



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS

“DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA TRADUCTOR DE TEXTO Y VOZ A LENGUAJE DE SEÑAS A TRAVÉS DE UNA INTERFAZ GRÁFICA CON UN AVATAR MEDIANTE DISPOSITIVO ANDROID”.

AUTORA:

MYRIAM PATRICIA CARGUACUNDO NUELA

DIRECTORA:

ING. PATRICIA CONSTANTE

**LATACUNGA
2017**



Introducción

- Capítulo I
Estado del arte
- Capítulo II
Diseño y selección de elementos
- Capítulo III
Implementación y pruebas
- Capítulo IV
Conclusiones
Recomendaciones

Introducción

- En la actualidad la tecnología ha ido evolucionando.
- Existen proyectos de guantes
- El 31.6% de las personas tienen un celular inteligente por lo cual se desarrolló este sistema traductor texto o voz a señales de señas.

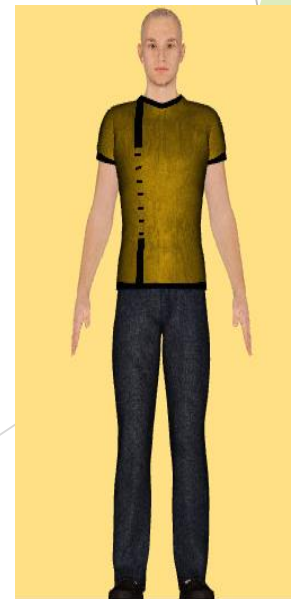


Objetivos

Objetivos Generales:

Diseñar e implementar un sistema traductor del lenguaje texto y voz a lenguaje de señas a través de una interfaz gráfica con un avatar mediante dispositivo Android.

Objetivos Específicos:



Justificación

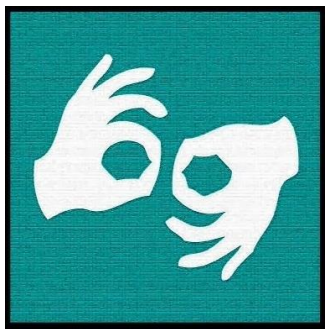
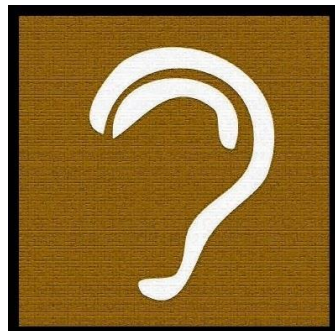
Este proyecto tendrá impacto social ya que su primer objetivo es ayudar a las personas con discapacidades auditivas y de lenguaje.

El proyecto es factible debido a que se realizará un sistema con una interfaz gráfica amigable que facilitará el mejoramiento del aprendizaje de las personas y también permitirá el aprendizaje para las personas que no saben el lenguaje de señas y se podrá realizarlo mediante el ingreso de señales de voz o texto.

Capítulo I

Estado del arte

Discapacidad

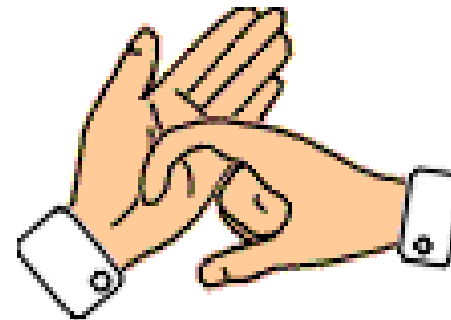
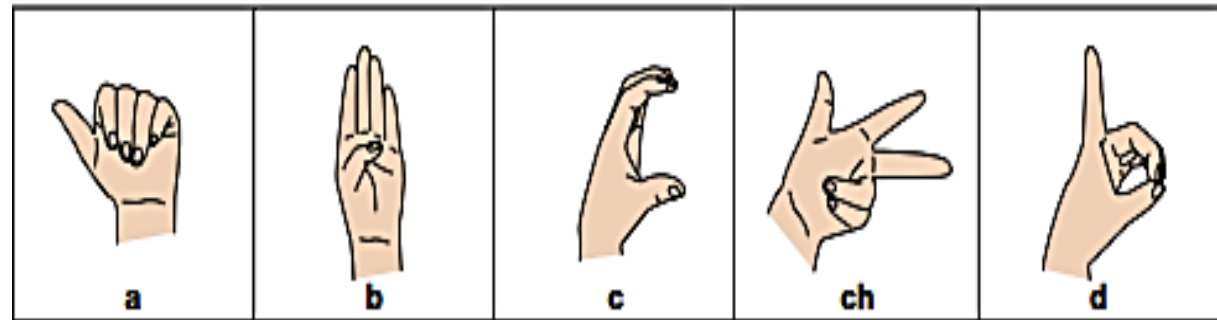


Según el Consejo Nacional de Discapacidades hasta Agosto 2017, el total de personas con alguna discapacidad es de 425,877 las cuales son:

TIPOS DE DISCAPACIDAD	TOTAL DE PERSONAS
Auditiva	54,795
Física	199,284
Intelectual	96,135
Lenguaje	5,621
Psicosocial	19,614
Visual	50,428

Lenguaje de señas

El lenguaje de señas se basa en las expresiones faciales, movimientos corporales y gestuales para comunicarse.



