia de pasos de modo que la computadora lleve a cabo un ciclo completo de ejecución de una instrucción, y hacer esto con todas las instrucciones de que conste el programa.

WHATSAPP DESKTOP: aplicación de mensajería disponible para usarse en computadoras, en la que se envían y reciben mensajes mediante Internet.

CIBERNÉTICA: ciencia que estudia los sistemas de comunicación y de regulación automática de los seres vivos y los aplica a sistemas electrónicos y mecánicos que se parecen a ellos.

RESUMEN

El desarrollo de este trabajo de grado se basa en la problemática que afrontan las personas con limitaciones motrices al momento de conectarse con el mundo vía Internet, ya que no les es posible utilizar el mouse y teclado de una forma tradicional, en este orden de ideas, este trabajo de grado generó aportes en esta temática desarrollando una interfaz que permite la interacción entre una persona con limitaciones motrices y la aplicación WhatsApp Desktop mediante comandos de voz.

El trabajo de grado utilizó un método Inductivo, teniendo un enfoque cualitativo, con alcances exploratorios, explicativos y experimentales. En el cual se realizó una investigación sobre sistemas afines que empleen señales de voz para el comando de interfaces humano-máquina, identificando técnicas y experimentos utilizadas en la comunidad científica. Con base en la información extraída, se definió la aplicación WhatsApp Desktop como instrumento de pruebas para implementar la interfaz, se diseñaron los componentes software que permiten una navegación básica en la aplicación y posteriormente se implementaron los algoritmos de procesamiento de voz e interacción con el sistema operativo del computador utilizando lenguaje Python.

Los alcances y limitaciones de la interfaz fueron cuantificados mediante pruebas y encuestas a un grupo de siete personas. Inicialmente se caracterizó la familiarización de los usuarios con Internet, la aplicación WhatsApp y sistemas de reconocimiento de voz. Posteriormente se realizó un conjunto de tres pruebas; la primera prueba evaluó el desempeño del sistema para reconocer un conjunto de 26 comandos donde cada usuario pronunció cada comando de forma independiente tres veces, evaluando un total de 546 comandos y obteniendo un porcentaje de acierto total del 91.76%.

La segunda prueba evaluó el desempeño del sistema para reconocer dictados, para esto se definió un dictado de 29 palabras el cual cada uno de los siete usuarios pronunció tres veces, evaluando un total de 609 palabras en dictado y obteniendo un porcentaje de acierto de 93%.

La tercera prueba evaluó el desempeño de la interfaz directamente sobre la aplicación WhatsApp Desktop, para esto se diseñó un protocolo de pruebas que combina comandos con dictados y permite al usuario abrir la aplicación, seleccionar un contacto y enviar un mensaje. Esta prueba fue repetida tres veces por cada uno de los siete usuarios y obtuvo un porcentaje de acierto del 91.16%.

Finalmente, se realizó una encuesta para cuantificar la percepción de los usuarios luego de interactuar con la interfaz desarrollada, en general un 40.6% y un 23.2% de los sujetos de prueba están de acuerdo y totalmente de acuerdo en que la interfaz

es una herramienta útil y mejora en la calidad de vida de personas con limitaciones motrices. De igual manera, manifestaron como trabajo futuro la necesidad de integrar signos de puntuación y tildes para una mejor calidad del texto enviado como mensaje.

ABSTRACT

The development of this degree work is based on the problems faced by people with motor limitations when connecting to the world via the Internet, since it is not possible for them to use the mouse and keyboard in a traditional way. In this order of ideas, this degree work generated contributions in this area by developing an interface that allows interaction between a person with motor limitations and the WhatsApp Desktop application through voice commands.

The degree work used an inductive method, taking a qualitative approach, with exploratory, explanatory and experimental scopes. In which research was conducted on related systems that use voice signals to command human-machine interfaces, identifying techniques and experiments used in the scientific community. Based on the information extracted, the WhatsApp Desktop application was defined as a test instrument to implement the interface, the software components that allow basic navigation in the application were designed and later the algorithms for voice processing and interaction with the computer operating system were implemented using Python language.

The scope and limitations of the interface were quantified through tests and surveys of a group of seven people. Initially, users' familiarization with the Internet, the WhatsApp application and speech recognition systems was characterized. Subsequently, a set of three tests were performed; the first test evaluated the system's performance in recognizing a set of 26 commands where each user independently pronounced each command three times, evaluating a total of 546 commands and obtaining a total success rate of 91.76%.

The second test evaluated the performance of the system to recognize dictations. For this purpose, a dictation of 29 words was defined, which each of the seven users pronounced three times, evaluating a total of 609 words in dictation and obtaining a 93% success rate.

The third test evaluated the performance of the interface directly on the WhatsApp Desktop application. For this, a test protocol was designed that combines commands with dictations and allows the user to open the application, select a contact and send a message. This test was repeated three times by each of the seven users and obtained a 91.16% success rate.

Finally, a survey was conducted to quantify the users' perception after interacting with the developed interface. In general, 40.6% and 23.2% of the test subjects agree and totally agree that the interface is a useful tool and improves the quality of life of people with motor limitations. Similarly, they expressed as future work the need to

integrate punctuation marks and accents for better quality of the text sent as a message.

INTRODUCCIÓN

En la actualidad las computadoras aportan soluciones para una gran parte de las necesidades del ser humano; las computadoras están casi en todos los ámbitos de nuestras vidas, empresas, bancos, escuelas, universidades y hogares, teniendo una interacción y un flujo de datos mediante un sistema de cómputo. La gran acogida de la computadora ha sido posible gracias a la disminución de costos y a la gran acogida del Internet, el cual permite una comunicación sin barreras y facilita la expansión del conocimiento llegando hasta las zonas más remotas del mundo (Avalos & Borba, 2020; I. De, 2019). Una de sus principales aplicaciones es la educación inclusiva, la cual reconoce que las personas con limitaciones motrices no han sido consideradas en el diseño de las interfaces y es necesario concebir nuevas formas de interacción que logren disminuir las barreras que limitan el aprendizaje y la comunicación. (Del & Kano, 2013)

Partiendo de la necesidad de implementar nuevos mecanismos de interacción entre las personas y los computadores, se propone en éste trabajo de grado titulado *Reconocimiento de voz para un sistema de interacción humano-máquina orientado a personas con limitaciones motrices*, el cual busca desarrollar una interfaz de computadora dirigida por comandos de voz que permite a personas con limitaciones motrices realizar tareas básicas de interacción con la aplicación WhatsApp Desktop, para esto se implementó el sistema software y se propuso un protocolo de pruebas, el cual después de ser ejecutado permitió establecer los alcances y limitaciones de la interfaz propuesta.

Este documento consta de siete capítulos, en el primer capítulo se describe el proyecto, en el capítulo 2 y 3 se plantean los objetivos y el marco de referencia, en el capítulo 4 se explica el desarrollo de la interfaz propuesta. Finalmente, en los capítulos 5, 6 y 7 se describen el protocolo de pruebas realizado, conclusiones y bibliografía.

1. DESCRIPCIÓN DEL PROYECTO

1.1 DEFINICIÓN DEL PROBLEMA

El uso de las computadoras en un plano regional y urbano se considera como un artefacto eficaz que dispone de instrumentos y programas que hacen más viable el análisis detallado de gran cantidad de información. Esto demuestra que el desempeño de los equipos de cómputo ha mejorado exponencialmente a la par con los avances en la ingeniería electrónica. Sin embargo, se puede denotar la poca actualización que han recibido los periféricos con los cuales el humano interactúa con la máquina, ya que conservamos el mouse y teclado como principales medios de comunicación con el ordenador, lo cual desencadena una experiencia de uso poco natural convirtiéndose en una limitante para personas con alguna discapacidad que les impida tener contacto físico con estos periféricos (Oswaldo et al., 2018)

Con base en lo anterior este trabajo de grado propone una forma alternativa de interacción humano-computador orientada a personas con limitaciones motrices la cual permite comandar la aplicación de escritorio WhatsApp Desktop utilizando comandos de voz.

1.2 ANTECEDENTES DEL PROBLEMA

El desarrollo de un asistente virtuales basados voz orientados a la manipulación del computador por parte personas con limitaciones motrices (Específicamente pacientes con lesiones medulares a nivel cervical C3-C6) presentan varios desafíos:

El primer desafío está relacionado con la adaptación de los asistentes virtuales a las nuevas formas de pronunciación de los fonemas por parte de los pacientes con tetraplejia, en algunos casos los pacientes deben contar con una traqueotomía que ayuda a la respiración cuando el camino habitual presenta alguna complicación (GARCÍA P et al., 2007), sin embargo, este procedimiento quirúrgico modifica el flujo de aire a través de las cuerdas bucales cambiando las características frecuenciales y de amplitud durante el habla, lo cual se reduce a una mayor robustez en el sistema de reconocimiento de voz (Rashid & Islam, 2017).

Un segundo desafío se encuentra en la calibración o configuración de los asistentes virtuales, especialmente en el caso de los pacientes tetraplégicos se procura tener una manipulación exacta del cursor a partir de la voz y es un gran reto debido a la gran cantidad de posibilidades y tiempo en la ubicación a detalle requeridos al generar posiciones efectivas, entendiendo que una calibración supervisada de los movimientos del cursor requiere un lazo cerrado de realimentación para alcanzar la exactitud, por esto es una necesidad latente concebir asistentes virtuales que

calibren con la menor cantidad de datos posible y el mejor tiempo la ubicación del cursor buscando optimizar esta funcionalidad de forma no supervisada (M. M. Pérez & Gómez, 2017)

Un tercer desafío está relacionado con la realimentación que existe entre el paciente y el sistema de inteligencia detrás de los asistentes virtuales, esta tarea es muy importante pues permite corregir o dar cuenta a la máquina de falsos positivos o verdaderos negativos buscando un aprendizaje activo de la máquina que permita perfeccionarse mientras interactúa con el usuario. A continuación, se muestran algunos estudios consultados bibliográficamente y sus características:

En la universidad Nacional del Este de Paraguay se desarrolló una interfaz la cual tiene como objetivo poder encender y apagar las luces de un recinto por medio de comandos de voz; para esto utilizaron el software de reconocimiento de voz *Dragon Naturally Speaking*, este interpreta los comandos dictados por parte del usuario y escribe la respuesta en el codificador de texto de la interfaz para posteriormente ser comparado con los comandos que ya están previamente guardados, si la instrucción dictada es correcta la acción será encender o apagar las luces. El autor reporta un funcionamiento del 98% en el reconocimiento de los comandos para realizar la acción de encendido y apagado de las luces del recinto.(Prieto, 2013)

NaviSAW y *MailSAW* es una herramienta multimodal la cual ayuda a las personas con discapacidades visuales a obtener acceso a internet, estas interfaces fueron integradas con el sistema SAW en la Universidad de Extremadura en España. El desarrollo propone controlar el navegador web y el correo electrónico utilizando los sentidos auditivos y hápticos de manera dinámica, esto con el fin de que el usuario con discapacidad visual tenga la posibilidad de escoger como desea interactuar con la herramienta multimodal y poder realizar tareas básicas en el navegador o en el correo electrónico, a su vez se incluye un periférico llamado *mouseSAW* este cuenta con 2 celdas táctiles, la celda del lado izquierdo tiene como función informar donde está situado el puntero y la celda derecha proporcionara información sobre el elemento que en ese momento se tiene seleccionado. (Gonzalez Rodriguez et al., 2007)

En la facultad de ciencias físicas y matemáticas de la Universidad Nacional de Trujillo situada en la ciudad de Perú, se presentó un sistema de empotrado para el control de un brazo robótico, el cual tendría instrucciones por medio de comando de voz (arriba, abajo, derecha, izquierda, parar, abre y cierra), para alcanzar el éxito del proyecto utilizaron el micro controlador *ref. dsPIC33EP256GP502*, para el reconocimiento de voz se utilizaron las técnica de *RNAs* (neuronas artificiales) y el *PCA* (análisis de componentes principales) donde la unión de estas dos técnicas dan como resultado el 100% de precisión para el reconocimiento de cada uno de los comandos ya previamente grabados. (E. A. P. De et al., 2012)

Finalmente, Lucía Cahuatijo de la Universidad Tecnológica Equinoccial, desarrolló un sistema que permite a personas con discapacidades abrir y cerrar una puerta mediante comandos de voz. Se utilizaron redes neuronales artificiales para el reconocimiento de los comandos dictados, con un entrenamiento de 50 muestras de voces diferentes, concluyendo que es recomendable utilizar una herramienta profesional para capturar la señal de voz; ya que por medio del micrófono de computador se generan más errores porque la señal no es limpia. (Ciencias De La Computación, 2013)

1.3 FORMULACIÓN DEL PROBLEMA

Uno de los avances tecnológicos importantes como lo ha sido la computadora, se convirtió en una herramienta esencial por las industrias, universidades, centros de investigación o simplemente en los hogares. Esta máquina con el pasar del tiempo ha desarrollado programas útiles para las diferentes actividades que se realizan a diario, sin embargo, podremos evidenciar la falta de actualización a los periféricos de interacción existentes entre el humano y la máquina, esto se podrá ver reflejado en las limitaciones que se crean para las personas con discapacidades, por este motivo surge la necesidad de desarrollar una interfaz que realice una interacción natural con el usuario y así obtener una experiencia satisfactoria a la hora de manipular la computadora. (Sánchez et al., 2020)

Por otro lado, una de las principales tareas que permite el Internet es la comunicación e interacción en tiempo real con el mundo cibernético, destacándose actualmente la aplicación WhatsApp (Zhang et al., 2020), la cual está integrada en casi todos los dispositivos móviles y cuenta con una interfaz para ser comandada desde un computador de escritorio, buscando una automatización de esta aplicación mediante comandos de voz podemos plantear el problema que se afrontará en este trabajo de grado como:

¿Cuál es la configuración de una interfaz de computador controlada por comandos de voz que permita a un usuario realizar tareas de comando básicas en el software de escritorio WhatsApp desktop?

2. JUSTIFICACIÓN

Con el pasar del tiempo la computadora se ha vuelto una de las principales herramientas que brindan una ventaja competitiva en las empresas ya que permiten hacer múltiples procesamientos en simultáneo aportando así una mayor fluidez y alta velocidad a la hora de realizar tareas repetitivas o de alta dificultad, en la actualidad el humano está rodeado de máquinas para las cuales es necesario desarrollar interfaces más amigables con el fin de que puedan facilitar la interacción con las personas mediante una mejor experiencia de usuario (C. De & De Empresas, 2019)

Este proyecto tiene como fin desarrollar una interfaz para ayudar a las personas que han sufrido una lesión raquímedular a nivel cervical entre las vértebras C3 y C6 comprometiendo la médula espinal y obstaculizando la transmisión del influjo nervioso, lo cual ocasiona una parálisis que afecta sus cuatro extremidades. Sin embargo, algunos pacientes con este diagnóstico han demostrado una respuesta verbal y ocular igual a un humano sano promedio, lo cual señala que se encuentran orientados en espacio y tiempo (Alexandra & Bejarano, 2020)

Con el desarrollo de este trabajo de grado la Fundación Universitaria Lumen Gentium Unicatólica generará nuevos conocimientos sobre los procesos de interacción entre humano-máquina, también creará un impacto social y de inclusión que se imparte desde la universidad, permitiendo socializar la investigación con otras universidades y dando a conocer el interés que se tiene por mejorar la calidad de vida de personas con limitaciones motrices.

3. OBJETIVOS

3.1 OBJETIVO GENERAL

Desarrollar una interfaz para computador que permita a personas con limitaciones motrices realizar tareas básicas de navegación en la aplicación WhatsApp desktop mediante comandos de voz.

3.2 OBJETIVOS ESPECÍFICOS

1. Determinar mediante una revisión bibliográfica las principales características de una interfaz de computador comandada por voz.
2. Implementar las herramientas software necesarias para controlar tareas básicas de navegación en la aplicación WhatsApp desktop.
3. Establecer un protocolo de pruebas para definir los alcances y limitaciones del sistema.

4. MARCO REFERENCIAL

4.1 ESTADO DEL ARTE

A continuación, se presenta los trabajos de investigación y productos comerciales encontrados en la revisión bibliográfica realizada. Las búsquedas fueron efectuadas utilizando GOOGLE SCHOLAR, IEEEEXPLORE Y SCIELO, utilizando las ecuaciones de búsquedas: interfaz comandada por voz, interfaz para navegador por medio de comandos de voz y automatización de sistemas de control por comandos de voz.

4.1.1 Referencias académicas

4.1.1.1 JNAVO Se enfoca en utilizar la voz para realizar comandos y representar acciones en un navegador web, este se pensó especialmente para personas invidentes, pero podrá ser usado por cualquiera persona que lo desee, su labor principal será realizar tareas básicas tales como atrás, adelante, detener y actualizar en el navegador, este tiene un segundo objetivo, ser acogido por tecnologías en el campo de reconocimiento y síntesis de voz y el campo de la Domótica. (David & Pérez, 2011)

Alcances:

- Presenta un análisis de una solución tecnológica para ayudar a personas invidentes.
- Se logra crear una especificación de comandos de voz aceptados por el sistema.
- Se da a conocer un diseño con técnicas que cumplen con los requerimientos demandadas por la necesidad de estas personas.

Limitaciones:

- El motor Microsoft speech tiene más facilidades y API'S para lenguaje como lo es C# o Visual Basic.
- Variación de la voz en la persona, ya que esta cambia a partir de su estado de ánimo o postura
- No tener un propio motor de reconocimiento y síntesis en Java.

Resultados:

Se realizaron pruebas de dictado de comandos de voz a ocho individuos con perfiles diferentes, incluyendo a una persona invidente. En lugares donde existía mucho

ruido, los resultados fueron aceptables debido a que la interfaz no logra reconocer todos los comandos dictados por el usuario. Se obtuvo un resultado del 2.3/5.0 en el funcionamiento del reconocimiento de dicha prueba. Por otra parte, se realizó el mismo dictado de comandos, pero en condiciones donde no había ruido ambiente, la calificación fue de 4.5/5.0 dando así un excelente funcionamiento.

4.1.1.2 Modelaje de una interfaz computarizada controlada por voz para cuadripléjicos que contribuya al desarrollo de aplicaciones que permitan comandar elementos de domótica en su domicilio (Enrique & Vivanco, 2017)

Esta interfaz es controlada por comando de voz para personas que tengan o han sufrido una lesión raquímedular, en la cual se presenta la pérdida de la sensibilidad y del movimiento en las extremidades (Moreno García, 2015). Mediante el procesamiento de la voz esta investigación propone comandar elementos de domótica en su domicilio, reduciendo la dependencia frente a otras personas y mejorando la calidad de vida de las personas al generar una sensación de autonomía.

La domótica permitirá la comunicación entre el usuario y un sistema de objetos conectados que forman una vivienda inteligente para esto el requisito mínimo sería que estos estén conectados a internet, el sistema domótico recopila información precedente de estos objetos y emitirá una orden automatizada para que estos la realicen por sí solos. La interfaz de usuario propuesta se compone de tres etapas:

- Alarma y mensajes:

Esta etapa deberá mostrar al usuario los errores o advertencias que puedan existir mientras se esté utilizando la interfaz, además realiza una realimentación constante al usuario indicando los comandos que han sido identificados y ejecutados.

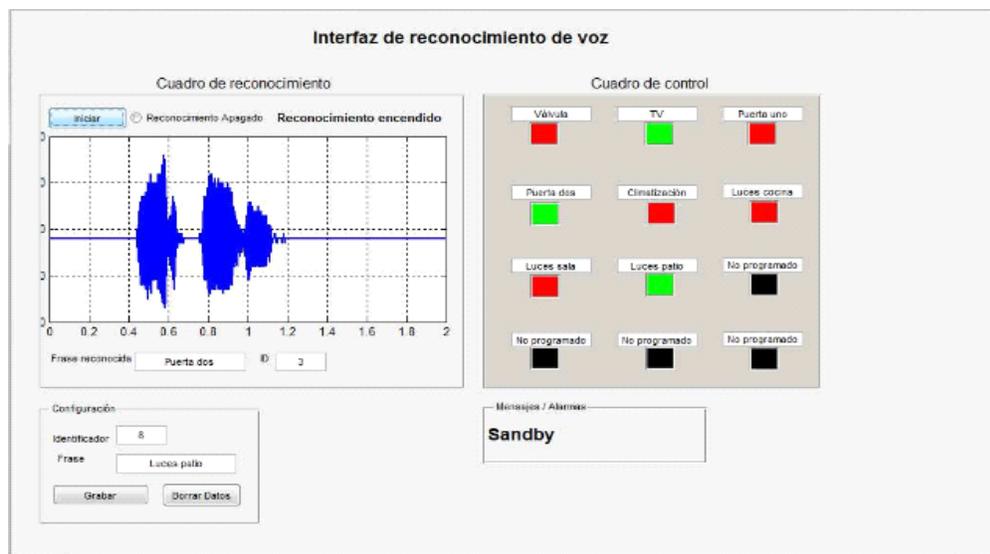
- Etapa de reconocimiento:

En esta etapa se realiza el procesamiento de las señales de voz para identificar los comandos pronunciados por el usuario. El sistema obtiene la voz por medio de un micrófono y se utilizó la tarjeta de sonido para digitalizar la señal de voz, esta fluctúa entre 100 Hz y 3.5 KHz con una frecuencia de muestreo de 11025 Hz a un tiempo de 2 segundos, posteriormente se utilizó como técnica de extracción de características de coeficiente *cepstrales* al obtener la energía central de cada frecuencia dada por el voz del usuario y realiza una clasificación mediante umbrales de energía definidos previamente para cada comando entrenado.

- Etapa de control:

Esta etapa es la encargada de realizar el accionamiento de cada comando una vez ha sido reconocido, en la Figura 1 se presentan los objetos que puede controlar la interfaz desarrollada, donde se observa que al lado izquierdo se presenta la señal de voz capturada por el sistema y en el lado derecho se encuentra el cuadro de control que tiene como función mostrar si cada dispositivo esta encendido (verde), apagado (rojo) o está en estado disponible para asignar una tarea (negro).

Figura 1. Etapa de control del marco referencial 4.1.1.2.



Alcances:

- Se desarrollo una interfaz para mejorar la calidad de vida de personas tetraplégicas mediante aplicación domótica controlada por voz.
- El sistema presentó un resultado del 98.6% para el reconocimiento del diccionario de comandos definidos, corroborando el buen funcionamiento del método de extracción de coeficiente de la escala de *MeI* utilizado.

Limitaciones:

- Se debe de conectar una tarjeta de serial de salida digitales con 9600 baudios de velocidad.
- Se debe de tener un micrófono externo de buena calidad y trabajar sin ruido ambiente para que no se distorsione la voz capturada.

Resultados:

La interfaz funciona de forma correcta y este fue el listado de comandos para acciones que se desarrollaron:

- Abrir o cerrar una llave de agua
- Encender/apagar el tv
- Encender/apagar luces sala
- Encender/apagar luces cocina
- Activar/Desactivar el movimiento de la silla de ruedas

La interfaz se configuró para 4 personas, esta arroja un resultado de 4.2/5.0 promediado entre las personas que realizaron la prueba y obtuvo un 98,6% de aciertos en los comandos evaluados.