Componentes del Sistema de Codificación de Voz a Texto

La voz humana

- La voz se produce dentro de los límites del número de vibraciones que mejor soporta el oído humano.

Fonología

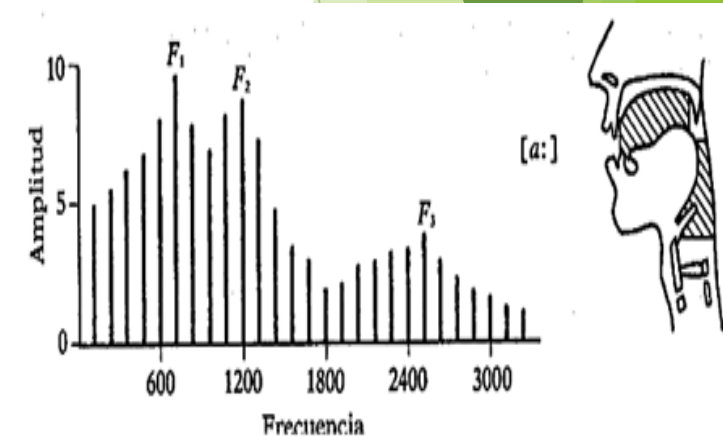
- Estudio de los sonidos de la lengua en cuanto a su carácter simbólico o de representación mental.

Fonética

- Se refiere a los sonidos en el habla, incluyendo su producción acústica y los procesos físicos y fisiológicos de emisión y articulación involucrados

Formantes

- Los articuladores alteran las resonancias vocales para formar los diferentes sonidos vocales.



Avatar

El avatar es una animación tridimensional de una fisonomía de un humano que realiza cada una de las señales según corresponda a las palabras ingresadas.



Brazo humano

Brazo

- El brazo se compone del hueso húmero, en el que se insertan distintos músculos para desarrollar su fuerza

Antebrazo

- Los huesos del antebrazo comienzan con la articulación del codo y desembocan en los huesos de la mano

Mano

- El esqueleto de la mano se subdivide en cuatro partes: el carpo, el metacarpo, la palma de la mano, las falanges

Dedos

- La falange proximal
- La falange medial
- La falange distal



Aparato de tamaño pequeño.

Capacidades de procesamiento

Conexión permanente o intermitente a una red

Memoria limitada



Capítulo II

Diseño y selección de elementos

Diseño

La navegación entre pantallas debe ser rápido, fácil acceso.

En la interfaz gráfica se consideran dos modos de operación de ingreso de palabras mediante teclado y voz para facilitar el manejo del sistema.

Avatar proporcional al cuerpo humano.

La aplicación del sistema debe ser compatible con versiones.

Diseño del sistema del avatar

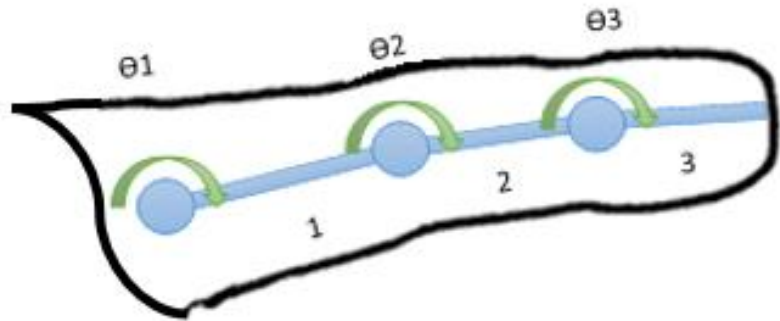
Fisonomía de un humano.