4.1.1.3 Automatización de una silla de ruedas controlada por comandos de voz.

En este proyecto se estudió las discapacidades motoras que impiden el desplazamiento de un sitio a otro, estas son obligadas a depender de otras o de dispositivos externos que le puedan brindar la ayuda de moverse, este trabajo quiere contribuir con el diseño de una silla de ruedas la cual podrá ser manejada por comando de voz utilizando los comandos: adelante, atrás, derecha e izquierda, (Pardo, A., Rubiano, Z., & Montagut, Y, 2003)

El módulo de reconocimiento que utilizó el sistema es el *VOICE DIRECT 365* este es de la compañía Sensory Circuit, incluye un Procesador Digital de Señales (DSP) que es el encargado de procesar la señal mediante una red neuronal perceptrón multicapa programada en la memoria EEPROM el cual guardó los patrones que ya fueron previamente entrenados, la capacidad para este módulo fue de 15 palabras.

Para esta interfaz con la silla de ruedas se usó un microcontrolador PIC 16f84A producido por la empresa Microchip, su tarea era recibir las señales provenientes del DSP y accionar los motores de la silla de ruedas utilizando la técnica de control de velocidad modulación por anchura de pulso (*PWM*).

Alcances:

- El diseño de la silla propuesta presenta un impacto social importante para las personas con limitaciones motrices, ya que permite coordinar los movimiento de un aparato electrónico sin tener contacto físico con este mediante el procesamiento de comandos de voz.

Limitaciones:

- El reconocimiento de voz implementado no presentó un funcionamiento del 100%, ya que confunde palabras que puedan ser similares o confundir las palabras reconocidas.

Resultados:

La voz puede ser una gran salida para ayudar a estas personas que cuentan con una dificultad física de desplazamiento, se generó esta silla de ruedas como una solución funcional para la comunicación entre humano y máquina la cual le garantizara una mayor sensación de independencia al usuario y a su vez cumpliendo con algunas de las tareas de control físico que no podían realizar.

4.1.1.4 Domotic Voice

Esta investigación desarrolló una interfaz denominada Domotic Voice la cual tiene como fin realizar el reconocimiento de voz para las órdenes del sistema domótico *LOXONE* el cual se encuentra ubicado en las instalaciones civiles de la Universidad Politécnica Salesiana Sede Cuenca. El desarrollo de esta tecnología ayudó a personas con discapacidades visuales o motoras a realizar tareas básicas tales como abrir y cerrar una puerta o una ventana, encendido y apagado de equipos dentro de la vivienda, selección de temperatura ambiente, encendido y apagado de las luces, motorizar persianas para abrir y cerrarlas; todo esto se logró hacer con la siguiente selección de equipo (*Álvarez & Armando, 2016*):

La tarjeta seleccionada fue la EasyVR 2.0, la cual brinda la capacidad de reconocer de hasta 6 idiomas y soporta hasta 32 voces de usuarios, para la configuración de este se necesitan dos programas que se pueden encontrar gratis en internet uno es el EasyVR-Commander que permite grabar y reconocer los comandos pronunciados por el usuario y el otro programa es QuickSynthesis5 que es encargado de realizar una síntesis de voz para realimentar al usuario sobre el comando reconocido actualmente.

A nivel de software y hardware de uso libre pensando en la masificación del producto, puesto que no se tendrán que adquirir ninguna licencia para la programación del microcontrolador, en este orden de ideas, se seleccionaron los módulos de Arduino; Arduino uno y Arduino nano, los cuales disponen de un microcontrolador *ATmega328P* y un *ATmega328*.

El protocolo de comunicación entre el sistema domótico y el sistema de reconocimiento de voz fue *ZigBee*, este permite realizar una red inalámbrica de corta distancia. Para la automatización de la vivienda se configurarán dos módulos

de XBee lo cual funcionarán como maestro y esclavo para general la comunicación, la distancia de alcance de los módulos es de 30 metros en interiores y de hasta 100 metros con una antena dipolo. El Maestro será el primer módulo, este va a contener las órdenes que reconoce el sistema *Domotic Voice* para activar el servidor de domótica y esclavo será el segundo módulo que tendrá como fin recibir las órdenes provenientes del XBee maestro, esto para que haga las labores de un *gateway* que sirve de intermediario entre el reconocimiento de voz y el servidor domótico.

Finalmente, para alimentar el sistema se seleccionó una batería *Turnigy de Lipo* recargable de dos celdas, con un voltaje de 7.4 voltios y una corriente de 800mA; esta batería es pequeña, ideal para este tipo de proyectos ya que con esta se alimentará el módulo de Arduino uno y este alimentara a su vez a los demás componentes del reconocimiento de voz, es ligera y fácil de cambiar al momento de realizar un mantenimiento.

Alcances:

- En las pruebas realizadas se concluyó que el sistema *Domotic Voice* posee una diferencia de 0.6 milisegundos con respecto a la aplicación "Control por voz gratis" de Android, lo cual hace que este sistema ejecute las órdenes en un tiempo menor a esta aplicación desarrollada para dispositivos Android.

Limitaciones:

- La voz femenina, posee una diferencia notoria al resto de espectros en comparación de la masculina, lo cual se debe a un tono de voz más agudo que provoca que el cálculo de la distancia euclidiana esté alejada a los coeficientes de la voz del patrón, lo cual provoca que el sistema sea incapaz de reconocer órdenes que contengan especialmente vocales cerradas, ya que estas son más distorsionadas en una voz aguda.
- Los ruidos ambientales pueden cambiar la diafonía de las palabras impidiendo que el sistema relacione bien los comandos dictados con los que ya se encuentran guardados en la base de datos.
- Se necesita de llevar siempre un micrófono consigo para poder emitir las ordenes que deberán ser ejecutas por el sistema domótico.
- Se debe de entrenar el sistema para que pueda entender las órdenes dictadas por una persona que tenga una tonalidad de voz independiente

Resultados:

Se obtuvo un rango de 2.3 dB entre el valor mínimo de 41 dB y el valor máximo de 69.4 dB medidos en condiciones normales, es decir cuando el laboratorio se encontraba de manera solitaria donde las condiciones son las deseadas, para esta situación Domotic Voice tiene un funcionamiento correcto, los comandos dictados son soportados y ejecutados correctamente. Pero si se pone a prueba en unas condiciones del valor mínimo de 74.6 dB y el valor máximo 86.3 dB el sistema comienza a tener fallas, es decir si hay muchas personas dentro del laboratorio y todas están hablando, Domotic Voice no sabrá cual será la orden que se le está dando interpretando todas como una orden y haciendo así que el sistema colapse.

4.1.2 Referencias comerciales

4.1.2.1 Dragon Naturally Speaking

Es el software de reconocimiento de voz líder del sector, este lo usan los médicos, abogados y distintos profesionales para convertir un discurso en texto utilizando un menor esfuerzo con gran velocidad y precisión (Nuance Communications, 2020).

Este software contiene otras funcionalidades tales como actualizaciones de estado y búsqueda en internet, a su vez permite la opción de controlar el computador por comandos de voz, es una muy buena característica para crear órdenes con las cuales se disminuye el tiempo dedicado a la documentación, estos procesos por lo general son realizados por un agente virtual, que realiza procesos repetitivos y puede programarse para que realice la creación de documentos, hojas de cálculo, presentaciones, rellenar informes basados en formularios utilizando la voz.

4.1.2.2 SpeecheXter

Es una aplicación profesional de voz a texto multilingüe y gratuita destinada a ayudar con la transcripción de documentos, libros, informes, publicaciones de blog, entre otros mediante la voz. El diccionario personalizado de SpeecheXter permite agregar comandos cortos para insertar datos de uso frecuente tales como signos de puntuación, números de teléfono y direcciones. (*SpeecheXter | Type with Your Voice!*, 2020)

Este software es de especial uso para las personas que puedan tener alguna dificultad para usar sus manos, que presenten dislexia o discapacidades que limitan el uso de las entradas convencionales de información mediante el teclado. Esta tecnología es compatible con el navegador Chrome, permitiendo minimizar esfuerzos de escritura y aprender la correcta pronunciación de las palabras en el idioma escogido. También permite crear notas de texto, correos electrónicos, redacción de libros y publicaciones de blog.

4.1.2.3 Speechlogger

Es un software de reconocimiento de voz y traducción en la web. Este software cuenta con un sistema de puntaje o puntuación automática, guardado automático, marcas de tiempo, capacidad de edición de texto, transcripción de archivos de audio y opciones de exportación utilizando el API de reconocimiento de Google (*Software de Reconocimiento Del Habla y Traducción Instantánea*, 2020).

La interfaz *cloud speech* api permite convertir audio en texto en más de 80 idiomas, todo esto se hace recibiendo el audio por medio del micrófono del dispositivo receptor. El software funciona utilizando una red neuronal de gran alcance que se enfoca en el procesamiento del lenguaje natural y soporta audio en distintos tipos de formatos de compresión tales como *FLAC*, *AMR*, *PCMU* y *linear-16*.

4.1.2.4 Bing speech api

Este API de Microsoft ofrece algoritmos que permiten integrar el reconocimiento de voz en las aplicaciones que corren en sistema operativo Windows de forma nativa y contiene soporte para 48 lenguajes distintos (*Bing Speech API | Microsoft Azure*, 2020).

También se podrá crear aplicaciones y servicios que hablen naturalmente con el usuario (*Chatbot*) mejorando la experiencia del usuario al utilizar esta tecnología, contiene sintetizadores de voz con voces y permite realizar traducciones automáticas en tiempo real, transcribiendo y traduciendo conversaciones de la vida real, esto es compatible con aplicaciones móviles, de escritorio y web impulsando una solución rápida y precisa para este requerimiento.

4.2 MARCO TEÓRICO

A continuación, se presentan los apartados teóricos y conceptuales que enmarcan el desarrollo de este trabajo de grado:

4.2.1 Interacción Humano-Máquina

La interacción que se lleva a cabo entre humano-máquina es una doctrina que se enfoca en definir sistemas computacionales que apoyan a las tareas diarias del ser humano; el mundo de la tecnología avanza muy rápido, por lo tanto, trae consigo cambios importantes con los cuales se han podido desarrollar varias formas de interacción, tales como interfaces táctiles, hápticas, gestuales y de reconocimiento de voz. En otras palabras, la interacción humano-máquina brinda al usuario una

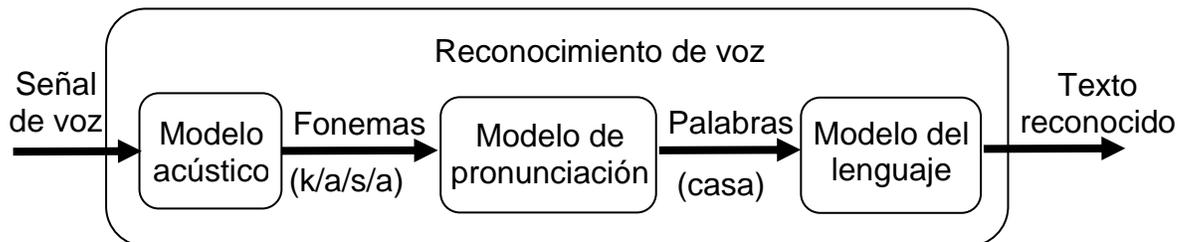
forma de interactuar más natural y novedosa a diferencia del formato clásico como el mouse y teclado.

La interacción entre el humano y las máquinas depende de un intercambio de información en ambas direcciones entre el operario y el sistema, usualmente se considera que el operario controla todas las acciones de la máquina por medio de la información que introduce y las acciones que realiza sobre esta, pero también es necesario considerar que el sistema realimenta con cierta información al usuario por medio de distintas señales, para indicar el estado del proceso o las condiciones del sistema (Castro-molinares & Ph, 2019).

4.2.2 Reconocimiento de Voz

Corresponde al proceso de capturar señales de voz generadas por el usuario a través de un micrófono y transformarlas a texto. Este sistema se compone de tres modelos como se presenta en la Figura 2, en la cual el *modelo acústico* inicia con un comportamiento probabilístico al recibir la señal de voz y como resultado muestra la distribución de los fonemas reconocidos; seguido el *modelo de pronunciación* toma estos fonemas y realiza una búsqueda dentro de un diccionario con un idioma establecido para formar las palabras. Finalmente, para que el texto reconocido tenga un sentido lógico, el *modelo del lenguaje* organiza las palabras y muestra la secuencia que más se ajusta a la señal de voz.(J. Pérez et al., 2017)

Figura 2. Arquitectura interna de un sistema de reconocimiento de voz

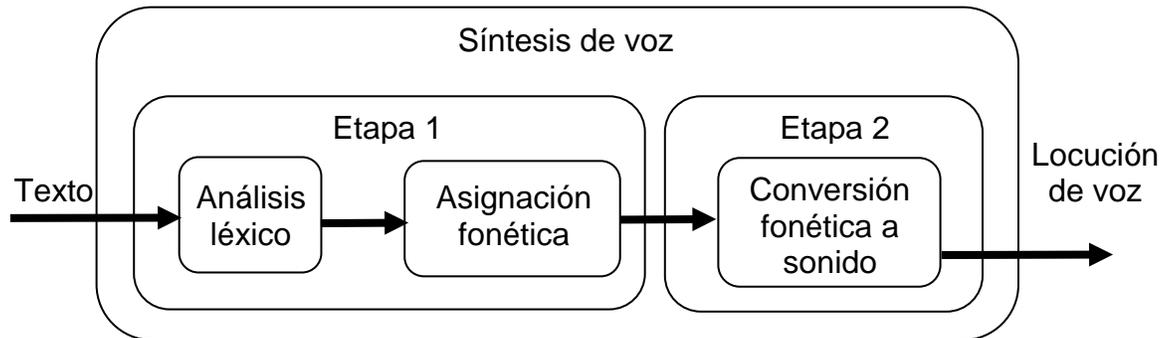


4.2.3 Síntesis de voz

Su funcionalidad es la de convertir un texto escrito en una locución artificial de voz mediante una concatenación de sonidos pregrabados. La efectividad de lograr una similitud en las características de la voz depende de su almacenamiento de palabras enteras u oraciones que dan el contexto a la frase; su estructura normalmente está compuesta por dos etapas internas como se presenta en la Figura 3, en la cual se puede observar que la primera etapa se divide en dos bloques, donde el texto ingresado es convertido en subunidades fonéticas gracias a un *análisis léxico*, seguido del bloque de *asignación fonética* donde se reconocerá las sílabas, hiatos,

diptongos, etc. Y Finalmente, la segunda etapa unificará las divisiones fonológicas para formar las palabras y convertirlas en sonidos controlando el tono y la duración de los fonemas para dar una síntesis de mayor calidad. (Barrobés & Ruiz, 2017)

Figura 3. Arquitectura interna de un sintetizador de voz.



4.2.4 Sistema Operativo

Es un conjunto de programas especialmente coordinados y dirigidos para todos los servicios que el usuario desee realizar, este permite regular los aspectos básicos del sistema a su vez se encarga de administrar todos los periféricos de una computadora este es abiertamente el encargado de mantener la integridad del sistema. Para tener una visión del apartado a continuación se muestran unos ejemplos de sistemas operativos para PC:

- Microsoft Windows
- GNU/Linux
- Mac OS X
- Google Chrome OS
- Ubuntu

4.2.5 Acceso a la información para personas con limitaciones motrices.

El desarrollo de este proyecto de investigación es fundamental para mejorar la calidad de vida de las personas tetraplégicas, la cuales han perdido una de las partes más importantes de la vida “la autonomía”, la libertad de poder realizar acciones valiéndose por sí mismas y poder acceder a la información del mundo mediante Internet (Rashid & Islam, 2017).

4.3 MARCO CONCEPTUAL

La interfaz propuesta mejora la calidad de vida de las personas permitiéndoles controlar la aplicación WhatsApp Desktop mediante el procesamiento de la voz humana. Por esta razón es necesario entender los conceptos que se relacionan con el funcionamiento de un sistema de reconocimiento de voz genérico.

4.3.1 Captura y pre-proceso

La información del mensaje hablado es procesada después de ser muestreada con una frecuencia de muestreo promedio de 11025Hz y cuantificada en 16 o 32 bits por muestra, respetando el teorema de Nyquist (Muestrear como mínimo al doble de la menor frecuencia del sistema). Luego se realiza un acondicionamiento de la señal para eliminar señales no deseadas como el ruido ambiental, dejando solo como información útil la voz de la persona. Algunas técnicas de pre-proceso son la extracción espectral de ruido y la detección automática de segmentos hablados. Adicionalmente se pueden ejecutar operaciones sobre los datos en dos dominios diferentes, en el tiempo y en el dominio frecuencial mediante la transformada de Fourier, ya que ambos dominios presentan cualidades propias que pueden ser beneficiosas para la futura extracción de características (Enrique & Vivanco, 2017).

4.3.2 Extracción de características

De la serie de datos adquiridos se extrae información relevante de características altamente discriminatorias permitiendo hacer diferenciaciones y clasificaciones. Algunas de las características importantes extraídas de la voz humana son coeficientes de predicción lineal que modelan el tracto vocal mediante un sistema auto-regresivo; los coeficientes de energía, delta y aceleración que brindan información importante del fragmento de voz considerado cuando se hace una segmentación mediante fonemas; por último los coeficientes cepstrales los cuales representan el espectro corto de potencia de una señal basada en la transformada del coseno del espectro de potencia (Mosquera et al., 2011).

4.3.3 Clasificación y decisión

Generalmente en el estado del arte de sistemas de reconocimiento de voz se usan los modelos ocultos de Markov (HMM) los cuales describen modelos estadísticos con transiciones de probabilidad. Estos modelos estadísticos son útiles para el análisis de una serie de observaciones en tiempo discreto, tales como un flujo de muestras acústicas extraídas de una señal de voz. Dichos modelos son muy apropiados para el reconocimiento de voz debido a su capacidad inherente de representar eventos acústicos de duración variable y a la existencia de algoritmos

eficaces para computar automáticamente los parámetros del modelo a partir de los datos de entrenamiento.

Por otra parte, las Redes Neuronales Artificiales, (RNA) que simulan el aprendizaje cognitivo humano a nivel neuronal en sistemas de capas externas y ocultas mediante redes interconectadas masivamente en paralelo de elementos simples (usualmente adaptativos), intentan interactuar con los 16 objetos del mundo real del mismo modo que lo hace un sistema nervioso biológico (Caicedo et al., 2009). Mediante entrenamiento se puede enseñar a la red neuronal artificial a discriminar eventos entre diversas clases. Otro de los principales métodos para el reconocimiento del habla es conocido como método de plantillas. Una idea clave en el método de las plantillas es obtener secuencias de tramas habladas para conformar un patrón, por ejemplo, una palabra, para luego medir distancias espectrales y comparar los patrones.

Otra alternativa es la de usar alguna forma de programación dinámica para alinear temporalmente los patrones y proceder a contar diferencias en la velocidad del habla, tanto a través de diferentes hablantes, así como de diferentes repeticiones de las palabras. La metodología de la aproximación por plantillas está bien desarrollada y provee un buen desempeño para una variedad de aplicaciones prácticas (Rabiner et al, 1993). Entre estas técnicas se encuentran la del clustering de K vecinos más cercanos la cual se basa en la suposición de que los prototipos más cercanos tienen una probabilidad a posteriori similar. La desventaja de este método radica en que la estimación resultante no es una función de densidad de probabilidad verdadera y que los datos de entrenamiento deben mantenerse almacenados (Nope et al, 2008).

4.4 MARCO CONTEXTUAL

Este Proyecto se desarrolló en la ciudad de Santiago de Cali, fue realizado por estudiantes de Último Semestre de Ingeniería en Sistemas en la Fundación Universitaria Lumen Gentium Unicatólica. Este trabajo de grado es parte fundamental del proyecto de investigación de convocatoria interna *Asistente virtual para personas tetrapléjicas controlado por voz y gestos faciales (Primera Fase)*, se realizó la implementación de la interfaz utilizando herramientas de software libre en Python y las pruebas se realizaron con personas sanas, ya que no se pudieron realizar pruebas en personas tetrapléjicas debido a la emergencia Internacional Sanitaria del Coronavirus (COVID-19), sin embargo se logró una caracterización de los alcances y limitaciones de la interfaz propuesta el cual dejando claro que funcionará de forma correcta en personas con limitaciones motrices, todos los usuarios realizaron las pruebas de forma voluntaria aceptando el *Formato de Consentimiento Informado* desarrollado para esta investigación (ANEXOS).

4.5 MARCO LEGAL

Los investigadores participantes del proceso tendrán garantías plenas de equidad, respeto y justicia, en lo concerniente a todas las actividades relacionadas con el proyecto. Asimismo, se les asignará los recursos necesarios para realizar sus actividades de forma cómoda, sin tener inconvenientes por falta de recursos.

Los participantes de la investigación podrán cambiar de opinión en cuanto a la permanencia en el proyecto, en caso de que este afecte su salud física, mental o su dignidad. De la misma manera los investigadores tendrán el derecho de reconocimiento (derechos de autor) o créditos relacionados con el trabajo realizado en el transcurso del proyecto de investigación. Tampoco serán discriminados, por ninguna condición, en la selección para vincularse al proyecto.

Todas las actividades de investigación se realizarán con el consentimiento informado de padres de familia, directivos institucionales, los mismos sujetos de investigación e instancias pertinentes. También gozarán de pleno anonimato, los resultados de la investigación no los afectarán divulgando información sensible. De la misma forma, en el presente proyecto no se recopilan datos sensibles de identificación a los sujetos de investigación. Los sujetos de investigación podrán retirarse en cualquier momento de esta, sin que esto sea motivo de represalias, juzgamientos o burlas.

5. METODOLOGÍA

5.1 MÉTODO Y TIPO DE ESTUDIO

La metodología propuesta para el desarrollo del proyecto está basada en el método Inductivo, teniendo un enfoque cualitativo y con alcances experimentales. Se partirá de un estudio bibliográfico de sistemas similares o afines donde se empleen señales de voz para el comando de interfaces humano-máquina. Posteriormente, se definirán los requerimientos del sistema de procesamiento de voz para captar comandos y controlar la aplicación computacional WhatsApp Desktop. Se definirán protocolos de prueba para validar las hipótesis planteadas y se evaluará el desempeño general del sistema.

5.2 TÉCNICA E INSTRUMENTOS

Para el desarrollo del proyecto se empleó los siguientes dispositivos:

- Computador: Procesador de 1,8 GHz o superior. Doble núcleo o superior recomendado. 4 GB de RAM recomendado, Espacio en disco duro: hasta 130 GB de espacio disponible, en función de las características instaladas; las instalaciones típicas requieren entre 20 y 50 GB de espacio libre.
- Audífonos y micrófonos: En el proceso se utilizó una diadema inalámbrica que unifica estos dos dispositivos, pero no hubo problema al usarse por separado.
- Conexión a Internet.

A nivel de herramientas de software se utilizó:

- Sistema operativo Windows 10.
- Visual Studio Code 1.40.2
- Python versión 3.5
- Whatsapp de Escritorio.