Análisis cinemático

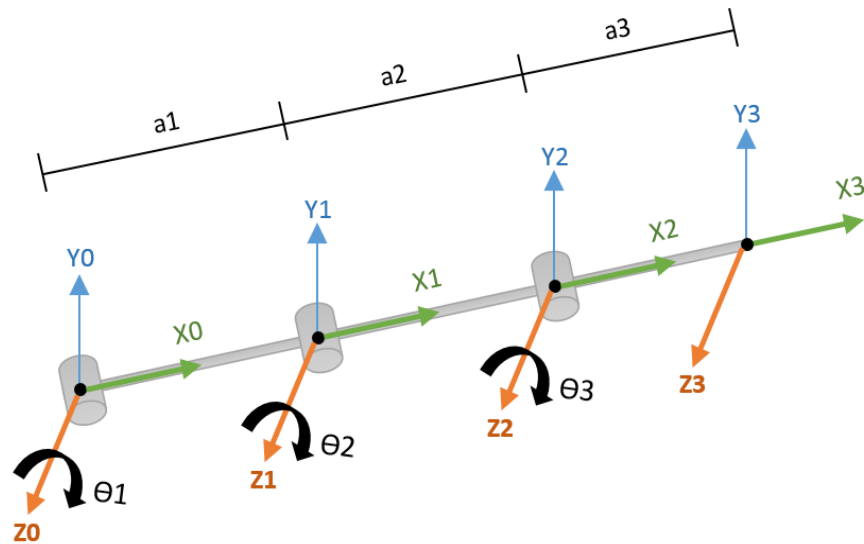
El análisis cinemático directo



Análisis cinemático del dedo índice

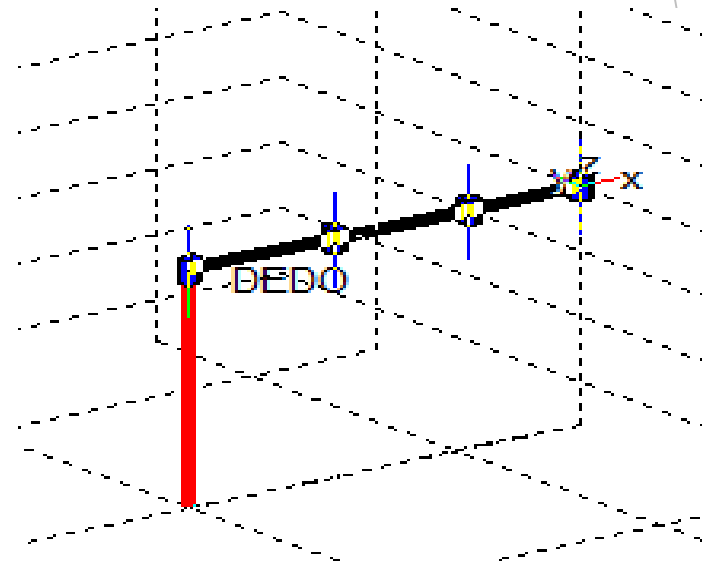


Grados de libertad y eslabones

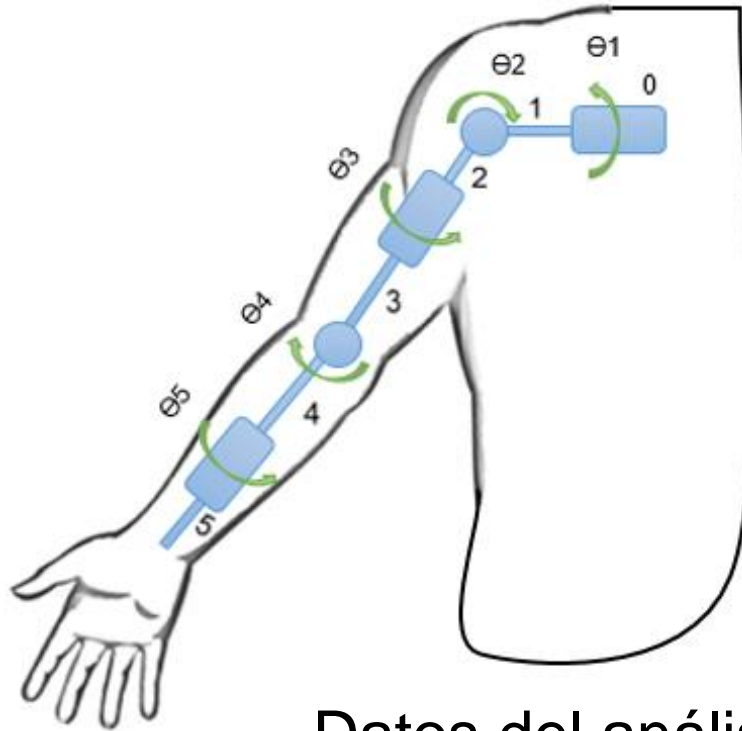


Articulación	Θ	d	a(mm)	α
1	q_1	0	$a_1=27$	0
2	q_2	0	$a_2=25$	0
3	q_3	0	$a_3=21$	0