Y por último las librerías de Python que se utilizaron para el desarrollo de la interfaz en el reconocimiento de voz fueron:

- SpeechRecognition: Librería para realizar el reconocimiento de voz, con soporte para varios motores y API, en línea y fuera de línea. (Python Speech Recognition, 2020)

- Pyaudio: Enlaces para **PortAudio v19**, librería de flujo de entrada / salida de audio multiplataforma. (Pham, 2020)

y para la Síntesis de voz:

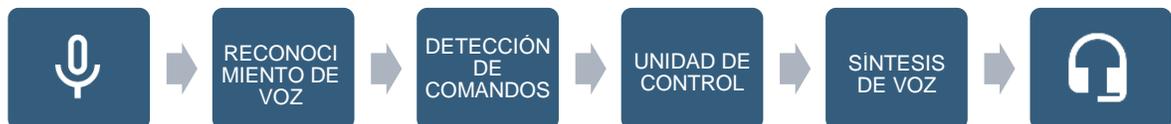
- Pyttsx3: Librería de texto a voz (TTS) para Python 2 y 3 la cual funciona sin conexión a Internet. (Bhat, 2020)

La biblioteca de Python ofrece una gran variedad de módulos que realizan distintos tipos de funciones, por esta razón luego de una comparación bibliográfica se escogieron estas librerías, principalmente por la facilidad de importarlas y la calidad del rendimiento en diferentes sistemas operativos. Se identificaron librerías robustas que utilizan algoritmos avanzados de redes neuronales para reconocimiento de voz, de las cuales se destacó *Cloud Speech-to-Text de Google (Cloud Speech-to-Text - Reconocimiento de Voz | Cloud Speech-to-Text, 2020)* por ser gratuita y de rápido acceso y configuración, siendo idónea para este proyecto.

5.3 ARQUITECTURA DEL SOFTWARE IMPLEMENTADO

La Figura 4 presenta el diagrama de bloques del procesamiento de la información que se extrae de las señales de audio, el primer bloque representa el reconocimiento de voz que se encarga de convertir la señal de voz en una cadena de texto. El segundo bloque detecta si la cadena de texto contiene o no los comandos establecidos para el manejo de la aplicación, de lo contrario, la interfaz asimila de que es un texto libre que se activa durante un dictado. En el bloque de unidad de control se ejecuta el comando de forma automática en el sistema operativo, finalmente el bloque de síntesis de voz genera una realimentación al usuario que confirma el comando identificado y ejecutado.

Figura 4. Diagrama de bloques general del sistema.



5.3.1 Reconocimiento de voz

Este es el bloque que se encarga de ser intermediario entre el usuario y el aplicativo, para su principal funcionamiento se utiliza la librería *SpeechRecognition* la cual constantemente escucha al usuario y reconoce el contenido del mensaje, retornando una cadena de texto que será utilizada para reconocer los comandos presentes en el mensaje. En la Figura 5 se presenta la configuración de la librería utilizada.

Figura 5. Ejemplo de implementación Librería *SpeechRecognition*.

```
import speech_recognition as sr

global r
r = sr.Recognizer()
global source
with sr.Microphone() as source:
    r.adjust_for_ambient_noise(source, duration = 1)
    audio = r.listen(source, phrase_time_limit = None)
    try:
        response = r.recognize_google(audio, language="es-CO")
        print("Entendi: " + response + "")
    except sr.UnknownValueError:
        print("No te entendi")
    except sr.RequestError as e:
        print("GSR; {}".format(e))
        print("No logre reconocer ningun audio")
```

La figura anterior muestra que para hacer uso del reconocimiento de voz es necesaria la librería *PyAudio*, la cual se encarga de crear los canales para *PortAudio* y le permite a Python grabar o reproducir audio. Esta herramienta empieza su reconocimiento con una toma de muestra del ruido ambiente mediante la función *adjust_for_ambient_noise* que cumple una función importante, ya que escucha el espectro de energía en el ambiente y permite tener una referencia de ruido que se deberá excluir más adelante. Al ser una acción parametrizable, se estableció (1s) un segundo de escucha, pues se identificó que entre más tiempo el proceso de reconocimiento de voz se torna lento y no ofrece una buena experiencia al usuario. La función *listen* obtiene el audio del micrófono y la función *recognize_google* es la encargada de reconocer el mensaje pronunciado por el usuario y esta necesita de una conexión a internet para poder reconocer la voz ya que funciona mediante el API de Google; una vez ejecutado, éste retorna la cadena de caracteres que corresponde a las palabras pronunciadas por el usuario "El texto reconocido".

5.3.2 Detección de comandos

Para la detección de comandos el primer paso es recibir la cadena de texto y compararla con los comandos establecidos, en este trabajo de grado se desarrolló un diccionario de 17 comandos los cuales son presentados en la primera columna de la Tabla 1.

Tabla 1. Lista de comandos y acciones.

RECONOCIMIENTO DE VOZ	SINTESIS DE VOZ	COMANDO EJECUTADO
abrir WhatsApp	“abriendo WhatsApp”	os.system('WhatsApp.exe')
arriba	“subiendo al siguiente chat”	'ctrl' + 'shift' + 'tab'
Abajo	“bajando siguiente chat”	'ctrl' + 'shift'
bajar	“bajando”	auto.scroll(-10)
subir	“subiendo”	auto.scroll(10)
minimizar	“minimizando WhatsApp”	'win' + 'm'
restaurar o maximizar	“abriendo WhatsApp”	'win' + 'shift' + 'm'
aumentar letra o acercar	“aumentando letra”	'ctrl' + '+'
reducir letra o alejar	“reduciendo letra”	'ctrl' + '-'
Nuevo	“creando nuevo chat”	'ctrl' + 'n'
Perfil	“mostrando perfil”	'ctrl' + 'p'
escribir + texto	“claro, escribiendo” + texto	auto.typewrite(texto)
atrás o escape	“atrás”	'esc'
eliminar conversación	“eliminando conversación”	ctrl' + 'backspace'
aceptar o Enter	“aceptar”	'enter'
chat de + nombre de la persona	“muy bien”	'ctrl' + 'f' auto.typewrite(nombre)
borrar todo	“muy bien”	'ctrl' + 'a' 'backspace'

A nivel algorítmico se usó la función *in()*, esta función es nativa de Python y no pertenece a ninguna librería, se utilizó con el objetivo de comparar una palabra o carácter en una cadena, esta retorna un valor booleano, donde *true* es si la palabra o carácter está contenida en la cadena **o false** si no lo está, un ejemplo de cómo funciona la función se presenta en la siguiente figura 6.

Figura 6. Ejemplo de implementación de la función *in ()*.

```
textol = "manzana"
texto2 = "manzana"
resultado = textol in texto2;
print(resultado)
>>> True
```

5.3.3 Control del sistema operativo

Este módulo de interacción permite comandar el computador sin tener contacto físico, emulando presionar combinaciones de teclas (las cuales son generadas como falsas interrupciones al sistema operativo). Las combinaciones de teclas asignadas para cada uno de los comandos implementados se presentan en la columna 3 de la Tabla 1.

A nivel algorítmico para implementar el control en el sistema operativo se utilizaron dos librerías; la primera es la librería *os* que pertenece a la biblioteca estándar de Python, una de sus funciones llamada *system* permitió ejecutar la aplicación de WhatsApp Desktop. Luego de tener la aplicación de escritorio abierta se utilizó la librería *pyautogui*, de la cual se utilizaron las funciones *press* y *hotkey* que emulan la acción de presionar las teclas de forma virtual; Las teclas soportadas se describen en la Figura 7. (Harrison, 2018; PyAutoGUI · PyPI, 2020)

Figura 7. Lista de teclas soportadas por la Librería *pyautogui*.

```
[ '\t', '\n', '\r', ' ', '!', '"', '#', '$', '%', '&', "'", '(',  
)', '*', '+', '-', '.', '/', '0', '1', '2', '3', '4', '5', '6', '7',  
'8', '9', ':', ';', '<', '=', '>', '?', '@', '[', '\\', ']', '^', '_', '`',  
'a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j', 'k', 'l', 'm', 'n', 'o',  
'p', 'q', 'r', 's', 't', 'u', 'v', 'w', 'x', 'y', 'z', '{', '|', '}', '~',  
'accept', 'add', 'alt', 'altleft', 'altright', 'apps', 'backspace',  
'browserback', 'browserfavorites', 'browserforward', 'browserhome',  
'browserrefresh', 'browsersearch', 'browserstop', 'capslock', 'clear',  
'convert', 'ctrl', 'ctrlleft', 'ctrlright', 'decimal', 'del', 'delete',  
'divide', 'down', 'end', 'enter', 'esc', 'escape', 'execute', 'f1', 'f10',  
'f11', 'f12', 'f13', 'f14', 'f15', 'f16', 'f17', 'f18', 'f19', 'f2', 'f20',  
'f21', 'f22', 'f23', 'f24', 'f3', 'f4', 'f5', 'f6', 'f7', 'f8', 'f9',  
'final', 'fn', 'hanguel', 'hangul', 'hanja', 'help', 'home', 'insert', 'junja',  
'kana', 'kanji', 'launchapp1', 'launchapp2', 'launchmail',  
'launchmediaselect', 'left', 'modechange', 'multiply', 'nexttrack',  
'nonconvert', 'num0', 'num1', 'num2', 'num3', 'num4', 'num5', 'num6',  
'num7', 'num8', 'num9', 'numlock', 'pagedown', 'pageup', 'pause', 'pgdn',  
'pgup', 'playpause', 'prevtrack', 'print', 'printscreen', 'prntscrn',  
'prtsc', 'prtscr', 'return', 'right', 'scrolllock', 'select', 'separator',  
'shift', 'shiftright', 'shiftright', 'sleep', 'space', 'stop', 'subtract', 'tab',  
'up', 'volumedown', 'volumemute', 'volumeup', 'win', 'winleft', 'winright', 'yen',  
'command', 'option', 'optionleft', 'optionright']
```

Para ejemplificar mejor el uso de estas librerías, en la Figura 8 se presenta un ejemplo algorítmico el cual realiza 3 acciones; la primera es abrir la aplicación WhatsApp, la segunda es presionar la tecla de 'esc' (escape) para retroceder en el aplicativo y por último una combinación de teclas para eliminar.

Figura 8. Ejemplos de uso de la Librería *PyAudio*.

```
67 def abrir1():
68     os.system('WhatsApp.exe')
69
70 def escape():
71     auto.press('esc')
72
73 def borrar():
74     auto.hotkey('ctrl', 'backspace')
75
```

En la figura anterior, la función `os.system()` permite ejecutar comandos dados en forma de cadena de texto en una *subshell* y definir la ruta de la aplicación como una variable del sistema (*Python | Os.System() Method - GeeksforGeeks, 2020*). La función `pyautogui.press()` permite ejecutar de manera remota sin presionar físicamente una tecla específica, se puede presionar también un conjunto de teclas, esto se realiza mediante la función `hotkey` con la cual podemos emular todo el teclado (*Keyboard Control Functions — PyAutoGUI Documentation, 2020*)

5.3.4 Síntesis de voz

Este algoritmo es el encargado de dar una realimentación auditiva por parte del aplicativo y va dirigida al usuario que le da a entender el comando que fue reconocido y ejecutado. La lista de frases sintetizadas por cada comando reconocido se presenta en la columna 2 de la Tabla 1. Su funcionamiento necesita de una cadena de texto la cual será pronunciada por el sintetizador, en la Figura 9 se presenta un ejemplo de configuración de la librería utilizada

Figura 9. Ejemplo de implementación de la Librería *Pytttsx3*.

```
import pyttsx3
tts = pyttsx3.init()

def hablar(nombre):
    tts.say("Hola")
    tts.say("como estas "+nombre+"?")
    tts.runAndWait()

hablar('Daniel')
```

Como se observa en la figura anterior, para hacer uso de la librería es necesario realizar su inicialización con la función `pyttsx3.init()` la cual permitirá crear un entorno de vocalización, luego la función `say()` permitirá vocalizar y estructurar el texto que se va a sintetizar.

En la Figura 9 se puede evidenciar que esta herramienta permite pronunciar más de una línea usando la función *say()*, e internamente creará una cola de textos que estarán a la espera de ser reproducidos por la función *runAndWait()*, esta última realizará la confirmación que liberará el texto sintetizado; si existe una cola de textos se reproducirá una pequeña pausa entre cada uno de los textos pronunciados.

6. PRUEBAS Y RESULTADOS

6.1 PROTOCOLO DE PRUEBAS

Inicialmente se realizó un análisis de los sujetos de prueba mediante encuestas para definir el grado de familiaridad con la navegación en Internet, la aplicación WhatsApp y el uso de reconocimiento de voz.

La primera prueba evaluó el desempeño del reconocimiento de comandos, se estudió un total de 26 comandos los cuales fueron repetidos tres veces por cada usuario, y se analizó el número de comandos acertados y a su vez buscar el comando con menos aciertos.

La segunda prueba permitió evaluar el reconocimiento de dictados, se definió previamente un mensaje el cual cada usuario repitió tres veces y permitió identificar las limitaciones de dicha herramienta, a su vez nos da a conocer posibles errores futuros y las palabras con dificultad de reconocimiento.

La tercera prueba evaluó el comportamiento de los usuarios ante una situación controlada, donde se busca una integración entre todos los bloques de la aplicación y simular un uso cotidiano, se pidió a los usuarios que siguieran una rutina donde ellos pudieran abrir la aplicación, enviar mensajes y navegar entre las distintas conversaciones.

Como último punto, se buscó realizar un análisis cualitativo donde se implementó una encuesta de 10 preguntas dirigida a los sujetos de prueba para evaluar su experiencia de usuario y percepción sobre el funcionamiento de la interfaz desarrollada.

6.1.1 SUJETOS DE PRUEBA

Las pruebas fueron realizadas a un grupo de siete personas, donde cada persona aprobó y firmó el formato de consentimiento informado utilizado para este trabajo de grado, el formato de consentimiento informado utilizado está incluido en los anexos de este trabajo de grado.

Inicialmente se realizó una encuesta para medir el grado de familiaridad de los usuarios con las temáticas tratadas en este trabajo de grado, estas encuestas se evaluaron mediante una escala tipo Likert (Matas, 2018; J. G. Sánchez & Terrats, 2011) en la cual las posibles opciones de respuesta eran: Nunca, Casi nunca, Algunas veces, Casi Siempre y Siempre.

Por otro lado, las preguntas realizadas incluían la edad, el uso diario de internet, WhatsApp y sistemas de reconocimiento de voz. En la Tabla 2 se presentan los resultados obtenidos:

Tabla 2. Perfil usuarios de prueba

Usuario	1	2	3	4	5	6	7
Edad	20	54	25	23	46	38	24
Internet	siempre	Algunas veces	siempre	siempre	Casi Siempre	Casi Siempre	siempre
WhatsApp	siempre	Algunas veces	siempre	siempre	Casi Siempre	siempre	siempre
Reconocimiento de Voz	Casi nunca	Algunas veces	Casi nunca	Algunas veces	Casi nunca	Algunas veces	Casi nunca

Se observó que el rango de edad de los usuarios oscila entre 20 - 54 años, además se calculó el porcentaje de familiaridad de los usuarios con internet, WhatsApp y reconocimiento de voz, la cual es presentada en la Tabla 3.

Tabla 3. Resultado porcentualmente sobre el grado de familiaridad de los usuarios de prueba.

Usuario	Nunca	Casi nunca	Algunas veces	Casi siempre	Siempre
Internet	0%	0%	14.29%	28.57%	57.14%
WhatsApp	0%	0%	14.29%	14.29%	71.43%
Reconocimiento de Voz	0%	57.14%	42.86%	0%	0%

En los resultados obtenidos se observa que el grupo de personas cuya edad es menor a 25 años tienen un uso de internet más amplio, con un total del 57.14% respecto al 42.86% del segundo grupo que tienen una edad superior a 38 años ya que la usan algunas veces en su día a día.

Las pruebas muestran que la mayoría de los usuarios usan siempre y casi siempre la aplicación de WhatsApp lo que da un total del 85.59%, pues se comunican a diario por motivos personales o laborales, por otro lado, el resto de las personas lo usan algunas veces, pero no dejan de usar la aplicación.

En los resultados podemos ver que los usuarios no usan con frecuencia los comandos de voz; con un 42.86% los usuarios votaron algunas veces para referenciar que la aplicación la usan en ocasiones específicas y esto se evidencia

en personas mayores de edad que constituyen el 28,57% de la opinión respecto al 14,28% que son los jóvenes.

6.1.2 RECONOCIMIENTO DE COMANDOS

En la prueba de comandos se solicitó a los voluntarios pronunciar los 26 comandos del aplicativo, se repitieron 3 veces cada uno de ellos evaluando un total de 546 comandos. En la Tabla 4 se presentan los resultados obtenidos por cada voluntario, en este figura se visualizan de color verde los comandos que se reconocieron de forma correcta y de rojo los comandos no reconocidos, en la última columna de la tabla se muestra el porcentaje total de aciertos por cada comando y el porcentaje total de la prueba global.

Tabla 4. Resultados de la prueba de comandos.

PERSONA NUMERO	1	2	3	4	5	6	7	ACIERTOS (%)
activar modo prueba	█	█	█	█	█	█	█	57.14
desactivar modo prueba	█	█	█	█	█	█	█	52.38
ajustar tiempo de escucha	█	█	█	█	█	█	█	100.00
modo ayuda	█	█	█	█	█	█	█	100.00
navegación	█	█	█	█	█	█	█	100.00
configuración	█	█	█	█	█	█	█	100.00
escritura	█	█	█	█	█	█	█	100.00
abrir WhatsApp	█	█	█	█	█	█	█	100.00
cerrar WhatsApp	█	█	█	█	█	█	█	95.24
arriba	█	█	█	█	█	█	█	100.00
abajo	█	█	█	█	█	█	█	90.48
bajar	█	█	█	█	█	█	█	76.19
subir	█	█	█	█	█	█	█	100.00
minimizar	█	█	█	█	█	█	█	95.24
restaurar o maximizar	█	█	█	█	█	█	█	85.71
aumentar letra o acercar	█	█	█	█	█	█	█	95.24
reducir letra o alejar	█	█	█	█	█	█	█	100.00

Es necesario mencionar que esta prueba se realizó bajo las mismas condiciones ambientales de 53,7 dB medidos con la aplicación Sonómetro (Sound Meter) (Google Play, 2020).

El dictado fue evaluado comparando palabra por palabra entre las reconocidas por el sistema y el dictado original contenido por 28 palabras. En la Tabla 5 se presentan los resultados obtenidos:

Tabla 5. Resultados de la prueba de reconocimiento de voz.

	Usuario 1	Usuario 2	Usuario 3	Usuario 4	Usuario 5	Usuario 6	Usuario 7	Total
Dictado 1	92%	92%	96%	92%	96%	96%	90%	93%
Dictado 2	96%	89%	92%	92%	96%	92%	96%	93%
Dictado 3	92%	96%	96%	89%	92%	92%	96%	93%
Total	93%	92%	95%	91%	95%	93%	94%	93%

En el reconocimiento del dictado se observó buenos resultados con un acierto global del 93% y con una desviación estándar del 2%; se evidencia que el voluntario con menor desempeño tuvo un resultado de un 91% de aciertos respecto al usuario con mejor desempeño que fue un 95%.

Por otro lado, se identificó que la palabra “WI-FI” se reconoció muy pocas veces debido a que es una palabra compuesta por siglas, evidenciando las dificultades que presenta la herramienta para reconocer abreviaturas o palabras en inglés si el usuario no tiene una buena pronunciación.

6.1.4 INTERFAZ FINAL CON WHATSAPP DESKTOP

En esta prueba se pidió a los voluntarios abrir la aplicación de mensajería mediante el comando “abrir WhatsApp”, donde el aplicativo responde de forma verbal que el proceso de abrir la aplicación se estaba ejecutando, como siguiente paso el usuario dicta el comando “enviar mensaje a Daniel”, donde el aplicativo responde preguntando por el mensaje que desea enviarle al contacto, seguido de esto se pide al usuario pronunciar la frase “¿hola, como estas?”, donde el aplicativo filtra por el nombre “Daniel” en la lista de contactos, pega el mensaje dicho por el usuario y finalmente envía automáticamente el mensaje.

Para un segundo caso se pidió al usuario eliminar la conversación mediante el comando “eliminar conversación”, el aplicativo responde de forma verbal que la conversación fue eliminada, luego se pidió navegar entre los distintos chats que

tenían iniciados, mediante el comando “ir al chat de Daniel” o simplemente “chat de Daniel”, y por consiguiente este fue dirigido hacia el chat del contacto; Ya dentro del chat presente se pidió usar la opción de enviar mensajes cortos, diciendo el comando “escribir” y seguido dictar el mensaje “¿hola, como estas?”, el aplicativo añade el mensaje dicho por el usuario y con el comando “aceptar” se confirma y envía el mensaje para así finalizar la prueba. A continuación, en la Tabla 6 se muestra el protocolo de pruebas.

Tabla 6. Protocolo de pruebas final.

Tipo de Interacción	Contenido de la Interacción
Comando	Abrir WhastApp
Comando	Enviar mensaje a Daniel
Dictado	¿hola, como estas?
Comando	Eliminar conversación
Comando	Ir al chat de Daniel / Chat de Daniel
Dictado	Escribir, ¿hola, como estas?
Comando	Aceptar / Enter

A continuación, en la tabla 7 se muestra los resultados obtenidos; contiene el promedio por cada uno de los siete usuarios y el porcentaje total de la prueba en la última columna.

Tabla 7. Resultados protocolo de pruebas final.

PERSONA NUMERO	1			2			3			4			5			6			7			TOTAL (%)
	2	1	1	0	0	1	0	1	0	0	0	1	2	0	1	1	1	0	1	0	0	
NUMERO DE FALLAS	2	1	1	0	0	1	0	1	0	0	0	1	2	0	1	1	1	0	1	0	0	8.84
TOTAL, ACIERTOS POR PRUEBA (%)	71	86	86	100	100	86	100	86	100	100	100	86	71	100	86	86	86	100	86	100	100	91.16
TOTAL, ACIERTOS POR PERSONA (%)	80.95			95.24			95.24			95.24			85.71			90.48			95.24			91.16

En esta prueba se obtuvieron resultados muy similares en todos los voluntarios. En la interacción verbal las acciones que obtuvieron un mejor desempeño fueron: “Abrir WhatsApp”, “Eliminar conversación”, “Escribir” y “Aceptar”. Por el contrario, los que obtuvieron menor resultado fueron: “Enviar mensaje a” y “Ir al chat de”.

Esto demuestra que si el usuario tiene dos contactos con el mismo nombre el aplicativo selecciona al primero de la lista, y provoca el envío del mensaje a una persona incorrecta.

Finalmente, en la prueba se obtuvo un resultado general del 91.16% de aciertos frente al 8.84% de error promedio, donde el mayor número de comandos erróneos por usuario fueron dos. La Tabla 7 nos muestra que la primera persona en realizar la prueba tuvo un menor resultado de acierto con un 80.95% debido a que presentó un ritmo acelerado al pronunciar los comandos frente a la última persona que tuvo un error más bajo por que realizó las pruebas con más calma.