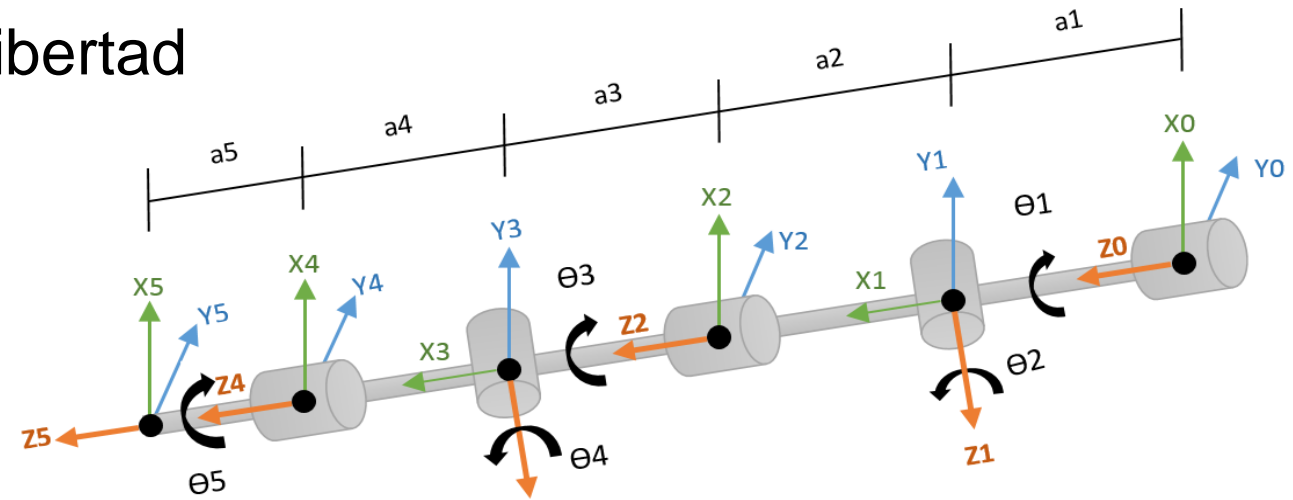
Datos del análisis D – H



Análisis cinemático del brazo

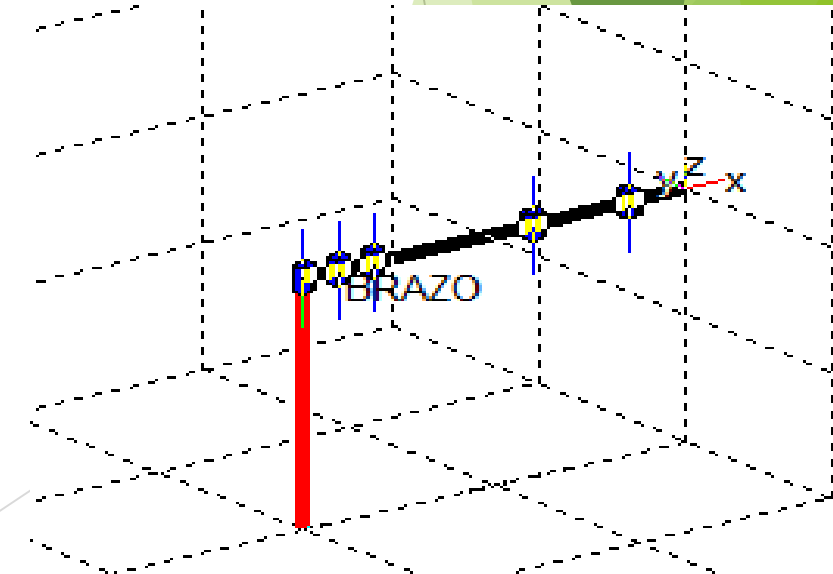


Grados de libertad y eslabones

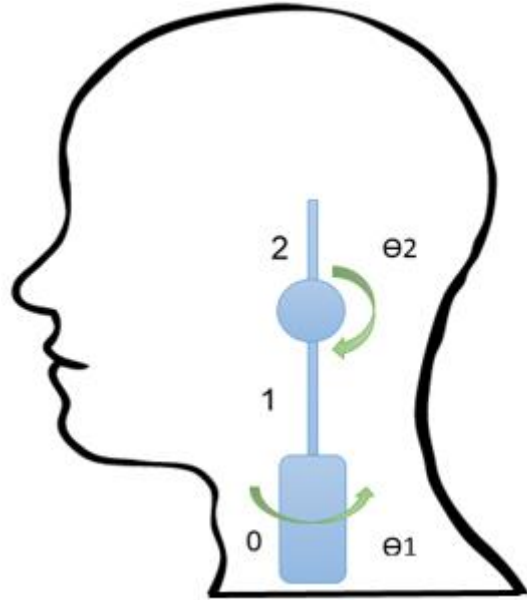


Datos del análisis D – H

Articulación	θ	d	$a(\text{cm})$	α
1	q_1	0	$a_1=9$	0
2	q_2	0	$a_2=9$	0
3	q_3	0	$a_3=42$	0
4	q_4	0	$a_4=25$	0
5	q_5	0	$a_5=15$	0