6.1.5 EXPERIENCIA DE USUARIO

Para finalizar el protocolo de pruebas se realizó una encuesta para medir la experiencia del usuario con el aplicativo utilizando como métrica de evaluación la escala de Likert (Matas, 2018; J. G. Sánchez & Terrats, 2011), La encuesta que se realizó fue la siguiente:

- 1) La Interfaz reconoce la voz.
- 2) La interfaz reconoce todos los comandos de navegación.
- 3) La interfaz reconoce lo dictado, al momento de enviar un mensaje a un contacto.
- 4) La ejecución de los comandos en la interfaz es rápida.
- 5) Las personas con limitaciones motrices tienen impedimento al momento de conectarse con el mundo vía internet.
- 6) Sabiendo que la interfaz está pensada para personas con tetraplejia, La herramienta ayuda a la comunicación virtual.
- 7) La interfaz es de fácil uso para el usuario.
- 8) La interfaz se puede utilizar en cualquier ambiente social sin interferencia en el reconocimiento de voz.
- 9) La interfaz cumple el objetivo de realizar tareas básicas de navegación en la aplicación WhatsApp desktop.
- 10) Pensando en la necesidad de las personas con tetraplejia, La interfaz es una herramienta que se puede recomendar.

Las personas que realizaron las pruebas anteriores fueron las encargadas de responder esta encuesta. Al finalizar se calculó la frecuencia de las respuestas y este fue el resultado:

Tabla 8. Resultados de la experiencia del usuario.

	Opción respuesta <i>likert</i>	Preguntas de la encuesta experiencia de usuario										Total	Equivalente (%)	
		1	2	3	4	5	6	7	8	9	10			
1	Totalmente en desacuerdo	0	0	0	0	0	0	0	0	0	0	0	0	0%
2	En desacuerdo	0	2	0	1	0	0	0	3	0	0	6	8.6%	
3	Ni de acuerdo ni en desacuerdo	1	3	2	2	3	1	2	2	2	1	19	27.1%	
4	De acuerdo	3	2	4	3	3	2	5	0	2	5	29	41.4%	
5	Totalmente de acuerdo	3	0	1	1	1	4	0	2	3	1	16	22.9%	

Partiendo de la gráfica anterior, se pudo identificar que un 85.71% de las personas encuestadas están de acuerdo de que la interfaz reconoce su voz y al ser la primera pregunta de la encuesta nos benefició mucho saber que la mayoría de los usuarios evidenciaran un buen trabajo de reconocimiento de voz por parte del aplicativo.

De igual forma la segunda pregunta es muy importante, sin embargo, hubo un cambio respecto a la primera, pues a pesar de que los usuarios reconocen que el aplicativo reconoce la voz, un 71% menciona que ellos pueden ver molesto la posibilidad de repetir algún comando.

Por otro lado, un total del 85.71% de las personas, están de acuerdo y totalmente de acuerdo que el aplicativo puede ayudar a las personas con tetraplejia y reducir la limitación de comunicación con otras personas. Con base en la pregunta 8, el 42.85% las personas afirman que no es posible usar la aplicación en cualquier ambiente, recomiendan un espacio silencioso que les permita tener un uso estable de la aplicación.

La encuesta reúne además observaciones constructivas por parte de los sujetos de prueba, un ejemplo de estas puede ser un soporte para *emojis*, que les permita tener conversaciones más naturales y les permita expresar ciertos estados de ánimo en sus mensajes.

Otro ejemplo fue que los mensajes no se enviaran de forma automática, pues si el aplicativo filtraba otro contacto por error no hay opción de reversar dicha acción y puede afectar la experiencia de usuario, por lo demás estaban de acuerdo que para personas que no tengan la posibilidad de usar un periférico tradicional les sería muy útil la aplicación.

7. ALCANCES Y LIMITACIONES

7.1 ALCANCES

El presente proyecto cumple con el alcance de desarrollar un software que permite a los usuarios usar WhatsApp desktop mediante comandos de voz.

Los aspectos puntuales que caracterizan su desarrollo fue el uso de tecnologías de reconocimiento de voz independientes del locutor (cualquier persona lo puede utilizar), la integración de un sintetizador de voz, la implementación en el sistema operativo Windows mediante algoritmos en Python.

7.2 LIMITACIONES

Las presentes limitaciones que restringirán el desarrollo de la interfaz:

- Dependencia del espacio en el que se use la herramienta (Ruido Ambiente).
- La identificación de un tiempo muerto de procesamiento al realizar el reconocimiento de voz (Propio del algoritmo).
- Es necesario contar con una conexión activa a Internet para que el sistema funcione.
- La dependencia de una buena pronunciación de los comandos y el control del habla; en el proceso se identificó que en ocasiones el aplicativo confundía las palabras porque las personas hablaban muy rápido.

8. CONCLUSIONES

Se desarrollo una herramienta que, a partir de comandos de voz, permite una emulación del teclado sin tener un contacto físico con el computador, dicha interfaz podrá ser usada por personas con discapacidades motoras o que deseen controlar de forma verbal la aplicación WhatsApp Desktop.

El sistema de reconocimiento de voz permitió realizar las tareas de navegación más comunes la aplicación de WhatsApp: enviar un mensaje, ir a una conversación en específico o crear una nueva conversación. Se realizaron pruebas con un grupo de personas de 7 personas con diferentes edades y hábitos sobre el uso de internet, en las pruebas se utilizaron 672 comandos de voz y 42 dictados, obteniendo un mayor porcentaje de acierto en los dictados con un 93% y un menor porcentaje en los comandos con un 91.76% de aciertos.

En los comandos de voz se observó que las palabras con estructura fonética similar presentan fallos pues existía la posibilidad de confundir dos comandos entre sí, un ejemplo es “activar” y “desactivar” que al ser muy similares si el usuario habla antes de que el micrófono se active puede llegar a confundirlos los comandos, también se observó como los que terminaban con “ar” tuvieron más fallos con un 7.70% de las pruebas realizadas.

En los dictados se observó que no se reconocieron los signos de puntuación, pero si se reconoció el texto pronunciado, a su vez se pudo observar cómo existe una relación directamente proporcional en el reconocimiento de palabras en inglés y la pronunciación del usuario, dando evidencia de una ayuda poca asistencia a un lenguaje diferente al parametrizado.

El sistema de interacción por comandos de voz fue evaluado por personas que no tenían limitaciones motrices, en base a su experiencia con el aplicativo se concluyó que la herramienta puede mejorar la calidad de vida de personas con limitaciones motrices, pues le permite controlar la aplicación de mensajería sin utilizar el teclado o mouse, adicionalmente tener más privacidad al momento e interactuar con otras personas sin asistencia de terceros.

9. RECOMENDACIONES

Una vez concluido el desarrollo de la interfaz, se considera tener en cuenta los siguientes aspectos para mejorar su desempeño:

- Implementar nuevas librerías de reconocimientos de voz, citando entre estas la Speech-to-Text de Google Cloud, ya que al obtener una base de datos más amplia permite reconocer más palabras e identificar diferentes idiomas mejorando la experiencia del usuario, incluso en ambientes ruidosos.
- Realizar la migración de los algoritmos desarrollados en este trabajo de investigación, incitando el uso de librerías con licencia de libre distribución para poder compartir y comercializar este desarrollo.

BIBLIOGRAFÍA

Alexandra, V., & Bejarano, L. (2020). *Universidad Técnica del Norte Facultad Ciencias de la Salud carrera de terapia física médica*. <http://repositorio.utn.edu.ec/handle/123456789/9899>

Álvarez Puluchi, Christian Gerardo, Cajas Fuela, Miguel Armando (junio de 2016), *Desarrollo de una interfaz p*. Retrieved May 23, 2020, from <https://www.google.com/search?client=firefox-b-d&q=%5B4%5D%09Álvarez+Puluchi%2C+Christian+Gerardo%2C+Cajas+Fuela%2C+Miguel+Armando+%28junio+de+2016%29%2C+Desarrollo+de+una+interfaz+p>.

Avalos, M., & Borba, D. (2020). *El uso responsable y seguro de internet: Aportes para la conformación de la ciudadanía digital*. <https://books.google.es/books?hl=es&lr=&id=0w7bDwAAQBAJ&oi=fnd&pg=PT2&dq=computadoras+en+la+sociedad+actual+2020&ots=m9zvUxSUi0&sig=AHZRr3xZEaH5leJM9TNLn4902I>

Barrobés, H., & Ruiz, M. (2017). *Síntesis del habla* Helenca Duxans Barrobés Marta Ruiz Costa-jussà PID_00188071. [https://www.exabyteinformatica.com/uoc/Audio/Procesamiento_de_audio/Procesamiento_de_audio_\(Modulo_8\).pdf](https://www.exabyteinformatica.com/uoc/Audio/Procesamiento_de_audio/Procesamiento_de_audio_(Modulo_8).pdf)

Bhat, N. M. (2020). *pyttsx3 · PyPI*. Pypi.Org. <https://pypi.org/project/pyttsx3/>

Caicedo E., López J. (2009) *Una aproximación práctica a las redes neuronales artificiales*. Universidad del Valle (Cali, Colombia), ISBN 978-958-670-676-1

Castro-molinares, S. P., & Ph, D. (2019). *Universidad Cooperativa de Colombia , sede Santa Marta*. <https://repository.ucc.edu.co/handle/20.500.12494/14492>

Ciencias De La Computación, I. Y. (2013). *Universidad Tecnológica Equinoccial Facultad de Ciencias de la Ingeniería Carrera de Informática y Ciencias de la Computación Prototipo que permita el abrir y cerrar una puerta por medio de la voz utilizando redes neuronales trabajo previo a la obtención*.

Cloud Speech-to-Text - Reconocimiento de voz | Cloud Speech-to-Text. (2020)
Retrieved May 22, 2020, from <https://cloud.google.com/speech-to-text?hl=es>

David, J., & Pérez, A. (2011). *Jnavo: Prototipo de navegador web controlado por voz.* Pereira : Universidad Tecnológica de Pereira.

De, C., & De Empresas, A. (2019). *Unidad Académica de Ciencias Empresariales.*
<http://repositorio.utmachala.edu.ec/handle/48000/13900>

De, Manuel, L., & Sisniegas, A. (2012). *Universidad Nacional de Trujillo, Facultad de Ciencias Físicas y Matemáticas - Diseño de la arquitectura de un sistema empotrado para el control de un brazo robótico mediante comandos de voz.*

Del, M., & Kano, R. L. (2013). Tecnología Y Discapacidad: Una Mirada Pedagógica. *Revista Digital Universitaria UNAM*, 14(12), 1607–6079.
<http://www.revista.unam.mx/vol.14/num12/art53/>

Enrique, I. W., & Vivanco, J. (2017). *Tesis Maestría Wilmer Jimenez.* Universidad Católica de Santiago de Guayaquil.

GARCÍA P, D., Castillo M, J., & Castillo C, J. (2007). Complicaciones respiratorias de la tetraplejia: Una mirada a las alternativas terapéuticas actuales. *Revista Chilena de Enfermedades Respiratorias*, 23(2), 106–116.
<https://doi.org/10.4067/s0717-73482007000200005>

Gonzalez Rodriguez, J., Sanchez-Figueroa, F., Lozano, A., & Macias, M. (2007) *NaviSAW y MailSAW herramientas multimodales para el acceso a Internet para usuarios con discapacidad visual 1.* Retrieved May 23, 2020, from <https://www.google.com/search?client=firefox-b-d&q=NaviSAW+y+MailSAW+herramientas+multimodales+para+el+acceso+a+Internet+para+usuarios+con+discapacidad+visual+1>

Google Play. (2018). *Sonómetro (Sound Meter) - Apps en Google Play.*
https://play.google.com/store/apps/details?id=com.gamebasic.decibel&hl=es_CO

Harrison. (2018). *Python Programming Tutorials.* 2018-08.