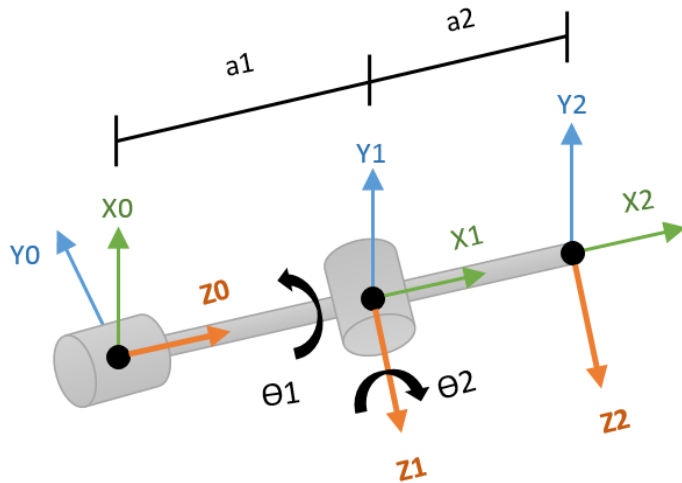


Análisis cinemático de la cabeza y el cuello



Articulación	Θ	d	a(cm)	α
1	q_1	0	$a_1=2$	0
2	q_2	0	$a_2=4.8$	0

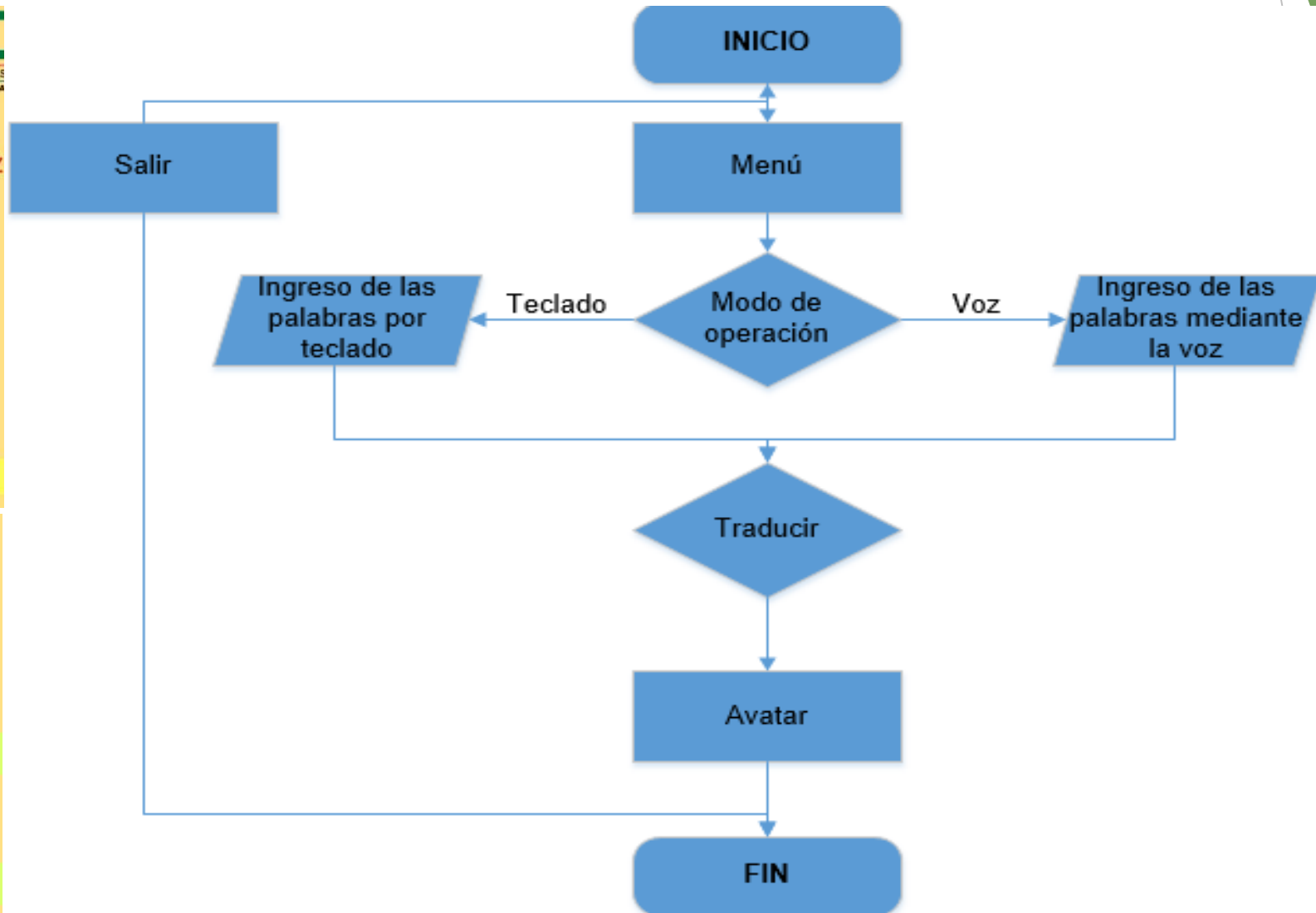
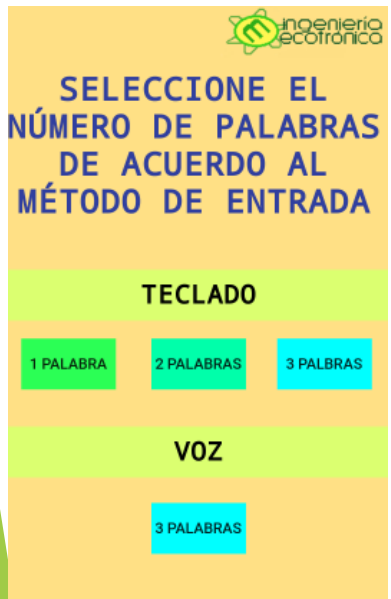
Grados de libertad y eslabones



Datos del análisis D – H



Diseño de la interfaz gráfica



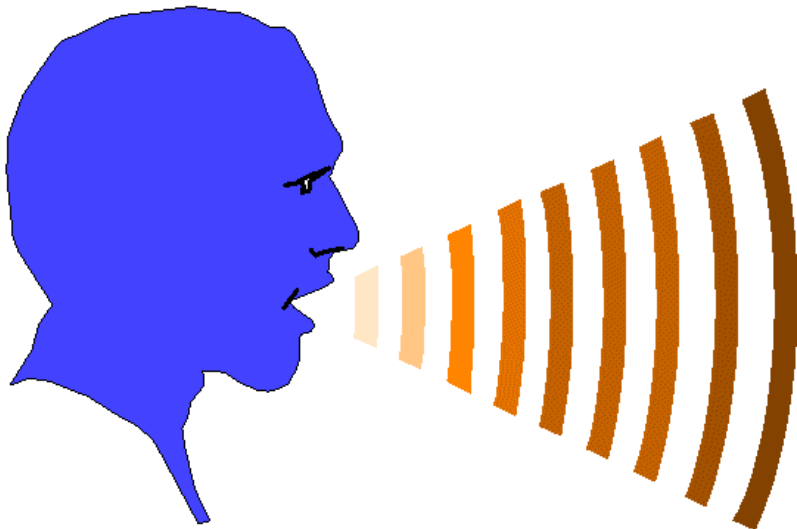
Capítulo III

Implementación y pruebas



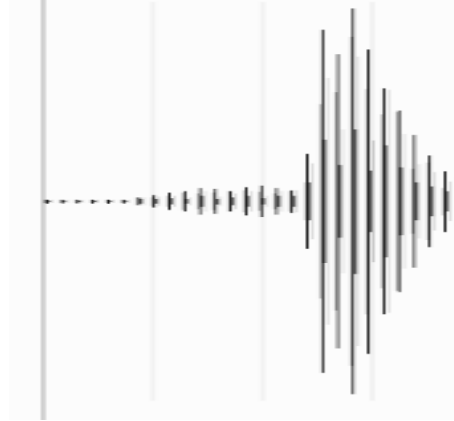
Procesamiento de la señal de voz

- Preprocesamiento: Convierte la entrada de voz a una forma que el reconocedor pueda procesar.
- Reconocimiento: Identifica lo que se dijo mediante redes neuronales.
- Comunicación: Envía la palabra al sistema que lo requiere.



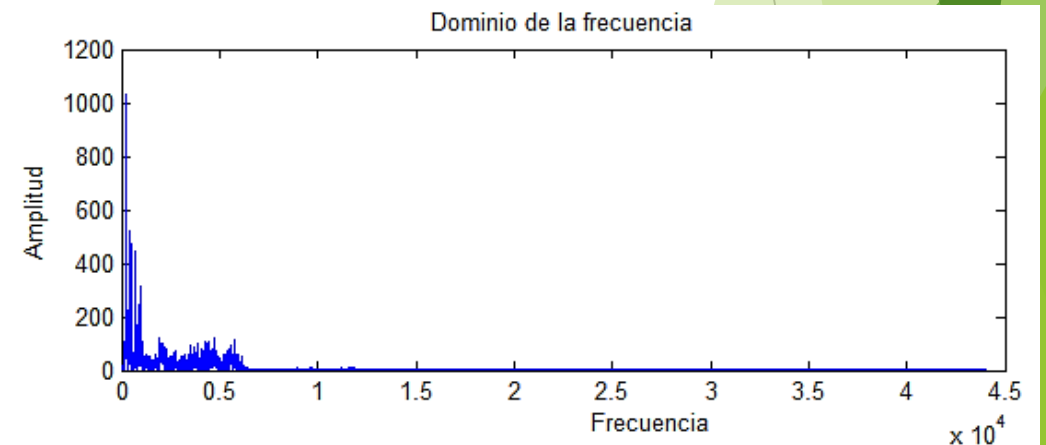
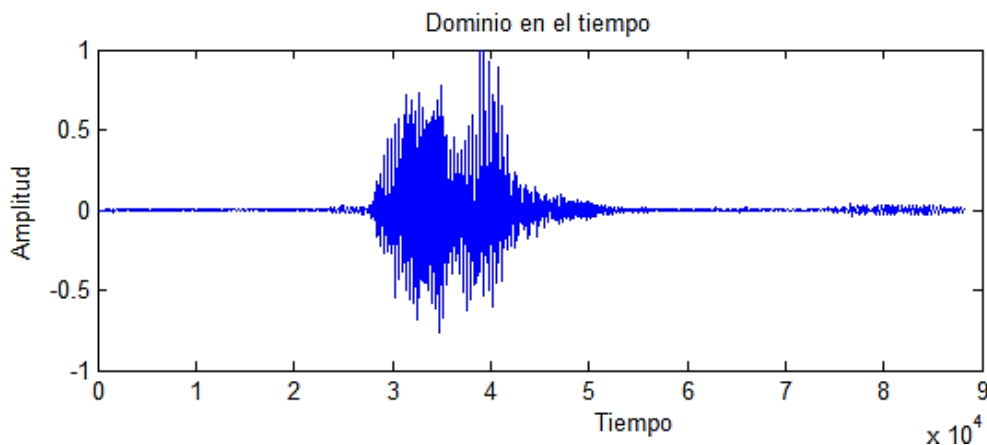
Entrada de la voz

Se activa el motor de reconocimiento de voz para ingresar la palabra.
La palabra de la señal ingresa de forma análoga al sistema como se indica en la figura



Pre-procesamiento

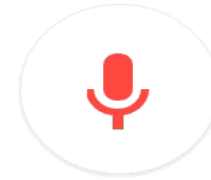
Pasa al dominio de la frecuencia.



Reconocimiento

En esta fase recibe la señal del preprocesamiento y se envía a una red neuronal recurrente que utiliza GOOGLE NOW.

Habla ahora



Comunicación

En esta última fase después de haber realizado el reconocimiento envía la palabra en forma de texto.

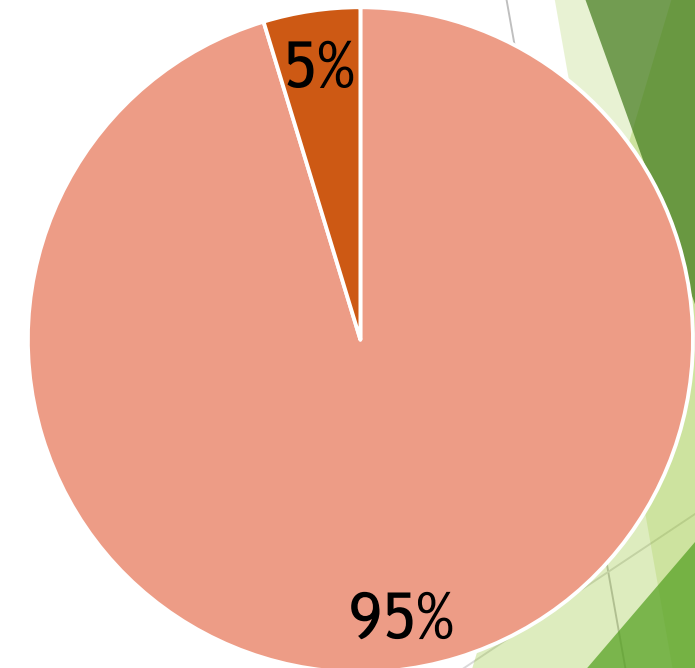
Implementación de la aplicación Android



Pruebas de reconocimiento de voz

Número de personas	Número de palabras	Reconoció	No reconoció
10	17	162	8

La muestra de palabras es 17



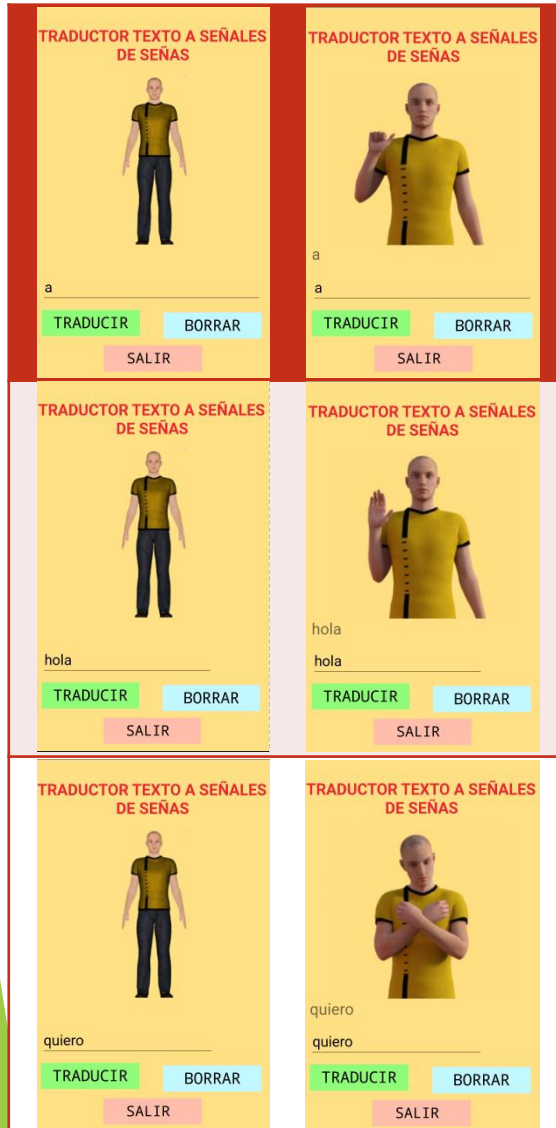
■ Reconoció ■ No reconoció

$$\% = \frac{|valor\ actual - valor\ antiguo|}{|valor\ antiguo|} \times 100\%$$

$$e\% = \frac{|162 - 170|}{|170|} \times 100\%$$

$$e\% = 4,70\% \approx 5\%$$

Pruebas base de datos

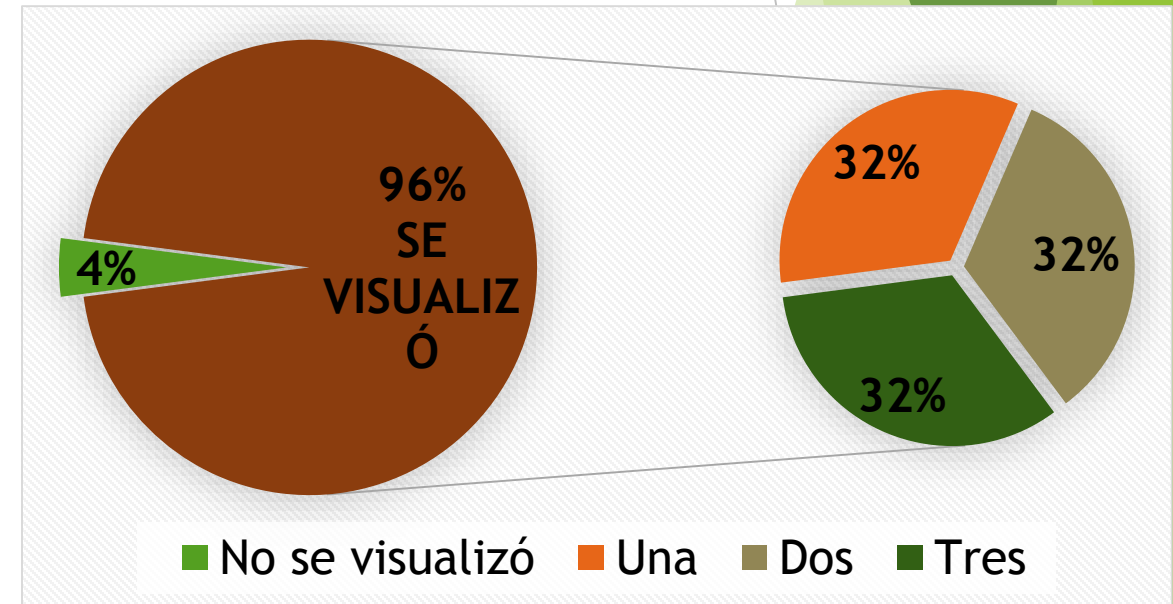


Número de personas	Número de palabras	Visualizó	No se visualizó
10	17	163	7

$$e\% = \frac{|valor\ actual - valor\ antiguo|}{|valor\ antiguo|} \times 100\%$$

$$e\% = \frac{|163 - 170|}{|170|} \times 100\%$$

$$e\% = 4,11\% \approx 4\%$$



Validación de la hipótesis



¿El diseño e implementación de un sistema traductor de texto, voz y su correspondiente interpretación a lenguaje de señas con un avatar mediante dispositivos Android facilitará la comunicación de las personas sordomudos?

El nivel de confianza estimado para el proyecto es del 0.05% y el grado de libertad 16 y se tiene un valor de 26.2962.

El valor obtenido es 38.2 es mayor a 26.2962, por lo tanto se acepta la hipótesis de trabajo y se rechaza la hipótesis nula.



Costo del proyecto

Una vez finalizado cada una de las etapas de diseño, implementación y pruebas de funcionamiento se debe indicar el valor económico.

El costo del proyecto es de 500 dólares

Descripción	Valor total
Celular Samsung S5	600.00
Software libre Android Studio	-
Software libre Daz Studio	-
Horas de investigación	300.00
Horas de programación	200.00



Capítulo IV

Conclusiones y Recomendaciones

- El presente sistema traductor de texto y voz a lenguaje de señas a través de una interfaz gráfica con un avatar mediante dispositivo Android se culminó exitosamente aprobándose la hipótesis de trabajo.
- La orientación del proyecto es para personas sordomudas y sus familiares, para mejorar su comunicación y aprendizaje, las mismas que sean personas independientes y se desenvuelvan en su entorno.
- Se realizó el avatar con cada una de las señas que se encuentran en la base de datos las mismas que estén perfectamente definidas para que pueda ayudar en su comunicación y enseñanza a las personas sordomudas y sus familiares

- La programación del sistema se desarrolló en Android Studio ya que es una plataforma de muy fácil uso y el diseño de la interfaz gráfica se realizó con un entorno amigable, que sea de fácil uso con el usuario y con una combinación de colores.
- Las señas del avatar se realizó en DAZ Studio que es un software libre y de fácil manipulación que permite realizar cada una de las señas mediante puntos y que en un futuro se puede implementar en un robot porque ya se tiene los ángulos de cada señal.
- Se realiza el análisis cinemático de cada una de las partes superiores del avatar y se verificó el movimiento de cada una de sus articulaciones mediante el Software de Matlab.

Recomendaciones

- El lenguaje dactilológico se debe realizar correctamente en el avatar para que pueda el usuario utilice sin tener ningún inconveniente.
- El número de palabras que consta en la base de datos es de 120 que puede ser ampliada, así como también se puede agregar el deletreo de las palabras que no se encuentren en la base de datos.
- Las animaciones de las señales no deben ser muy pesadas para que la ejecución de la seña no se demore y se pueda obtener mejores resultados en el sistema traductor.