<https://pythonprogramming.net/python-3-os-module/>

Keyboard Control Functions — PyAutoGUI documentation. (2020). Retrieved May 22, 2020, from <https://pyautogui.readthedocs.io/en/latest/keyboard.html>

Matas, A. (2018). Diseño del formato de escalas tipo Likert: Un estado de la cuestión. *Revista Electronica de Investigacion Educativa*, 20(1), 38–47. <https://doi.org/10.24320/redie.2018.20.1.1347>

Moreno García, S. (2015). Trauma Raquimedular Vaisman. In *revistas.unal.edu.co* (Vol. 7, Issue 1). <https://revistas.unal.edu.co/index.php/morfologia/article/view/50590>

Mosquera H., Oliveros E. Loaiza H. (2011) Interfaz Humano-Máquina Audiovisual, Grupo en Percepción y Sistemas Inteligentes, Trabajo de Grado en Ingeniería Electrónica, Universidad del Valle

Nuance Communications. (2020). *Dragon Speech Recognition - Get More Done by Voice | Nuance.* Nuance Communications, Inc. <https://www.nuance.com/dragon.html>

Nope S., Loaiza H., Caicedo E. (2008) Estudio Comparativo de Técnicas Para el Reconocimiento de Gestos Por Visión Artificial. 2008. Grupo en Percepción y Sistemas Inteligentes, Universidad del Valle 2008. Publicado en la Revista Avances en Sistemas e Informática, Vol. 5 - No.3, Diciembre de 2008, Medellín, ISSN 1657-7663

Oswaldo, Y., Uicab, Q., & Menéndez Domínguez, H. (2018). Extensión de las fronteras del uso de equipos de cómputo: De pacientes con ELA a personas con discapacidad motriz severa. In *aihc.amexihc.org* (Vol. 3, Issue 1). <http://aihc.amexihc.org/index.php/aihc/article/view/38>

Pardo, A., Rubiano, Z., & Montagut, Y. (2003). Automatización de una silla de ruedas controlada por comandos de voz. *Revista Colombiana de Tecnologías de Avanzada*, 2, 6. https://www.google.com/search?client=firefox-b-d&sxsrf=ALeKk03UIT7lkfGblsZCUB1UYjGlr9EqYg%3A1590282251804&ei=C8jXtvNMMnH_QawmouACw&q=AUTOMATIZACIÓN+DE+UNA+SILLA+DE+RUEDAS+CONTROLADA+POR+COMANDOS+DE+VO%5D+Universidad+de

+Pamplona&oq=AUTOMATIZACIÓN+DE+UNA+S

Pérez, J., Díaz, J., & Harari, I. (2017). *Ejecución de comandos de voz mediante Web Speech API*. *Sedici.Unlp.Edu.Ar*, 964–970. <http://sedici.unlp.edu.ar/handle/10915/63851>

Pérez, M. M., & Gómez, J. G. (2017). *Estudio del efecto del entrenamiento sobre la marcha mediante ayuda robótica en combinación con realidad virtual en niños con parálisis cerebral infantil/tesis*. <https://repositorioinstitucional.ceu.es/handle/10637/8566>

Pham, H. (2017). *PyAudio* · *PyPI*. *PyPI.Org*. <https://pypi.org/project/PyAudio/>

Precios - Bing Speech API | Microsoft Azure,2020. (2020). Retrieved May 23, 2020, from <https://azure.microsoft.com/es-es/pricing/details/cognitive-services/speech-api/>

Prieto, F. (2013). *Control de luces por voz, una aplicación orientada a personas con capacidades especiales 1*. 1–2. <https://www.google.com/search?client=firefox-b-d&q=Control+de+luces+por+voz%2C+una+aplicaci%27on+orientada+a+personas>

PyAutoGUI · *PyPI*. (2020). Retrieved May 22, 2020, from <https://pypi.org/project/PyAutoGUI/>

Python | os.system() method - GeeksforGeeks. (2020). Retrieved May 22, 2020, from <https://www.geeksforgeeks.org/python-os-system-method/>

Python Speech Recognition. (2019). *SpeechRecognition* · *PyPI*. *PyPI.Org*. <https://pypi.org/project/SpeechRecognition/>

Rashid, A. O., & Islam, S. (2017). Percutaneous tracheostomy: A comprehensive review. In *Journal of Thoracic Disease* (Vol. 9, pp. S1128–S1138). AME Publishing Company. <https://doi.org/10.21037/jtd.2017.09.33>

Rabiner, L., Biing-Hwang J., (1993) *Fundamentals of Speech Recognition*, ISBN

987-0130151575, Editorial Prentice Hall; 1 edition April 22 1993.

Sánchez, A. A., Conrado, M. V. B., & 2020, U. (2020) *Desarrollo de competencias investigativas y uso de herramientas tecnológicas en la gestión de información*. Scielo.Sld.Cu. Retrieved May 21, 2020, from http://scielo.sld.cu/scielo.php?pid=S1990-86442020000100096&script=sci_arttext&tlng=en

Sánchez, J. G., & Terrats, J. A. (2011). *Guía técnica para la construcción de escalas de actitud*. <https://www.odiseo.com.mx/2011/8-16/garcia-aguilera-castillo-guia-construccion-escalas-actitud.html>

Software de reconocimiento del habla y traducción instantánea. (2020). Retrieved May 23, 2020, from <https://speechlogger.appspot.com/es/>

SpeechTexter | Type with your voice! (2020). Retrieved May 23, 2020, from <https://www.speechtexter.com/>

Zhang, J., Muñoz, C., Como, L. H.-L. redes sociales, & 2020. (2020) *Oportunidades y retos para la comunicación integrada y comercial en redes sociales de mensajería: los casos de WhatsApp y WeChat*. Academia.Edu. Retrieved May 21, 2020, from <http://www.academia.edu/download/62125595/Libro-01-v11-HIGH-PRINT-DEF20200217-34468-ei379i.pdf#page=439>

ANEXOS

Anexo A. REPOSITORIO CODIGO FUENTE

Los algoritmos desarrollados en este trabajo de grado se encuentran alojados en el repositorio: <https://github.com/llkevinll/interfaz-de-voz/tree/master>

Anexo B. CONSENTIMIENTO INFORMADO PARA PARTICIPACIÓN EN INVESTIGACIÓN

INTERFAZ HUMANO-COMPUTADOR PARA PERSONAS CON LIMITACIÓN MOTRIZ DE MIEMBROS SUPERIORES BASADA EN GESTOS FACIALES

I. INFORMACIÓN

Usted ha sido invitado a participar en la **investigación Interfaz Humano-Computador para Personas con Limitación Motriz de Miembros Superiores Basada en Gestos Faciales**. Para decidir participar en esta investigación, es importante que considere la siguiente información. Siéntase libre de preguntar cualquier asunto que no le quede claro. Si no tiene preguntas ahora, usted podrá hacerlas en cualquier momento, a través de los datos de contacto abajo relacionados. Su participación en este estudio es completamente voluntaria.

El objetivo de este estudio es Desarrollar una interfaz humano computador basada en gestos faciales y detección de zonas de interés en una aplicación orientada al internet para personas con limitaciones motrices de miembros superiores.

Participación: Si usted autoriza participar en este estudio se le aplicarán las siguientes técnicas de recolección de la información:

1. Se harán pruebas para implementar una interfaz humano computador basada en gestos faciales y detección de zonas de interés en una aplicación orientada al internet para personas con limitaciones motrices de miembros superiores. Para esto el sujeto estará ubicado frente a una cámara de video y monitor de computador, para así poner a prueba el programa desarrollado.

Durante el desarrollo de la técnica o procedimiento se tomará anotaciones por parte del investigador o se hará uso de una grabadora de voz o cámara de video con el fin de revisar en detalle la información obtenida, sin embargo, esta grabación podrá ser interrumpida y/o retomada en el momento que usted lo considere, sin que esto conlleve algún perjuicio para usted o la institución que representa.

Riesgos: El estudio que se desarrollará ha sido clasificado como una investigación. Dado que el proyecto culminará con el desarrollo de una interfaz humana

computador la cual se ensayará con personas reales, el tipo de estudio propuesto es de tipo experimental. Básicamente este proyecto se apoyará en trabajos sistemáticos fundamentados en los conocimientos alrededor de los sistemas de reconocimiento de gestos de la cara que permitirá la creación de una interfaz humano computador basada en gestos faciales y detección de zonas de interés en una aplicación orientada al internet para personas con limitaciones motrices de miembros superiores Según el Artículo 11 de la Resolución 008430 de 1993, esta investigación no representa riesgo alguno para los participantes; sin embargo, siempre existe el riesgo de violación de la confidencialidad y por ello hemos adoptado todos los mecanismos necesarios para minimizar este riesgo, con la firma de los respectivos acuerdos de confidencialidad.

Beneficios: Usted no recibirá ningún beneficio directo, ni recompensa alguna, por participar en este estudio. No obstante, su participación permitirá generar información para determinar las principales características de una interfaz humano-computador para personas con limitaciones motrices de miembros superiores, implementar una técnica de visión por computador para la identificación de gestos faciales y la detección de zonas de interés en una aplicación de escritorio orientada a internet, desarrollar una interfaz software de generación de comandos para una aplicación de escritorio a partir de gestos y zonas de interés detectadas y ejecutar un plan de pruebas para definir los alcances y limitaciones del sistema.

Voluntariedad: Su participación es absolutamente voluntaria. Usted tendrá la libertad de preguntar lo que considere pertinente para proseguir, como también de detener su participación en cualquier momento que lo desee. Esto no implicará ningún perjuicio para usted.

Confidencialidad: La información recolectada no será usada para ningún otro propósito, además de los señalados anteriormente, sin su autorización previa y por escrito. Todas sus opiniones y aportes serán confidenciales, y mantenidas en estricta reserva. En las presentaciones y publicaciones de esta investigación, su nombre no aparecerá asociado a ninguna opinión particular. Hasta donde nos es posible, sus respuestas serán confidenciales y minimizaremos los riesgos asignando un código a su formulario y limitando el acceso a sus datos personales únicamente al investigador principal.

Conocimiento de los resultados: Usted tiene derecho a conocer los resultados de esta investigación. Para ello, una vez sean sistematizados y publicados los resultados de la investigación, se le hará llegar una copia del enlace que le permitirá acceder a la misma a la dirección de correo electrónico por usted aportada.

Datos de contacto: Si requiere más información o comunicarse por cualquier motivo relacionado con esta investigación, puede contactar al investigador principal responsable de este estudio:

Nombre: _____ Teléfonos: _____ E-mail: _____
También puede comunicarse con el Comité de Ética de la Investigación que aprobó este estudio:

Presidente (a) del Comité de Ética de la Fundación Universitaria Católica Lumen Gentium –Unicatólica-: Fabio Alberto Enríquez Martínez

Carrera 122 No. 12 - 409, Cali, Valle del Cauca, PBX (2) 5552767 Ext. 1110

Correo Electrónico: ceiunicatolica@unicatolica.edu.co

II. FORMULARIO DE CONSENTIMIENTO INFORMADO

Yo, _____,
acepto participar en el estudio INTERFAZ HUMANO-COMPUTADOR PARA PERSONAS CON LIMITACIÓN MOTRIZ DE MIEMBROS SUPERIORES BASADA EN GESTOS FACIALES, en los términos aquí señalados.

Declaro que he leído y comprendido, las condiciones de mi participación en este estudio. He tenido la oportunidad de hacer preguntas y estas han sido respondidas en su totalidad por el investigador responsable. También sé que se respetará la buena fe, la confidencialidad e intimidad de la información por mí suministrada, lo mismo que mi seguridad física y psicológica. No tengo dudas al respecto.

Firma Participante

Firma Investigador Responsable

c.c. _____

c.c.

Correo electrónico: _____

Lugar

y

Fecha:
