



**INSTITUTO POLITECNICO NACIONAL**  
*ESCUELA SUPERIOR DE INGENIERÍA MECÁNICA Y  
ELÉCTRICA*  
*UNIDAD “ADOLFO LÓPEZ MATEOS”, ZACATENCO*



---

---

**“DESARROLLO DE UN SISTEMA DE RECONOCIMIENTO DE  
VOZ PARA UN AUTOMÓVIL”**

**T E S I S**

PARA OBTENER EL TÍTULO DE  
**INGENIERO EN COMUNICACIONES Y ELECTRÓNICA**

P R E S E N T A N :

*JUAN ALBERTO CONDE BALANDRANO  
ALAN LEONARDO CRUZ ISLAS*

**ASESORES:**

*M. en C. MARCIAL MARGARITO SÁNCHEZ SÁNCHEZ  
M. en C. ITZALÁ RABADÁN MALDA*

*CIUDAD DE MÉXICO, FEBRERO 2019*

**INSTITUTO POLITÉCNICO NACIONAL**  
**ESCUELA SUPERIOR DE INGENIERÍA MECÁNICA Y ELÉCTRICA**  
**UNIDAD PROFESIONAL “ADOLFO LÓPEZ MATEOS”**

**TEMA DE TESIS**

QUE PARA OBTENER EL TÍTULO DE INGENIERO EN COMUNICACIONES Y ELECTRÓNICA  
POR LA OPCIÓN DE TITULACIÓN TESIS COLECTIVA Y EXAMEN ORAL INDIVIDUAL  
DEBERA (N) DESARROLLAR C. ALAN LEONARDO CRUZ ISLAS  
C. JUAN ALBERTO CONDE BALANDRANO

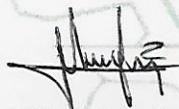
**“DESARROLLO DE UN SISTEMA DE RECONOCIMIENTO DE VOZ PARA UN AUTOMÓVIL”**

DISEÑAR UN SISTEMA PARA EL CONTROL DE UN AUTOMÓVIL, POR MEDIO DEL RECONOCIMIENTO DE VOZ DESDE UN DISPOSITIVO MÓVIL.

- ❖ INTRODUCCIÓN TEÓRICA
- ❖ DESARROLLO DEL PROYECTO
- ❖ PRUEBAS Y RESULTADOS

CIUDAD DE MÉXICO, A 05 DE FEBRERO DEL 2019.

**ASESORES**

  
M. EN C. MARCIAL MARGARITO SÁNCHEZ  
SÁNCHEZ

  
M. EN C. ITZALÁ RABADÁN MALDA

  
ING. GABRIEL VEGA REYES  
JEFE DEL DEPARTAMENTO DE  
INGENIERÍA EN COMUNICACIONES Y ELECTRÓNICA



**INSTITUTO POLITÉCNICO NACIONAL**

**PRESENTE**

Bajo protesta de decir la verdad los que suscriben, **Juan Alberto Conde Balandrano** y **Alan Leonardo Cruz Islas**, manifestamos ser autores y titulares de los derechos morales y patrimoniales de la obra titulada “**Desarrollo de un sistema de reconocimiento de voz para un automóvil**”, en adelante “**La Tesis**” y de la cual se adjunta copia en un **impreso y un CD**, para efecto de **resguardo de la biblioteca y para cumplir con el trámite de titulación**. Por lo que por medio de la presente y con fundamento en el artículo 27 de la Ley Federal del Derecho de Autor, se prohíbe el uso y/o explotación de la “**La Tesis**” en las formas y medios descritos en el fundamento legal citado, en virtud de que cualquier utilización por una persona física o moral distinta de los autores puede afectar o violar derechos autorales, industriales, distintos a los autores de la “**La Tesis**”.

En virtud de lo anterior, **EL IPN** deberá reconocer en todo momento nuestra calidad de autores de la “**La Tesis**” y limitarse a su uso en la forma arriba señalada.

Ciudad de México a 29 de marzo de 2019

**Atentamente**



**Juan Alberto Conde Balandrano**



**Alan Leonardo Cruz Islas**

## **INDICE.**

<b><i>Justificación</i></b>	<b><i>IV</i></b>
<b><i>Objetivo</i></b>	<b><i>V</i></b>
<b><i>Criterios de diseño</i></b>	<b><i>VI</i></b>
<b><i>Antecedentes</i></b>	<b><i>VII</i></b>
<b><i>CAPITULO I. INTRODUCCIÓN TEÓRICA.</i></b>	<b><i>1</i></b>
<i>Introducción.</i>	<i>1</i>
1.1 <i>Fundamentos de la voz.</i>	<i>2</i>
1.1.1 <i>Frecuencia voz hablada.</i>	<i>2</i>
1.1.2 <i>Frecuencia fundamental de la voz.</i>	<i>3</i>
1.1.3 <i>Características subjetivas de un sonido.</i>	<i>4</i>
1.1.4 <i>Formantes.</i>	<i>5</i>
1.1.5 <i>Presión acústica.</i>	<i>6</i>
1.2 <i>Reconocimiento de voz.</i>	<i>7</i>
1.2.1 <i>Características de los sistemas de reconocimiento de voz.</i>	<i>7</i>
1.2.2 <i>Principios de funcionamiento de los sistemas de reconocimiento de voz.</i>	<i>8</i>
1.2.3 <i>Ventajas y desventajas.</i>	<i>9</i>
1.2.4 <i>Debilidades del sistema de reconocimiento de voz.</i>	<i>9</i>
1.3 <i>Validación del rendimiento del sistema.</i>	<i>10</i>
1.4 <i>Corpus.</i>	<i>13</i>
1.5 <i>Digitalización de la voz.</i>	<i>14</i>
1.6 <i>Ruido.</i>	<i>16</i>
6.1 <i>Niveles de ruido.</i>	<i>17</i>
1.7 <i>Microcontrolador.</i>	<i>18</i>
1.7.1 <i>ATMEL ATmega328.</i>	<i>18</i>
1.7.2 <i>Esquema del microcontrolador.</i>	<i>19</i>
1.7.3 <i>¿Por qué ATmega328?</i>	<i>20</i>
1.8 <i>Bluetooth.</i>	<i>21</i>
1.8.1 <i>Modulo bluetooth HC-05.</i>	<i>21</i>
1.8.2 <i>Características.</i>	<i>22</i>

<b>CAPITULO II. DESARROLLO DEL PROYECTO.</b>	<b>23</b>
2.1 Diagrama a bloques del sistema.	23
2.2 Reconocimiento de voz del dispositivo móvil.	24
2.2.1 Características	24
2.2.2 Compatibilidad y requisitos del sistema.	25
2.2.3 Desarrollo de la aplicación.	25
2.3 Configuración del módulo bluetooth HC-05.	26
2.4 Programación del Microcontrolador ATmega328.	28
2.4.2 Diseño del circuito.	29
2.5 Etapa de potencia.	31
2.6 Ubicación del sistema.	32
<b>CAPITULO III. PRUEBAS Y RESULTADOS</b>	<b>33</b>
3.1 Reconocimiento de voz a texto.	33
3.2 Módulo bluetooth.	35
3.2.1 Test de configuración.	35
3.3 Aplicación móvil	38
3.4 Prototipo Anubis 1.0	42
3.5 Prototipo Anubis 2.0	43
<b>Costos</b>	<b>50</b>
<b>Conclusiones</b>	<b>51</b>
<b>Anexos</b>	<b>53</b>
<b>Referencias</b>	<b>56</b>

## **JUSTIFICACIÓN.**

Es común que a un conductor se le lleguen a extraviar, olvidar o romper las llaves de su automóvil. Con el fin de tener un método alternativo para el control de un vehículo, resulta una mejor opción implementar un sistema capaz de manipular un automóvil por medio de la voz desde un dispositivo móvil, un aparato de uso común e indispensable en la vida de millones de personas.

El propósito principal es que el sistema brinde seguridad y facilidad de uso al conductor, logrando resolver los problemas antes mencionados, además de ahorrar tiempo y esfuerzo al controlar el automóvil.

## **OBJETIVO.**

Diseñar un sistema para el control de un automóvil, por medio del reconocimiento de voz desde un dispositivo móvil.

## **CRITERIOS DE DISEÑO.**

- *Funcionamiento en ambientes con ruido de fondo moderado (60 dB).*
- *Bajo consumo eléctrico (5 w).*
- *Interfaz de fácil uso (Android 5.0 o superior)*
- *Seguro (contraseña, vinculación y comandos de voz)*
- *Accesible.*

## **ANTECEDENTES.**

Durante los últimos años han surgido nuevos sistemas humano-computadora que combinan varias tecnologías para permitir el acceso y transferencia de información a través de la voz. Estas interfaces basadas en voz involucran principalmente la tecnología de reconocimiento y síntesis de voz, siendo el reconocimiento de voz el proceso de transformar una señal a texto, y la síntesis de voz el proceso de transformar el texto a una secuencia de sonidos.

La historia del reconocimiento de voz comienza en el año 1870 cuando Alexander Graham Bell quiso desarrollar un dispositivo capaz de proporcionar la palabra visible para la gente que no escuchara. Bell no tuvo éxito creando este dispositivo, sin embargo, el esfuerzo de esta investigación condujo al desarrollo del teléfono.

Más tarde en los años cuarenta cuando se desarrolló un dispositivo capaz de visualizar la señal acústica sobre papel y que naturalmente fue el espectrógrafo. A partir de este momento comenzó a vislumbrarse la posibilidad de realizar sistemas para el reconocimiento automático de voz. En 1950 después del intento de Bell, de crear la primera máquina de reconocimiento de voz. La investigación fue llevada a los laboratorios de AT&T. El sistema tuvo que ser entrenado para reconocer el discurso de cada locutor individualmente, una vez especializada la maquina tenía una exactitud de un 99% de reconocimiento.

En 1952 cuando K.H. Davis, R. Biddulph, S. Balashek, de los Laboratorios Bell construyen el primer dispositivo de reconocimiento, capaz de discriminar con cierta precisión los diez dígitos ingleses pronunciados de forma aislada por un único locutor. El dispositivo era totalmente electrónico gracias al creciente avance de la tecnología.

Durante los 60's, se comenzaron a realizar aplicaciones para los sistemas de reconocimiento de voz, desde ese momento el reconocimiento de voz ha estado bajo intensas investigaciones y desarrollos. Desde no poder reconocer más que un conjunto pequeño de palabras aisladas (típicamente dígitos y unos pocos comandos o palabras clave) pronunciadas por un único locutor, hasta reconocer vocabularios medios y grandes por múltiples locutores. Estos sistemas han hecho grandes avances, en la industria militar y de la salud, lo que ha dado paso a la automatización de muchos procesos.

# ***CAPITULO I. INTRODUCCIÓN TEÓRICA.***

## **INTRODUCCIÓN.**

La comunicación verbal es la forma más habitual de transmitir información entre personas. La inteligencia artificial permite la comunicación verbal entre seres humanos y computadoras, capaces de procesar la señal de voz emitida por el ser humano y reconocer la información contenida en ésta, convirtiéndola en texto o emitiendo órdenes previamente establecidas. El reconocimiento de voz puede controlar diferentes sistemas, mediante el uso de órdenes que pueden ser muy simples, o muy complejas.

Actualmente existen sistemas de reconocimiento de voz con funciones distintas que han ido tomando un papel de gran importancia en la sociedad y tienen como finalidad lograr seguridad al usuario. El reconocimiento de voz es un tema de tecnología conocida que está en plena expansión comercial y cada vez se están buscando más aplicaciones que se le pueden dar para facilitarnos la vida en el trabajo o en la rutina diaria.

Se espera que en un futuro el reconocimiento de voz, sea de utilidad constante, no sólo en las industrias, sino también dentro de nuestros hogares, ya que sus aplicaciones pueden ser variadas y es de gran conveniencia al ahorrarnos tiempo y esfuerzo en la realización de las tareas que efectuamos a diario.

## 1.1 FUNDAMENTOS DE LA VOZ.

La voz es el sonido que producimos cuando el aire expelido de los pulmones pasa a través de la laringe; justo allí se encuentran alojadas las cuerdas vocales y cuando están en contacto con el aire entran en un proceso de vibración. Cuando el aire sale por nuestra boca lo hace ya en forma de sonido. Y es tal vez el principal de los muchos sonidos que percibimos.

Los sonidos consisten en variaciones en la presión del aire a través del tiempo y a frecuencias que son percibidas por el oído humano. Una de las maneras de representar el sonido es a través de una gráfica, en la Figura 1.1 se muestra la gráfica de una señal de voz.

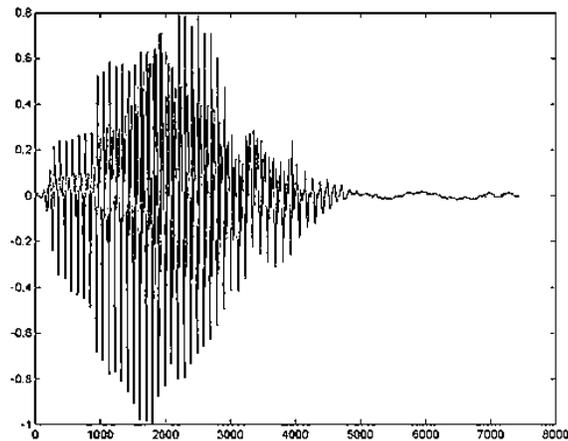


Figura 1.1 Representación gráfica de una señal de voz.

*Los límites de audición son frecuencias que pueden estimular al oído humano para que sean percibidas en el cerebro como una sensación acústica. Estos límites de frecuencia se extienden de aproximadamente **20 Hz** hasta los **20 KHz**. Y de **0 dB** a **120 dB** para presión sonora.*

### 1.1.1 FRECUENCIA VOZ HABLADA.

La inteligibilidad oral se debe a las altas frecuencias. Para que el habla sea comprensible, el rango de frecuencias conversacionales de la voz humana está entre 250 Hz y 3000 Hz, algunos fonemas se encuentran cerca de los 4000 Hz y los 8000 Hz.

### 1.1.2 FRECUENCIA FUNDAMENTAL DE LA VOZ.

La frecuencia fundamental es la frecuencia de vibración de las cuerdas vocales, esta frecuencia brinda información sobre la velocidad a la que vibran las cuerdas vocales al producir un sonido. Esta vibración casi periódica es diferente en el caso de los hombres que al de las mujeres.

A partir de una serie de estudios sobre la frecuencia fundamental de la voz (*Russell, Penny y Pemberton, 1945*), encontraron que la frecuencia fundamental masculina está entre los 100 y los 150 Hz, mientras que la frecuencia fundamental femenina está entre los 200 y los 250 Hz.

Llegando a la conclusión que en la frecuencia fundamental masculina hubo una disminución del tono de habla en la adultez temprana y media, luego la frecuencia cada vez es mayor en las últimas edades como se muestra en la Tabla 1.1. Y para la frecuencia fundamental femenina hubo un tono agudo a temprana y mediana edad, y en la edad adulta la frecuencia disminuye como se muestra en la Tabla 1.2.

Tabla 1.1 Frecuencia fundamental masculina.

<b>Rango de edad</b>	<b>Frecuencia de voz (Hz)</b>
20 – 29	119.5
30 – 39	112.2
40 – 49	107.1
50 – 59	118.4
60 – 69	112.2
70 – 79	132.2
80 - 89	146.3

Tabla 1.2 Frecuencia fundamental femenina.

<b>Rango de edad</b>	<b>Frecuencia de voz (Hz)</b>
20 – 29	227.1
30 – 39	214.2
40 – 49	213.5
50 – 59	214
60 – 69	209.3
70 – 79	206.4
80 - 89	197

### 1.1.3 CARACTERÍSTICAS SUBJETIVAS DE UN SONIDO.

Las cualidades o parámetros que permiten diferenciar un sonido de otro son *intensidad, duración, tono, timbre y pitch*. (Figura 1.2).

- **Intensidad:** La intensidad hace referencia a la fuerza con que un sonido llega a nuestro oído. La intensidad depende de la amplitud de la vibración. De esta manera tendremos sonidos fuertes, medios o débiles.
- **Duración:** Permite diferenciar entre un sonido corto y un sonido largo cuando el resto de sus cualidades o parámetros son idénticos.
- **Tono:** Permite la distinción entre un sonido grave y un sonido agudo. Las frecuencias altas o frecuencias bajas.
- **Timbre:** Esta cualidad permite reconocer a una persona por su voz, aun cuando posean la misma intensidad y duración. El timbre no es medible, sólo es descriptible.

No existen sonidos puros (*de una sola frecuencia*). Los sonidos reales son compuestos de vibraciones de muchas frecuencias, precisamente la proporción de las frecuencias componentes es lo que caracteriza el timbre de los sonidos.

- **Pitch:** El pitch es la capacidad del oído o sistema auditivo para ubicar un sonido en una escala no necesariamente musical.

Aunque el tono y la intensidad del habla están determinados principalmente por la vibración de las cuerdas vocales, su espectro está fuertemente determinado por las resonancias del tracto vocal. Los picos que aparecen en el espectro sonoro de las cuerdas vocales, independientemente del tono, se denominan *formantes*.

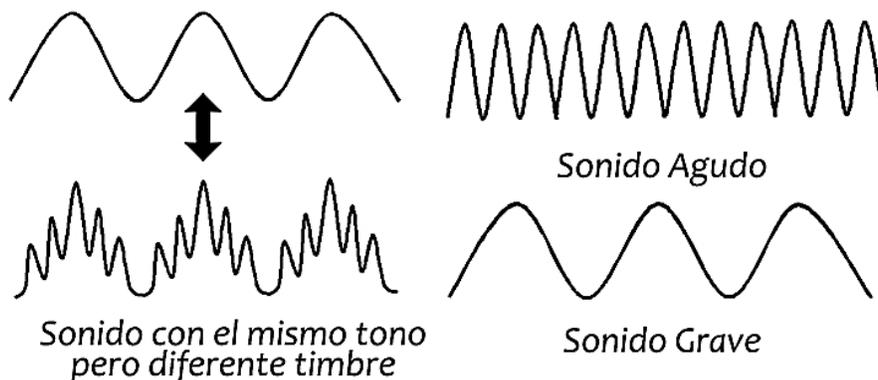


Figura 1.2 Características del sonido.

### 1.1.4 FORMANTES.

Las formantes son bandas de frecuencia donde se concentra la mayor parte de la energía sonora. Permiten distinguir los sonidos del habla humana, sobre todo las cuerdas vocales. También sirven para los sistemas de reconocimiento de voz en la autenticación de la voz, ya que cada sonido del habla humana tiene una marca característica de formantes, es decir, hace un reparto diferente de la energía sonora entre los diferentes formantes, lo cual permite clasificarlos o categorizarlos.

Los formantes del sonido de la voz humana son particularmente importantes porque son componentes esenciales en la inteligibilidad del habla. La distinción de los sonidos vocales se puede atribuir a las diferencias en sus tres primeras frecuencias formantes.

Una onda sonora puede descomponerse en un gráfico espectral. Este tipo de gráficos contienen la información acerca de las frecuencias que integran un sonido y cuáles son sus respectivas amplitudes.

En la Figura 1.3 se puede apreciar el espectro correspondiente a una señal de voz, siendo la primera la correspondiente a la frecuencia fundamental o primer armónico  $F_0$ , después se puede distinguir bien tres formantes;  $F_1$ ,  $F_2$  y  $F_3$ .

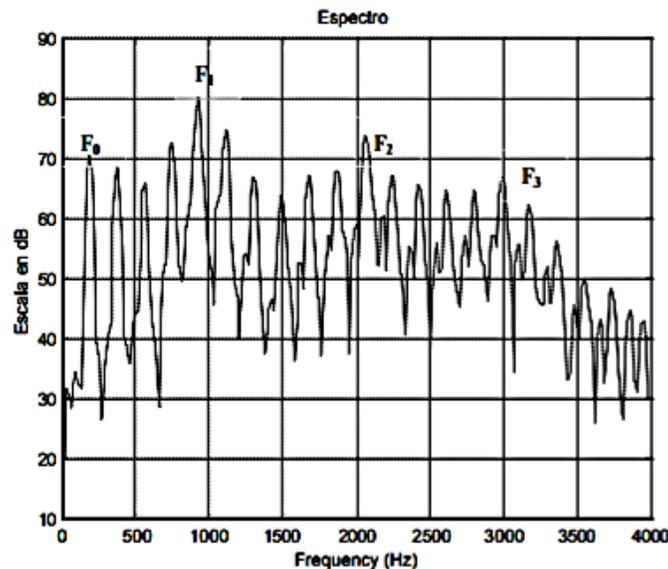


Figura 1.3 Espectro de una señal de voz.

### 1.1.5 PRESIÓN ACÚSTICA.

Es la diferencia de presión producida por un locutor, y la presión estática del entorno, es decir, la presión atmosférica que es igual a 1,013 milibar o 760 mm de Hg (milímetros de mercurio). Su unidad de medida es el pascal. *Pa*; Las presiones sonoras audibles varían entre los 20  $\mu$ Pa y los 20 Pa. Normalmente se expresa la presión sonora en decibels y se denomina Nivel de Presión Sonora (*NPS*). El nivel de referencia en decibels corresponde a 0 dB mientras que el nivel máximo corresponde a 120 dB. El rango de audición es entonces de 120 dB. (*Tabla 1.3*).

*Tabla 1.3 Gama de niveles de intensidad de la voz.*

<i>EMISION</i>	<i>NIVEL SONORO (dB).</i>
<i>Nivel mínimo de la voz humana</i>	20
<i>Mujer conversando en voz baja</i>	25
<i>Hombre conversando en voz baja</i>	30
<i>Mujer conversando en voz normal</i>	50
<i>Hombre conversando en voz normal</i>	55
<i>Mujer hablando en publico</i>	60
<i>Hombre hablando en publico</i>	65
<i>Mujer hablando esforzándose</i>	70
<i>Hombre hablando esforzándose</i>	75
<i>Mujer cantando</i>	80
<i>Hombre cantando</i>	85
<i>Nivel máximo de la voz humana</i>	90

## **1.2 RECONOCIMIENTO DE VOZ.**

El reconocimiento de voz es la capacidad de convertir, la voz humana a un código comprensible por la computadora. En este caso se convierte en caracteres.

### **1.2.1 CLASIFICACIÓN DE LOS SISTEMAS DE RECONOCIMIENTO DE VOZ.**

Los sistemas de reconocimiento de voz se clasifican de acuerdo a las siguientes restricciones:

#### *A. Dependencia vs. independencia del locutor.*

En un sistema dependiente del locutor, solo se reconoce el vocabulario de una persona en particular.

Cuando es independiente del locutor se trata de un sistema que puede reconocer la voz de cualquier persona.

#### *B. Palabras aisladas vs. habla continua.*

El habla continua es cuando un locutor habla de forma natural, con las pausas que son necesarias al hablar.

Sin embargo cuando son palabras aisladas, el locutor habla más lento, esto va a dar como resultado que la probabilidad de reconocimiento sea mayor.

#### *C. Tamaño del vocabulario.*

Otro factor muy importante es el tamaño del vocabulario, ya que a medida que va creciendo el vocabulario también va aumentando la dificultad para reconocer la similitud entre palabras que se parezcan, por ejemplo, cuando se confunde una palabra por otra o cuando tarda más en reconocer.

#### *D. Variabilidad y ruido.*

El ruido puede degradar bastante nivel de desempeño del reconocedor, y este es producido por: el ruido ambiental, suspiros, música, etc. Otros factores son el ruido producido por el locutor, la calidad del micrófono, entre otros.

### **1.2.2 PRINCIPIOS DE FUNCIONAMIENTO DE LOS SISTEMAS DE RECONOCIMIENTO DE VOZ.**

Este tipo de sistemas deben ser capaces de trabajar de tres formas distintas:

- a) **Entrenamiento:** En esta fase se registran las palabras de referencia correspondientes.
- b) **Funcionamiento:** En esta fase a partir de la señal de voz, el sistema tomará decisiones acerca del reconocimiento del vocabulario.
- c) **Actualización:** En esta fase el sistema tendrá la opción de incorporar o dar de baja palabras o frases, así como poder actualizar o mejorar modelos del sistema.

#### *ENTRENAMIENTO DEL SISTEMA.*

Para el correcto funcionamiento del sistema de reconocimiento de voz es necesaria una fase previa a la fase de funcionamiento. En esta fase, denominada entrenamiento, se registra las palabras que va a reconocer inicialmente el sistema, siendo almacenados en la base de datos.

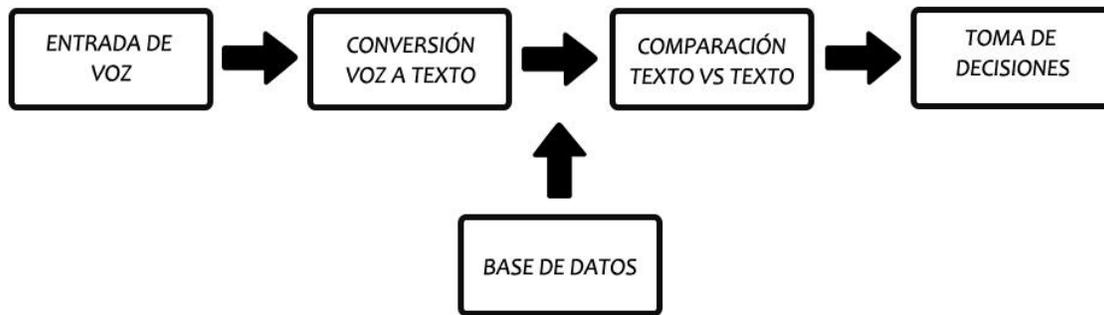
#### *FUNCIONAMIENTO DEL SISTEMA.*

Una vez obtenida la señal de voz, la primera tarea del sistema será procesarla para obtener de forma eficiente la información presente en la señal de voz. Esta información es convertida en texto para después almacenarla y compararla con el texto almacenado en la base de datos, para hacer la comparación entre la palabra obtenida en ese momento y la palabra almacenada.

Por último, el sistema deberá tomar una decisión acerca de la semejanza de las palabras. Este proceso se realiza en el módulo de toma de decisiones que es de mayor importancia ya que es la única salida observable por el usuario. En la Figura 1.4 se puede observar el diagrama.

#### *ACTUALIZACIÓN DEL SISTEMA.*

Durante la vida útil del sistema, éste deberá ser capaz de incorporar nuevas palabras o frases y dar de baja a otras, opcionalmente actualizar o mejorar el sistema.



*Figura 1.4 Diagrama a bloques de un sistema de reconocimiento de voz.*

### **1.2.3 VENTAJAS Y DESVENTAJAS.**

#### **VENTAJAS**

- Ahorro de tiempo al ingresar.
- Menor trabajo para encender los sistemas internos.
- Pequeñas dimensiones.
- Fácil manejo para el usuario.
- Seguridad.
- Fiabilidad.

#### **DESVENTAJAS**

- Problemas de entendimiento en ambientes donde existe altos niveles de ruido.
- Fluidez del habla en las personas.
- Precisión al distinguir palabras erróneas.

### **1.2.4 DEBILIDADES DEL SISTEMA DE RECONOCIMIENTO DE VOZ.**

El análisis de la señal de voz y su posterior reconocimiento deben superar algunos problemas, algunos de ellos son, ruido y distorsión de los entornos donde se utiliza el sistema, inclusive hasta la correcta pronunciación, el comportamiento, el estado de ánimo y la salud del usuario. En el sistema se ha de tener muy en cuenta la variabilidad que posee la señal de voz, pues el individuo no puede repetir de forma completamente exacta una palabra o frase.

Aunque el reconocimiento de voz puede ser una buena opción para resolver problemas de seguridad debido sobre todo a su fácil implementación, se ha de tener en cuenta la posible susceptibilidad a las variaciones del micrófono o del ruido que se pueda generar dentro del sistema puede dar lugar a un aumento de la tasa de falsos positivos.

### 1.3 VALIDACIÓN DEL RENDIMIENTO DEL SISTEMA.

El rendimiento de un sistema de reconocimiento de voz conlleva tomar reiteradas pruebas que se promedian para definir una medida de referencia. Posteriormente, cuando el usuario intente autenticarse, se tomará una o varias pruebas, que se compararan con las medidas de referencia que el sistema tenga registradas, para determinar si existe un elevado nivel de concordancia con alguna de ella.

Una vez que la prueba se ha aplicado, varios índices se utilizan para evaluar la prueba:

- Verdadero positivos (VP): Número de casos que la prueba declara positivos y que son verdaderamente positivos.
- Falsos positivos (FP): Número de casos que la prueba declara positivos y que en realidad son negativos.
- Verdadero negativo (VN): Número de casos que la prueba declara negativos y que son realmente negativos.
- Falsos negativos (FN): Número de casos que la prueba declara negativos y que en realidad son positivos.
- Tasa de falsos positivos (FPR): Proporción de casos negativos que la prueba detecta como positivos.
- Tasa de falsos negativos (FNR): Proporción de casos positivos que la prueba detecta como negativo.
- Número de pruebas (N): Número total de veces que se realizaron las medidas.

*De manera que:*

$$N = VP + FP + FN + VN \quad (1.2)$$

Los falsos positivos, con frecuencia son atribuidos a fallas en los dispositivos o sus programas. Una falsa coincidencia no es un problema tan grave, puesto es muy difícil que se presente una de ellas, sin embargo, siempre hay una posibilidad, por más remota que sea.

Para un sistema de reconocimiento de voz, debe considerarse la tasa de falsos negativos (tasa de rechazos incorrectos, o recuento de no aceptaciones de la autenticación de usuarios válidos), y la tasa de falsos positivos (tasa de aceptaciones incorrectas, o recuento de las ocasiones en que se da por válida la autenticación de un usuario inválido). Cuando más bajo sea la tasa de error igual, más efectivo es el mecanismo de autenticación.

Los índices de evaluación se calculan de la siguiente manera:

- Sensibilidad (equivalente a la tasa de positivos verdaderos): Proporción de casos positivos que están bien detectadas por la prueba. La definición matemática es:

$$\text{Sensibilidad} = VP / (VP + FN). \quad (1.3)$$

- Especificidad (también llamada Tasa de verdaderos negativos): proporción de casos negativos que son bien detectadas por la prueba. La definición matemática es:

$$\text{Especificidad} = VN / (VN + FP). \quad (1.4)$$

- Prevalencia: la frecuencia relativa de los acontecimientos de interés en la muestra total.

$$\text{Prevalencia} = (VP + FN) / N. \quad (1.5)$$

- Valor Predictivo Positivo (VPP): Proporción de casos verdaderamente positivos entre los casos positivos detectados por la prueba. Se tiene:

$$\text{PPV} = TP / (TP + FP) \quad (1.6)$$

ó

$$\text{PPV} = \frac{\text{Sensibilidad} * \text{Prevalencia}}{[(\text{Sensibilidad} * \text{Prevalencia}) + (1 - \text{Especificidad})(1 - \text{Prevalencia})]}$$

Es un valor fundamental que depende de la prevalencia, un índice que es independiente de la calidad de la prueba.

- Valor predictivo negativo (VPN): Proporción de casos verdaderamente negativos entre los casos negativos detectados por la prueba. Se tiene:

$$\text{VPN} = VN / (VN + FP) \quad (1.7)$$

ó

$$\text{VPN} = \frac{\text{Especificidad} * (1 - \text{Prevalencia})}{[(\text{Especificidad} * (1 - \text{Prevalencia})) + ((1 - \text{Sensibilidad}) * \text{Prevalencia})]}$$

Este índice depende también de la prevalencia que es independiente de la calidad de la prueba.

- Razón de verosimilitud positiva (LR +): Esta relación indica a qué punto una persona tiene más posibilidades de ser positivo en la realidad cuando la prueba es que está diciendo es positivo. Se tiene:

$$LR + = \frac{\text{sensibilidad}}{(1-\text{especificidad})} \quad (1.8)$$

La RP + es un valor positivo o nulo.

- Razón de verosimilitud Negativa (LR-): Esta relación indica a qué punto una persona tiene más posibilidades de ser negativo, en realidad, cuando la prueba es que está diciendo es positivo. Se tiene:

$$LR- = \frac{(1-\text{sensibilidad})}{\text{Especificidad}} \quad (1.9)$$

El LR-es un valor positivo o nulo.

- Razón de oportunidades: Indica la cantidad de un individuo es más probable que sea positivo si el resultado es positivo, en comparación con los casos en que la prueba es negativa.

Por ejemplo, una razón de oportunidades de 2 significa que la probabilidad de que el caso positivo se produce es dos veces superior si la prueba es positiva que si es negativo. Se tiene:

$$\text{Razón de oportunidades} = \frac{VP*VN}{FP*FN} \quad (1.10)$$

## 1.4 CORPUS.

*Un corpus lingüístico es un conjunto de textos almacenados en formatos electrónicos y agrupados con el fin de estudiar palabras, significados, contextos, una lengua o una determinada variedad lingüística.*

En el reconocimiento de voz se usa el corpus lingüístico, este tipo de corpus es una colección de textos, las cuales están accesibles en forma de lectura dentro de una computadora.

En los sistemas de reconocimiento de voz, una parte fundamental la constituye el corpus que sirve precisamente para poder entrar y perfeccionar un reconocedor. Entre más grande, preciso, limpio y etiquetado sea el corpus, obtendremos mejores resultados.

Los sistemas de reconocimiento de voz están enfocados en las palabras y los fonemas de cada palabra, los sonidos son los que distinguen una palabra de otra en un mismo idioma. Por ejemplo, "tapa", "capa", "mapa", "napa", son palabras diferentes puesto que su sonido inicial se reconoce como fonemas diferentes en español. Para poder analizar y describir el habla, se debe de tomar en consideración los siguientes análisis:

- *Articulación:* Análisis de cómo el humano produce los sonidos del habla.
- *Acústica:* Análisis de la señal de voz como una secuencia de sonidos.
- *Percepción Auditiva:* Análisis de cómo el humano procesa el habla.

Los tres enfoques proveen ideas y herramientas para obtener mejores y más eficientes resultados en el reconocimiento de voz.

El uso del corpus es indispensable como tal, debe de contar con un orden, esto con la finalidad de facilitar la búsqueda de la información y de seguir una estructura.

Gracias a los constantes avances tecnológicos, tanto en dispositivos de almacenamiento, como en la evolución de interfaces digitales; y de acuerdo con los avances en los sistemas de reconocimiento de voz, un corpus puede contener infinidad de datos.

## 1.5 DIGITALIZACIÓN DE LA VOZ.

La mayor parte de señales de interés práctico, tales como las señales de voz, son analógicas. Cualquier sistema de control basado en un microcontrolador no puede interpretar señales analógicas, ya que sólo utiliza señales digitales. Para poder procesar señales analógicas a través de medios digitales, es necesario convertirlas a formato digital, es decir, transformarlas en señales binarias. Este procedimiento se denomina conversión analógico-digital (A/D), y los dispositivos que la realizan son convertidores A/D (ADC). La conversión A/D es un proceso de tres pasos, como se muestra en la Figura 1.6.

1. *Muestreo*. El muestreo consiste en tomar muestras de una señal analógica en un periodo de tiempo (usualmente en 1 segundo), para posteriormente cuantificarlas.

### TEOREMA DE NYQUIST.

El teorema de Nyquist-Shannon afirma que una señal analógica puede ser reconstruida, sin error, de muestras tomadas en iguales intervalos de tiempo. La razón de muestreo debe ser igual, o mayor, al doble de su ancho de banda. (Figura 1.5).

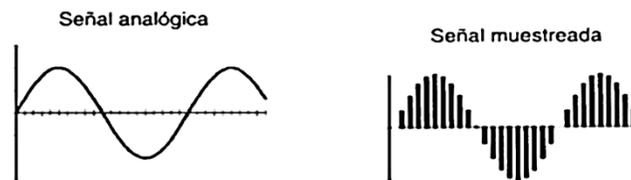


Figura 1.5 Teorema de Nyquist.

2. *Cuantificación*. En esta parte del proceso los valores continuos de la señal analógica se convierten en series de valores numéricos decimales discretos correspondientes a los diferentes niveles o variaciones de la amplitud que contiene la señal analógica original.

Por tanto, la cuantificación representa el componente de muestreo de las variaciones de valores de amplitud tomados en diferentes puntos de la señal, que permite medirlos y asignarles sus correspondientes valores en el sistema numérico decimal, antes de convertir esos valores en sistema numérico binario.

3. *Codificación*. Consiste en traducir los valores obtenidos durante la cuantificación en valores numéricos binarios equivalentes a los valores de amplitud que conforman la señal analógica original. El código más utilizado es el código binario, pero también existen otros tipos de códigos que son empleados.

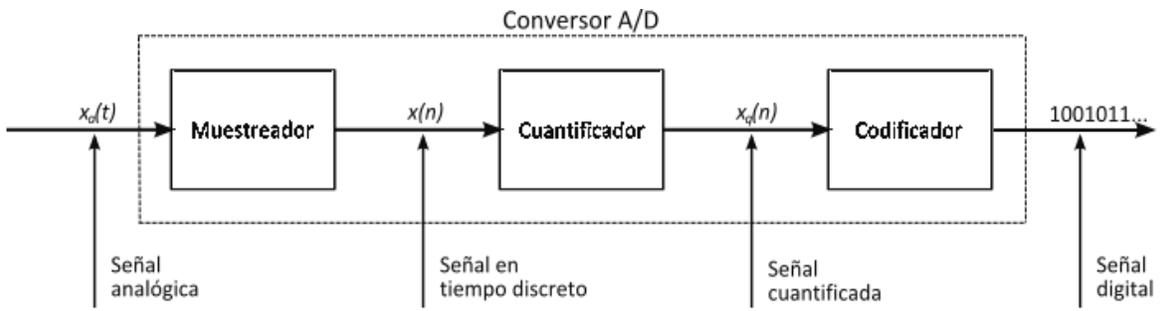


Figura 1.6 Partes básicas de un convertidor analógico-digital (A/D).

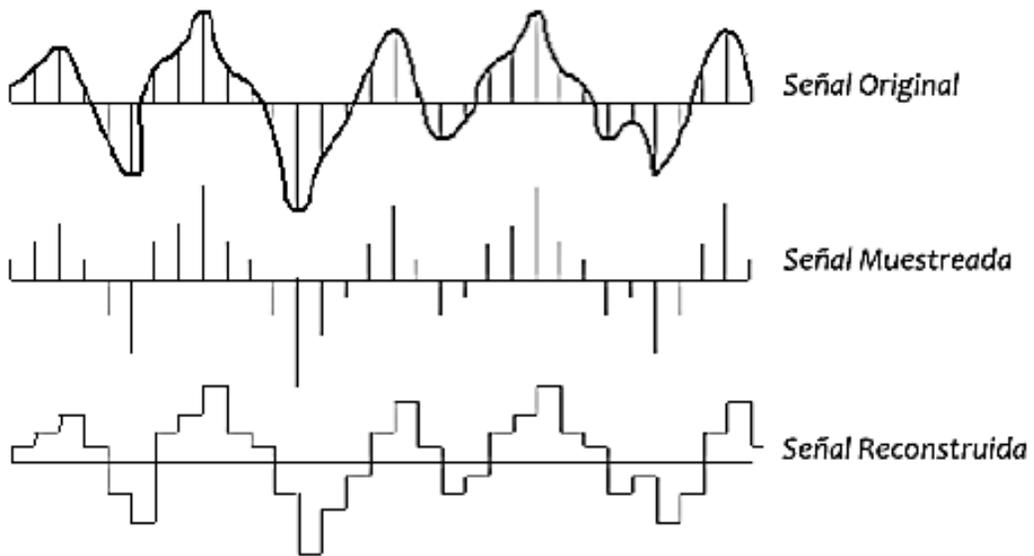


Figura 1.7 Conversión analógico digital (A/D).

## 1.6 RUIDO.

Existen muchos factores ambientales que afectan la correcta percepción de las señales de voz, tales como *el ruido y la calidad del micrófono*.

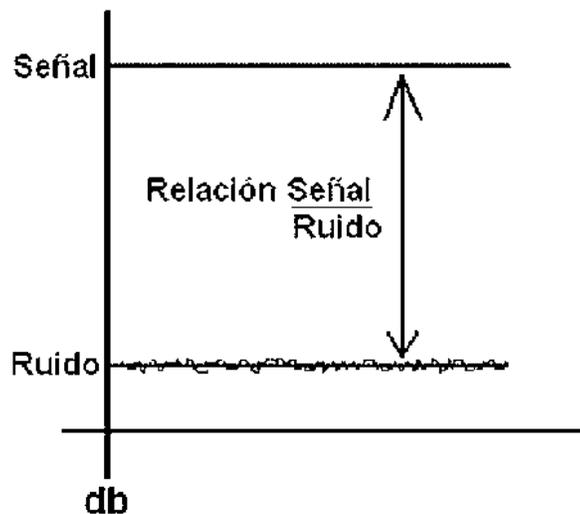
### - RUIDO.

Un ruido es una mezcla compleja de sonidos de varias frecuencias. Se define como aquellos sonidos aleatorios que de forma “oculta” transforman y enmascaran el sonido.

Es muy importante saber que ruido acústico, es igual a, aire en movimiento. Por lo tanto, donde pasa el aire, pasa el ruido. El ruido es una característica inherente a toda señal analógica y a los tratamientos electrónicos que pueda sufrir. Dado que es poco probable encontrar un entorno en perfecto silencio, es importante conocer la cantidad de ruido, en relación con la señal que se introduce en el sistema. La cantidad de señal comparada con la cantidad de ruido de fondo de una señal en particular (*ver Figura 1.8*), se conoce como relación señal a ruido (SNR). La relación señal-ruido matemáticamente está definida cómo:

$$SNR = \frac{P_s}{P_n} \quad (1.11)$$

Donde  $P_s$  es la potencia media de la señal deseada y  $P_n$  es la potencia media del ruido indeseado. Estos niveles de potencia deben medirse en los mismos puntos y en el mismo ancho de banda del sistema.



*Figura 1.8 Relación señal a ruido.*

Una buena relación señal a ruido es primordial en cualquier tipo de sistema, una SNR mayor, indica un ruido de fondo menos perceptible.

No obstante, sin ruido externo, el sistema es susceptible a captar ruido a través del micrófono, y aunque suene extraño, muchas veces el ruido proviene de la boca durante la pronunciación del mensaje.

- **MICRÓFONO.**

Los ruidos se producen en el origen mismo de la señal, ya que el micrófono no es perfecto. En el caso de los sonidos bilabiales, el micrófono es muy susceptible de detectar ráfagas de aire ocasionadas por los sonidos bilabiales lo que ocasiona ruido durante la pronunciación del mensaje.

Otras fuentes de ruido externo, tales como los automóviles, camiones, teléfonos, personas hablando y el ruido ambiente puede causar problemas con la exactitud del sistema. Las alteraciones pueden ser tan potentes que incluso se produzca la pérdida total de la información que se pretendía obtener.

### **1.6.1 NIVELES DE RUIDO.**

*Tabla 1.4 Niveles de ruido.*

<b>FUENTE DE SONIDO:</b>	<b>dB</b>
<i>Conversación en voz baja</i>	50
<i>Conversación normal</i>	60
<i>Ruido ambiente</i>	60
<i>Automóvil</i>	70
<i>Estacionamiento</i>	70
<i>Tráfico pesado</i>	80
<i>Teléfono sonando</i>	80
<i>Camión</i>	90
<i>Sirena</i>	90
<i>Martillo mecánico</i>	100
<i>Moto</i>	110
<i>Avión</i>	120

## **1.7 MICROCONTROLADOR.**

Un microcontrolador es un circuito integrado en el cual se pueden grabar instrucciones. Estas instrucciones se escriben utilizando un lenguaje de programación que permite al usuario crear programas que interactúan con circuitos electrónicos. El uso de un microcontrolador *ATmega328* nos permitirá dar mayor eficiencia al sistema y reducir el tamaño del sistema electrónico. (*Figura 1.9*).



*Figura 1.9 Microcontrolador ATmega328.*

### **1.7.1 ATMEL ATMEGA328.**

El ATmega328 es un circuito integrado de alto rendimiento que opera entre 1.8 y 5.5 volts. Su arquitectura le permite ejecutar instrucciones en un solo ciclo de reloj, alcanzando una potencia de 1 MIPS (millones de instrucciones por segundo), balanceando consumo de energía y velocidad de proceso.

Este microcontrolador se puede programar mediante la plataforma IDE (Entorno de Desarrollo Integrado), su lenguaje es derivado de C, es mucho más simple que el utilizado por otros microcontroladores. Se puede usar para hacer funcionar una gran variedad de interruptores y sensores, también para controlar luces, motores y otros actuadores físicos.

Actualmente el microcontrolador ATmega328 se usa comúnmente en múltiples proyectos y sistemas autónomos donde un microcontrolador simple, de bajo consumo de energía y bajo costo es requerido. Sus características son excelentes a pesar de su reducido tamaño, es compatible con accesorios y funciones como bluetooth, infrarrojo y sensores.

### 1.7.2 ESQUEMA DEL MICROCONTROLADOR.

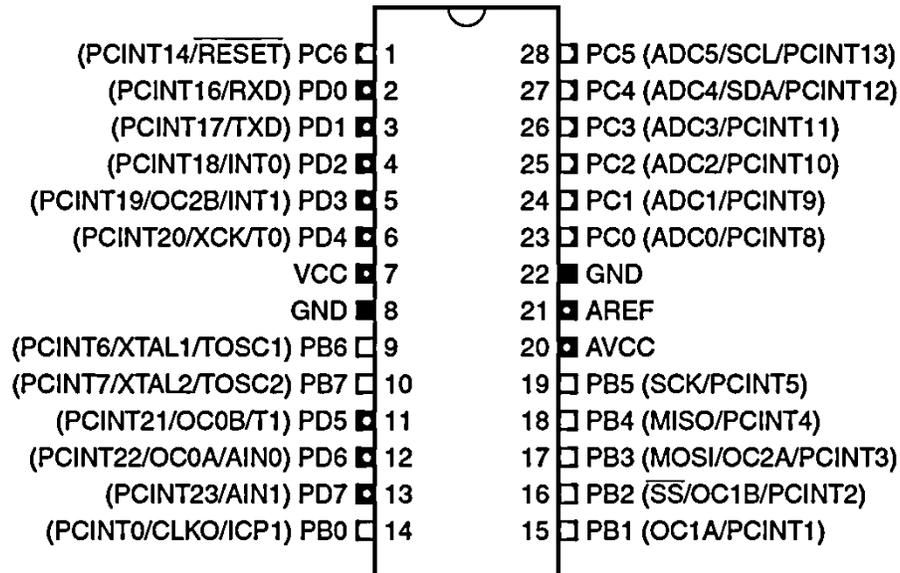


Figura 1.10 Esquema del microcontrolador.

Tabla 1.5 Especificaciones del microcontrolador

PARÁMETROS	VALORES
Flash	32 Kbytes
SRAM	2 Kbytes
Cantidad Pines	28
Frecuencia máxima de operación	20 MHz
CPU	8-bit AVR
Pines máximos de E/S	23/
Interrupciones internas	24
SPI	1
UART	1
Canales ADC	8
Resolución de ADC	10
Eeprom	1K
Canales PWM	6
Voltaje de operación	1.8-5.5 v
Timers	3

### **1.7.3 ¿POR QUÉ ATMEGA328?**

- *COSTOS.*

Una de las principales características del microcontrolador ATmega328 es su bajo costo. Actualmente, adquirir el microcontrolador resulta relativamente fácil y económico, debido a su amplia presencia en el mercado. De igual forma, la inmensa cantidad de sensores y placas de expansión que han sido diseñadas para la compatibilidad con el microcontrolador y sus bajos precios, permite que esta tecnología sea considerada para toda clase de proyectos.

- *DISPONIBILIDAD.*

Como se ha mencionado, el microcontrolador puede ser adquirido a bajo costo en el mercado. Poco a poco, a medida que pasa el tiempo se ha incrementado la presencia de este microcontrolador, dada a su creciente demanda.

- *FIABILIDAD.*

Dada su eficiencia y su gran capacidad para llevar a cabo diferentes funciones, es realmente compacta y es ideal para proyectos profesionales ya que la etapa de control se hace más fácil.

## 1.8 BLUETOOTH.

Bluetooth es una especificación tecnológica para redes inalámbricas que permite la transmisión de voz y datos entre distintos dispositivos mediante una radiofrecuencia segura (2,4 GHz). Los principales objetivos del bluetooth son:

- Facilitar las comunicaciones entre equipos móviles.
- Eliminar los cables y conectores entre estos.
- Facilita la sincronización de datos entre equipos personales.

Para una sincronización entre dos dispositivos bluetooth debe de existir un cliente (maestro) y un servidor (esclavo). El servidor es el dispositivo bluetooth que espera la conexión para recibir información, el cliente es el dispositivo que establece la sincronización y envía la información. Los dispositivos Bluetooth pueden comunicarse entre sí cuando se encuentran dentro de su alcance. Estos dispositivos se clasifican en referencia a su potencia de transmisión.

- Clase 1: 100 metros aproximadamente.
- Clase 2: 10 metros.
- Clase 3: 1 metro.

### 1.8.1 MODULO BLUETOOTH HC-05.

El módulo de bluetooth HC-05 es un módulo maestro-esclavo, en el modo maestro puede conectarse con otros módulos bluetooth, mientras que en el modo esclavo queda a la espera de recibir instrucciones. (Figura 1.11 y 1.12).

Es ideal para aplicaciones inalámbricas, ofrece una mejor relación de precio y características, su funcionamiento es de bajo consumo y es de fácil implementación con microcontroladores.

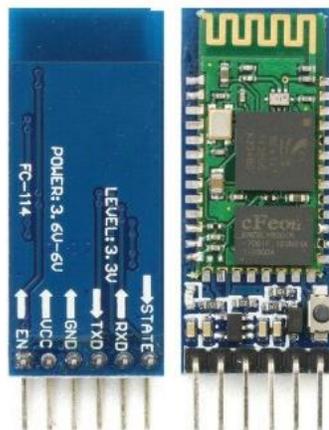
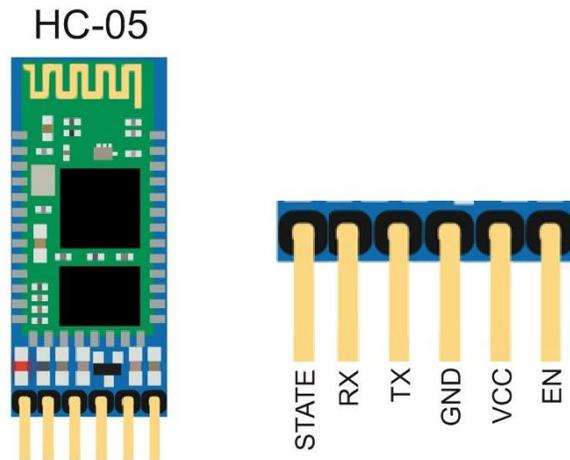


Figura 1.11 Modulo bluetooth HC-05.



*Figura 1.12 Pines del módulo bluetooth HC-05*

### **1.8.2 CARACTERÍSTICAS.**

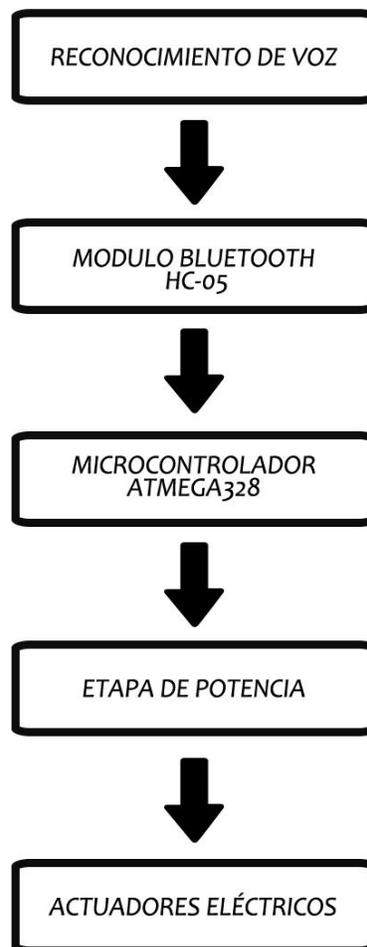
- Compatible con el protocolo Bluetooth V2.0.
- Voltaje de alimentación: 3.3VDC – 6VDC.
- Voltaje de operación: 3.3VDC.
- Frecuencia: 2.4 GHz
- Tasa de baudios por defecto: 9600.
- Alcance: 10 m.
- Tamaño: 4.4 cm x 1.6 cm x 0.7 cm.
- Peso: 30 g.
- Corriente de operación: < 40 mA.
- Corriente modo sleep: < 1mA.
- Seguridad: Autenticación y encriptación.

## **CAPITULO II. DESARROLLO DEL PROYECTO.**

### **2.1 DIAGRAMA A BLOQUES DEL SISTEMA.**

El siguiente diagrama a bloques representa el funcionamiento del sistema, además define la organización, así como el principio y su final.

La voz emitida por el usuario se procesa mediante el reconocimiento de voz de Google, donde se interpreta y convierte la señal de voz emitida en texto, después se envía esta información al módulo bluetooth y es compartida al microcontrolador donde se hace la comparación de la palabra almacenada y la palabra recibida, que, en caso de coincidir con alguna, enviará un pulso a la etapa de potencia donde activará los actuadores eléctricos. (Figura 2.1).



*Figura 2.1 Diagrama a bloques del sistema.*

## 2.2 RECONOCIMIENTO DE VOZ.

Android dispone de una herramienta muy interesante y hay gente que no suele utilizarla, el reconocimiento de voz. El sistema utilizará los motores de reconocimiento de voz de Google, por su precisión de reconocimiento que son niveles superiores al 90% y se mejora con el tiempo, conforme se perfecciona la tecnología interna de reconocimiento de voz de Google. (Figura 2.2).

La voz emitida por el usuario se procesa mediante el reconocimiento de Google, donde se interpreta utilizando algoritmos y convierte la señal de voz emitida en texto, Por supuesto, hay que hablar claro para que la aplicación nos entienda.

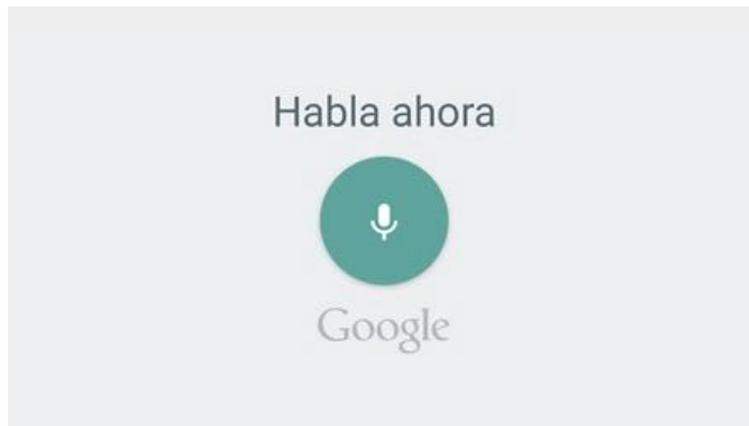


Figura 2.2 Reconocimiento de voz de Google.

### 2.2.1 CARACTERÍSTICAS.

- **Consigue resultados de forma rápida:** Google puede reconocer la palabra emitida por el usuario de manera rápida, casi inmediata, conforme no haya demasiado ruido alrededor.
- **Tratamiento de ruido:** Es capaz de procesar audio de diversos entornos ruidosos sin necesidad de reducción de ruido adicional.
- **Compatible con Android:** Es compatible con cualquier dispositivo Android, incluidos teléfonos y tabletas.
- **Vocabulario internacional:** Reconoce más de 80 idiomas y variantes, con un amplio vocabulario.
- **Reconocimiento en directo:** Proporciona resultados de reconocimiento mientras los usuarios siguen hablando.

## **2.2.2 COMPATIBILIDAD Y REQUISITOS DEL SISTEMA.**

El reconocimiento de voz de Google no requiere ninguna instalación, puede usarse en tabletas y teléfonos compatibles con Android.

En versiones muy antiguas de Android el reconocimiento de voz para funcionar requería tener una conexión de internet, pero actualmente nos permite funcionar sin conexión a internet. Para usar reconocimiento de voz sin conexión se tiene que descargar el vocabulario en el idioma que se utilizará.

Se descarga el idioma desde la *aplicación de Google > Configuración > Voz > Reconocimiento de voz sin conexión > Todo > Español (EE.UU.) > Descargar.*

Una vez descargado el idioma ya podemos usar el reconocimiento de voz en cualquier momento sin depender de internet.

## **2.2.3 DESARROLLO DE APLICACIÓN.**

Para poder utilizar el reconocimiento de voz de Google y que ayude a enviar la información al módulo de bluetooth, se necesita desarrollar una aplicación que sea de fácil manejo para el usuario.

La herramienta en línea App inventor, permite programación en bloques para la elaboración de aplicaciones tanto en lo estético, como el algoritmo que esta interpretará, destinadas al sistema operativo Android.

1. El primer paso es desarrollar un módulo de seguridad el cual servirá para dar acceso solo al usuario a la aplicación donde se emitirán las órdenes para el control del automóvil. Este módulo contará con el nombre del usuario y una contraseña previamente establecida para poder ingresar.
2. El segundo paso es hacer una conectividad entre el módulo bluetooth que trabaja como servidor (esclavo) y el dispositivo móvil con sistema operativo Android que opera como cliente (maestro). Para esto se crea un botón llamado “conectar” que se encargará de sincronizar el módulo bluetooth y el dispositivo móvil.
3. El tercer paso es deshacer el vínculo entre el módulo bluetooth y el dispositivo móvil cuando se acabe de utilizar la aplicación. Se crea un botón de “desconectar” que se encargará de desvincular el módulo bluetooth y el dispositivo móvil.
4. El cuarto paso es habilitar el reconocimiento de voz de Google para poder recibir la señal de voz, convertirla en texto y esta información enviarla al módulo bluetooth. Se crea un botón el cual activará el reconocimiento de voz para después pronunciar la palabra o frase que emitirá la orden.

## 2.3 CONFIGURACIÓN MÓDULO BLUETOOTH.

El módulo bluetooth HC-05 se puede configurar para cambiar parámetros como el nombre del dispositivo, contraseña, velocidad de transmisión, modo maestro/esclavo. Se configura a través de comandos AT enviados desde el monitor serie del IDE (Entorno de Desarrollo Integrado) mediante un convertidor USB-TTL, como se muestra en la Figura 2.3



Figura 2.3 Comunicación modulo bluetooth y USB-TTL.

Un conversor USB-TTL permite la comunicación entre un microcontrolador y un pc mediante el protocolo USB de forma sencilla.

Los comandos “AT” es un lenguaje de estándar abierto de comandos para configurar y parametrizar módems. Los caracteres “AT”, que preceden a todos los comandos, significan “Atención”, e hicieron que se conociera también a este conjunto de comandos como comandos “AT”.

Para comunicarnos con el módulo y configurarlo, es necesario tener acceso al módulo mediante una interfaz serial, en este caso es la plataforma IDE, el módulo recibirá y enviará la información a través de sus puertos TX y RX, como se muestra en la Figura 2.4.

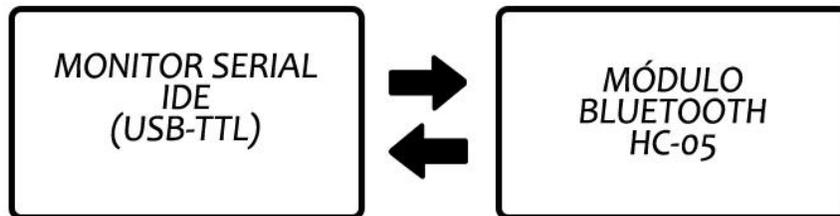


Figura 2.4 Diagrama de comunicación serial.

Para obtener un nivel de seguridad alto en el sistema se cambiará los parámetros, como el nombre del módulo, la contraseña, la velocidad se configura para que trabaje a la misma velocidad que el microcontrolador (9600 baudios) y que trabaje en modo servidor (esclavo) para que solo reciba información.

Para configurar el módulo bluetooth se necesitan los siguientes comandos:

- *Comando para cambiar el nombre del módulo.* Por defecto el módulo bluetooth se llama "HC-05" esto se puede cambiar con el siguiente comando AT:

***AT+NAME=ANUBIS***

- *Comando para cambiar código de vinculación.* Por defecto viene con el código de vinculación (Pin) "1234", para cambiarlo hay que enviar el siguiente comando AT:

***AT+PSWD=0506***

- *Comando para configurar la velocidad de comunicación.* La velocidad por defecto es de 9600 baudios y es la que se necesita para trabajar a la velocidad del microcontrolador.

***AT+UART=9600***

- *Comando para configurar el modo en el que trabajará el módulo.* Por defecto el módulo viene en modo esclavo. Y nos servirá para comunicarse con el dispositivo móvil.

***AT+ROLE=0***

## 2.4 PROGRAMACIÓN DEL MICROCONTROLADOR ATMEGA328

Para que el microcontrolador funcione conforme a las necesidades requeridas, se debe programar para que el microcontrolador lea cada línea de código, la interprete y la realice. La programación del microcontrolador se hará mediante la plataforma IDE y se compone de dos bloques:

- PRIMER BLOQUE (*SETUP*).

El primero contiene instrucciones de configuración, en el cual especificamos la velocidad de comunicación que trabajará el microcontrolador (9600 baudios), la declaración de los pines de salida y su modo de trabajo.

- SEGUNDO BLOQUE (*LOOP*).

El segundo contiene el código que se ejecutará continuamente, en el cual interpretará la información que reciba, revisará si la información es válida y la comparará con las palabras almacenadas, para obtener la similitud entre la palabra recibida y la palabra almacenada y así mandar un alto o un bajo a la salida. El código se ejecutará de forma cíclica, el cual se repite de forma indefinida hasta que se apague o se reinicie el microcontrolador. (Figura 2.5).

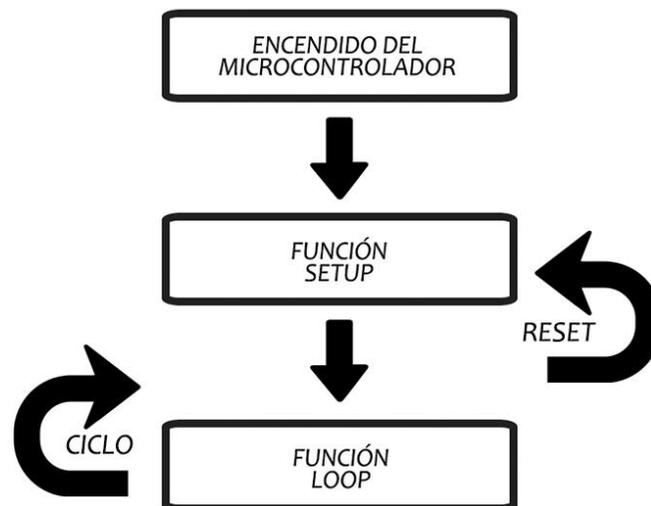


Figura 2.5 Diagrama a bloques de la programación.

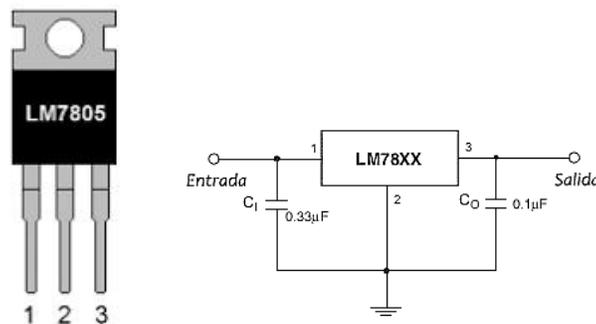
### 2.4.1 DISEÑO DEL CIRCUITO.

El circuito consta de dos etapas, la primera funciona como un regulador de voltaje, la cual nos entregará 5 volts para la alimentación del microcontrolador, ya que la fuente de alimentación será la batería del automóvil la cual nos entrega 12 o 13 volts. Así esta etapa nos reducirá el voltaje para el funcionamiento del microcontrolador.

La segunda etapa consta del microcontrolador en conjunto con el módulo bluetooth HC-05, en donde se estará recibiendo la información enviada por el dispositivo móvil.

#### - PRIMERA ETAPA.

La primera parte consta de un regulador de voltaje positivo LM7805, el cual se alimenta de una cantidad de voltaje determinada y es capaz de entregar una cantidad menor acondicionada para un equipo determinado. En este caso se alimenta de 12 volts de la batería y entrega un voltaje positivo de 5 volts a 1 Ampere de corriente para alimentar al microcontrolador. Debido a que el regulador absorbe el voltaje excedente suministrado, disipa calor es por eso que se utilizara un disipador de calor acorde al regulador que estamos utilizando. Es recomendable el uso de dos capacitores, uno en la entrada y el otro a la salida del regulador para evitar oscilaciones de alta frecuencia generadas internamente durante el proceso de regulación. (ver *Figura 2.6*).



*Figura 2.6 Regulador de voltaje LM7805.*

Como medida de protección, el circuito contará con un fusible de fusión rápida en paralelo con el varistor, de este modo cuando el voltaje supere el valor nominal, que en este caso son 12 volts, toda la corriente circulará a través del varistor, provocando el corte del fusible, impidiendo que la corriente pase a través del resto de los componentes, evitando dañarlos. (ver *Figura 2.7*).

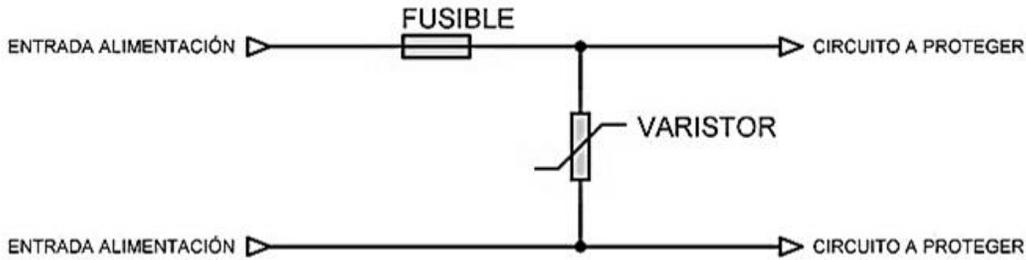


Figura 2.7 Diagrama eléctrico de protección al circuito.

- SEGUNDA ETAPA.

La segunda etapa consta del microcontrolador, el cual necesita componentes adicionales. Consiste de una resistencia acompañado de un botón de regreso para el reinicio del atmega328, un oscilador externo de 16 MHz, dos capacitores en serie a sus extremos para dar mayor estabilidad al microcontrolador, el módulo bluetooth se conecta en la recepción (RX) y transmisión (TX) de datos del ATmega328 y por último los pines de salida. (ver Figura 2.8).

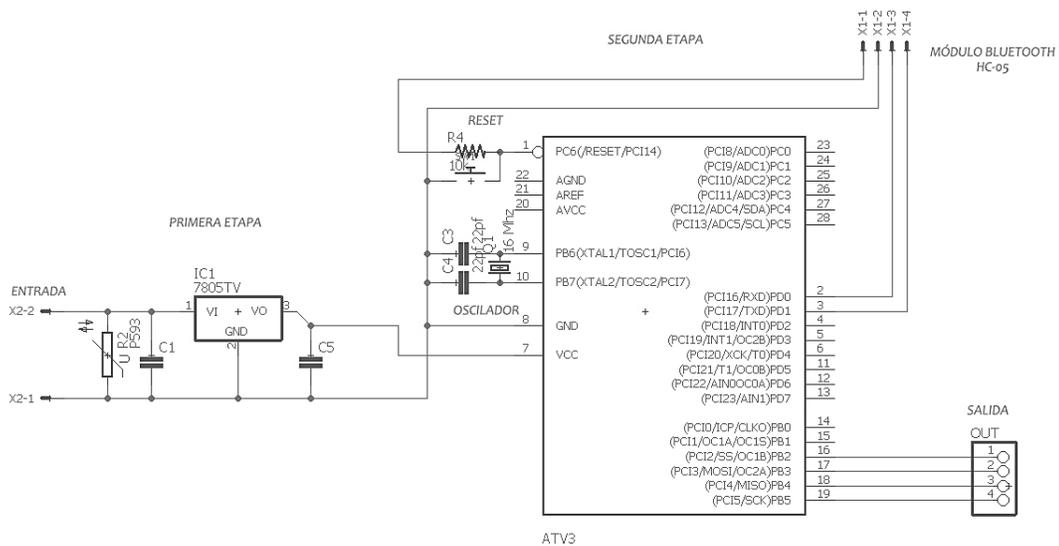


Figura 2.8 Diagrama eléctrico del circuito.

## 2.5 ETAPA DE POTENCIA.

La etapa de potencia consta de un arreglo de transistores, los transistores son dispositivos que pueden funcionar como interruptores en un circuito eléctrico, de forma que mantiene aislado al microcontrolador para evitar daños con la etapa de potencia, permitiendo de ese modo que señales de baja potencia accionen mecanismos de alta potencia, en este caso los actuadores eléctricos.

Los transistores dan la capacidad de un interruptor para poder controlarlo y la resistencia depende de la variación del voltaje de entrada con el que se va a trabajar. (ver Figura 2.9).

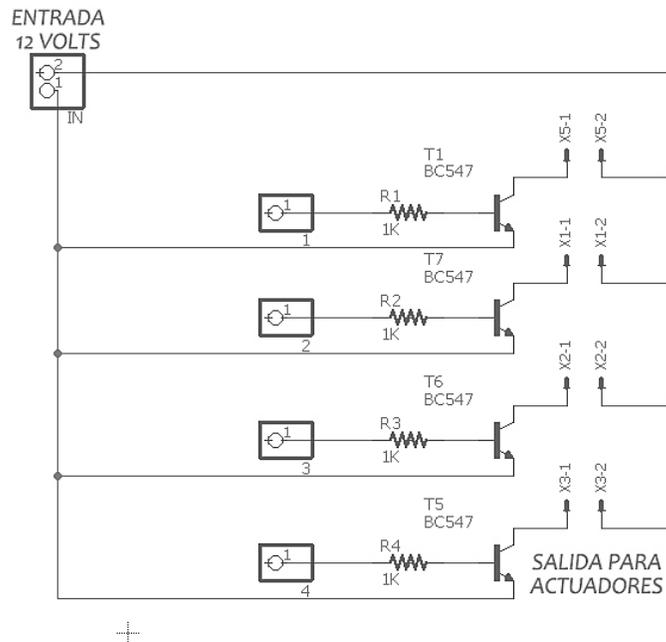


Figura 2.9 Diagrama eléctrico de la etapa de potencia.

El funcionamiento del transistor depende de la cantidad de corriente que pase por su base, Cuando no pasa corriente por la base, no puede pasar tampoco por sus otras terminales (colector-emisor), entonces es como si se tratara de un interruptor abierto y no circula corriente por la etapa de potencia.

Cuando llega un pulso positivo a la base del transistor, permite la circulación de corriente entre el colector y el emisor, en este caso permite el voltaje de 12 volts de la batería del automóvil. En este caso el transistor se comporta como si fuera un interruptor cerrado y permite el paso de corriente de la batería que alimentará a los actuadores eléctricos.

## 2.6 UBICACIÓN DEL SISTEMA.

Se propone que el sistema se encuentre en el interior del automóvil para que las conexiones no afecten con el funcionamiento del automóvil, además se tiene en cuenta que todos los vehículos tienen diferente arquitectura y, por ende, la ubicación del sistema puede variar según el automóvil. A continuación, se presentará un ejemplo de su ubicación en el siguiente vehículo, que debido a sus dimensiones y a las acciones que ejecutará resulta una mejor opción, se ubique debajo del asiento del conductor. (ver Figura 2.10).

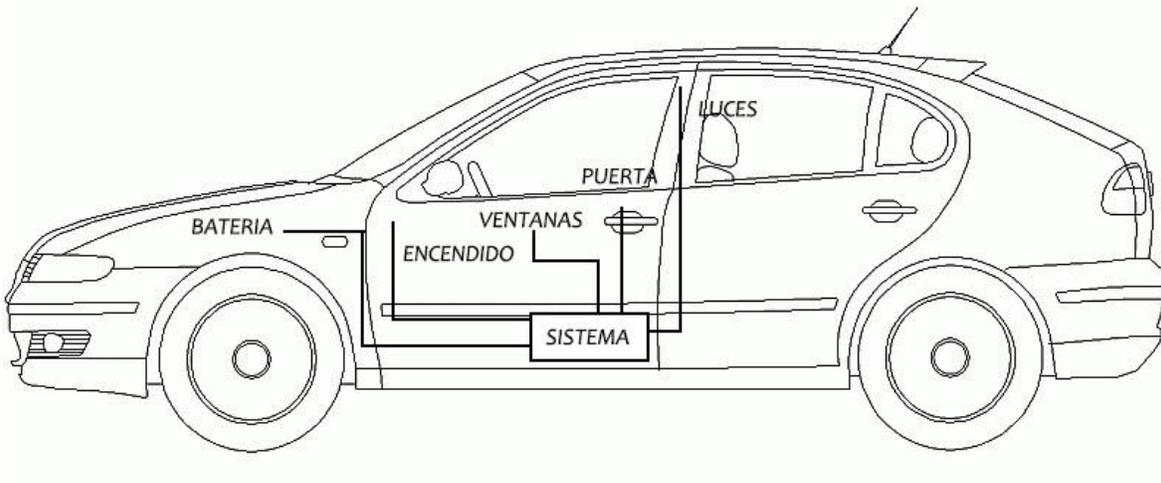


Figura 2.10. Ubicación del sistema en el automóvil.

### **CAPITULO III. PRUEBAS Y RESULTADOS.**

En este Capítulo se realizaron las pruebas de manera individual de acuerdo al diagrama a bloques del sistema (ver Figura 2.1), a excepción del microcontrolador, en cuyo caso las pruebas se hicieron con el prototipo *Anubis 2.0*.

#### **3.1 RECONOCIMIENTO DE VOZ A TEXTO**

La API Speech de la compañía Google®, (*Application Programming Interface por sus siglas en inglés*) o Interface de Programación de Aplicaciones en español, puede transmitir resultados de texto conforme vaya reconociendo la voz, de forma que el texto reconocido aparece inmediatamente mientras la persona habla. De esta manera las pruebas se realizaron en dos partes, y consistieron en dictarle a la API de un teléfono inteligente de la marca *Xiaomi* modelo *Mi A1*, una serie de frases y palabras aleatorias en cinco ocasiones cada una de ellas, con un ruido de fondo de 60 dB con el propósito de conocer los alcances y eficiencia de la misma.

##### **- PARTE UNO**

Las frases usadas fueron las siguientes:

- a) *“Pedro dicta a este equipo. Prefiere dictar que escribir y, sobre todo, lo prefiere al lápiz y papel”.*
- b) *“En el idioma está el árbol genealógico de una nación”.*
- c) *“Pero si el pensamiento corrompe el lenguaje, el lenguaje también puede corromper el pensamiento”.*
- d) *“Saber que se sabe lo que se sabe y saber que no se sabe lo que no se sabe”.*
- e) *“Todo cuerpo permanece en su estado de reposo o de movimiento rectilíneo uniforme a menos que otras fuerzas actúen sobre él”.*

A partir de lo cual se recabó la siguiente información (ver Tabla 3.1).

*Tabla 3.1 Eficiencia del API aplicada a frases.*

<b>Total de Repeticiones</b>	<b>Repeticiones Reconocidas</b>	<b>Eficiencia de la API (%)</b>
25	22	88

Donde los inconvenientes principales fueron, que, al leer frases con pausas determinadas por la puntuación, la API no siempre es capaz de reconocer el contenido en su totalidad.

- *PARTE DOS.*

Las palabras dictadas fueron:

- agua
- pez
- batería
- **luz**
- **cerrar**
- gorro
- **encender**
- persona
- **parar**
- libros
- casar
- manejar
- pavimento
- microcontrolador
- haz
- escápula
- otorrinolaringólogo
- **ventana**
- fugaz
- **abrir**
- universidad
- manglar
- entumir
- manipulación

De las cuales las palabras que están en **negritas**, son los comandos con los que el sistema activará los sistemas del automóvil. Obteniendo los resultados que se muestran en la Tabla 3.2.

*Tabla 3.2 Eficiencia de la API aplicada a palabras*

<b>Total de repeticiones</b>	<b>Repeticiones Reconocidas</b>	<b>Eficiencia de la API (%)</b>
120	103	85.83

En la parte dos del análisis se encontró que la API tiene dificultad para detectar palabras homófonas y palabras monosilábicas que comiencen con consonantes fuertes como las letras *p, t, q*, etc. Asimismo, no se tuvieron inconvenientes en reconocer las palabras que servirán como comandos para el sistema de reconocimiento de voz.

Por ende, se comprueba que la API cumple los requerimientos básicos del sistema de reconocimiento de voz, ya que tiene una eficiencia alta con respecto a los comandos a utilizar para activar los sistemas del automóvil.

## 3.2 MÓDULO BLUETOOTH

Como se trató previamente en el punto 2.3, el módulo bluetooth HC-05 funcionará en modo esclavo para establecer comunicación entre el sistema de reconocimiento de voz y el dispositivo móvil. Por lo tanto, las siguientes pruebas constaron en comprobar las características que ofreció el fabricante, así como el modo de comunicación, velocidad de comunicación, cambio de nombre al módulo y código de vinculación para establecer comunicación, mediante programación en la plataforma IDE.

Para nuestro sistema se necesitan los siguientes requerimientos:

- Modo de Conexión: **Esclavo**.
- Velocidad de comunicación: **9600 baudios**.
- Nombre: **ANUBIS**.
- Código de vinculación: **0506**.

La programación se realizó de la siguiente manera:

### 3.2.1 TEST DE FUNCIONAMIENTO

#### a) Estatus de comunicación

Lo primero es comprobar si nuestro bluetooth responde a los comandos AT.

Procedimiento:

**Enviar:** AT

**Recibe:** OK

Se recibió como respuesta un OK, entonces podemos continuar.

#### b) Configuración maestro-esclavo

Para el sistema se requirió el modo esclavo.

Procedimiento:

**Enviar:** AT+ROLE=<Role> Ej.: AT+ROLE=0

**Dónde:** <Role>

0 -> Esclavo

1 -> Maestro

### **c) Velocidad de comunicación**

Procedimiento:

**Enviar:** AT+UART=<Baud> ,< StopBit>,< Parity>

**Dónde:**

< Baud > equivale a una velocidad, los valores pueden ser: 4800, 9600, 19200, 38400, 57600, 115200, 23400, 460800, 921600 o 1382400.

< StopBit> es el Bit de parada, puede ser 0 o 1, para 1 bit o 2 bits de parada respectivamente. Para aplicaciones comunes se trabaja con 1 bit por lo que este parámetro normalmente se deja en 0.

< Parity> Es la paridad, puede ser 0 (Sin Paridad), 1 (Paridad impar) o 2 (Paridad par). Para aplicaciones comunes no se usa paridad, por lo que se recomienda dejar este parámetro en 0.

### **d) Cambio de nombre del módulo**

Procedimiento:

**Enviar:** AT+NAME=<Nombre> Ejemplo: AT+NAME=ANUBIS

### **e) Cambio el código de vinculación**

Procedimiento:

**Enviar:** AT+PSWD=<Pin> Ejemplo: AT+PSWD=0506

Teniendo entonces el módulo previamente programado y conectado a una fuente de 5V se trató de entablar comunicación un total de 50 veces a una distancia que oscile entre los 5m y 12m, ya que es la distancia a la que opera un módulo *clase 2* (ver Punto 1.8), mediante dos teléfonos inteligentes; uno de la marca *Xiaomi* modelo *Mi A1* y el segundo de la marca sur coreana *Samsung* modelo *J7 PRO*. Arrojan el análisis mostrado en la Tabla 3.3.

*Tabla 3.3 Eficiencia de conexión modulo bluetooth HC-05.*

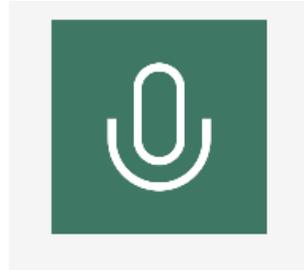
<b>Número de intentos</b>	<b>Conexiones exitosas</b>	<b>Eficiencia de conexión (%)</b>	<b>Tiempo promedio de conexión (segundos)</b>
100	95	95	5

Tomando en cuenta el análisis anterior se observó que el módulo bluetooth HC-05 responde de manera muy eficiente a distancias menores o iguales a 10 metros, lo que lo convierte en una opción viable para el diseño del sistema, ya que cubre de manera adecuada el rango de operación, y la conexión en un tiempo promedio de 5 segundos. Asimismo, fue posible cambiar el nombre del dispositivo y establecer conexión con el mismo mediante un código de vinculación previamente establecido en el código de programación.

### 3.3 APLICACIÓN MÓVIL

La aplicación móvil tendrá la función de comunicar al usuario con el sistema de reconocimiento de voz mediante el uso de un teléfono celular inteligente, con el cual el usuario dictará una palabra que se procesará en la API por medio de la aplicación, y que posteriormente la enviará al sistema para compararla con las instrucciones que previamente hayan sido programadas en el sistema.

La aplicación diseñada cuenta con logotipo como se muestra en la Figura 3.1.



*Figura 3.1 Logotipo de la aplicación de reconocimiento de voz.*

Además, posee una interfaz de fácil uso y para ingresar a ella es necesario contar con un usuario y contraseña, y evita también el mal uso del sistema, como se muestra en la Figura 3.2.



*Figura 3.2 Pantalla de acceso a la aplicación.*

Una vez que el usuario ingresa satisfactoriamente a la aplicación se mostrará la siguiente pantalla en la que resaltan 3 accesos con las leyendas: *HABILITAR BLUETOOTH*, *CONECTAR*, *DESCONECTAR*. Con los cuales la aplicación realiza la acción descrita al ser seleccionados por el usuario. (ver *Figura 3.3*).



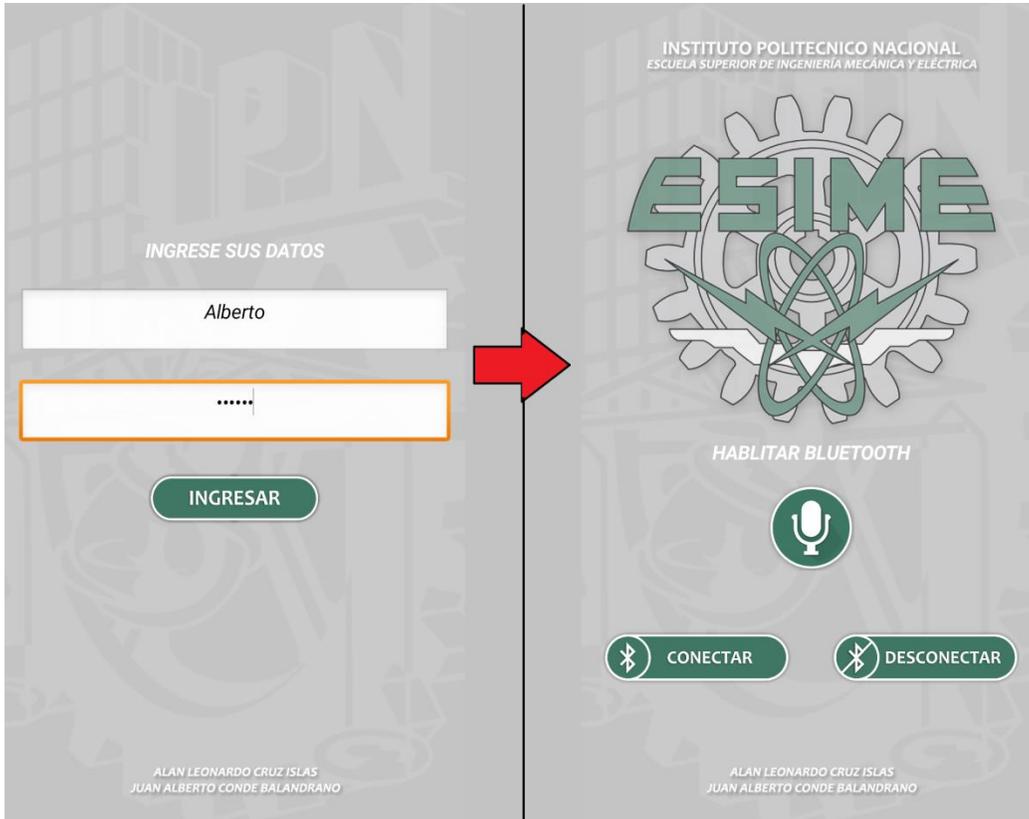
*Figura 3.3 Interfaz de inicio de la aplicación.*

A partir de lo descrito anteriormente se realizaron pruebas (en dos teléfonos inteligentes, uno de la marca *Xiaomi* modelo *Mi A1* y otro de la marca *Samsung J7 PRO*) a la aplicación para verificar su funcionamiento, así como descartar posibles fallas que pudieran ocasionar el mal funcionamiento del sistema.

- *PRIMERA PARTE*

La prueba constó en intentar acceder en 100 ocasiones con el usuario y contraseña válidos para la aplicación con el fin de corroborar su buen funcionamiento tal y como se muestra en la *Figura 3.4*. Teniendo en las 100 ocasiones éxito al tratar de ingresar a la aplicación.

- 1) USUARIO: **Alberto**  
CONTRASEÑA: **Anubis**



*Figura 3.4 Ingreso a la aplicación mediante usuario y contraseña válidos.*

Al constatar el acceso exitoso a la aplicación, se trató de ingresar nuevamente a la aplicación un total de 100 ocasiones con usuarios y contraseñas no válidos que pudieran tener parecido con el usuario y contraseña válidos. Resultados que se arrojaron en la Tabla 3.4.

*Tabla 3.4 Nivel de seguridad de la aplicación.*

<b>Intentos de ingreso con usuarios y contraseñas NO válidos</b>	<b>Ingresos exitosos</b>	<b>Seguridad (%)</b>
100	0	100

De tal manera que, al realizar las pruebas anteriores, se comprobó que la aplicación móvil tiene un alto nivel de seguridad y confianza para acceder a la misma, evitando el uso del sistema por personas no autorizadas por parte del usuario.

- SEGUNDA PARTE

La prueba constó en establecer conexión un total de 50 veces entre la aplicación y el dispositivo bluetooth. Para ello es necesario seguir las instrucciones que se muestran en el *Anexo A*.

Las pruebas arrojadas se muestran en la Tabla 3.5.

*Tabla 3.5 Efectividad de conexión entre la aplicación móvil y el dispositivo bluetooth.*

<b><i>Total de intentos</i></b>	<b><i>Intentos exitosos</i></b>	<b><i>Intentos sin éxito</i></b>	<b><i>Efectividad (%)</i></b>
50	47	3	94

Tenemos que existe una alta efectividad de conexión entre la aplicación móvil y el dispositivo bluetooth del sistema, a excepción de cuando el usuario olvida desvincular su dispositivo móvil, lo cual crea un error al intentar vincular con otro dispositivo nuevo, es por eso que se hacen ciertas recomendaciones de uso que se muestran en el *Anexo A*.

### 3.4 PROTOTIPO ANUBIS 1.0

El prototipo *Anubis 1.0* se construyó con la finalidad de probar y observar el funcionamiento en conjunto de los bloques anteriores al bloque de potencia, para en su caso descubrir fallas y errores de diseño que no se suscitaron en la prueba individual por bloque, y que pudieran afectar el funcionamiento del sistema.

*Anubis 1.0* fue montado en una placa de pruebas con los bloques de reconocimiento de voz, bluetooth, y microcontrolador ATmega328. Para propósitos de prueba se colocaron diodos emisores de luz (LED) en las salidas digitales del microcontrolador, que tendrán la función de activar o desactivar los sistemas manuales del automóvil mediante la etapa de potencia.

*Anubis 1.0* fue montado conforme al diagrama de la Figura 2.8, constando del microcontrolador ATmega328, el módulo bluetooth, diodos LED, conectores y un oscilador externo de 16 MHz. (ver Figura 3.5).

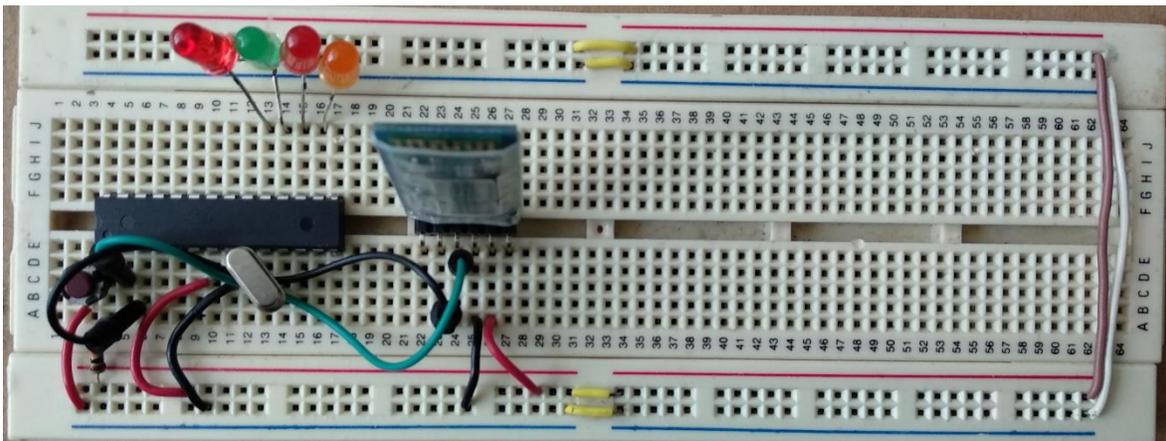


Figura 3.5 Vista superior de *Anubis 1.0*.

Las pruebas que se realizaron para comprobar el buen funcionamiento de *Anubis 1.0* fueron únicamente pronunciar las palabras clave y observar que responde a ellas mediante la iluminación de los diodos emisores de luz. Una vez que así fue comprobado se tomó la decisión de pasar el diseño a circuito impreso para poder realizar pruebas de funcionamiento más a fondo.

### 3.5 PROTOTIPO ANUBIS 2.0

Anubis 2.0 es el prototipo final del proyecto, el cual fue elaborado en diferentes etapas, todas ellas descritas a continuación:

**a) Diseño de placa de circuito impreso (PCB-Printed Circuit Board).**

El diseño de la placa consistió en realizar el acomodo de todos los componentes que tendrá el dispositivo, así como marcar las pistas que conducirán la corriente eléctrica. Para esta parte se utilizó el software Eagle 7.6.0 que se especializa en el diseño de circuitos impresos. (Figura 3.6).

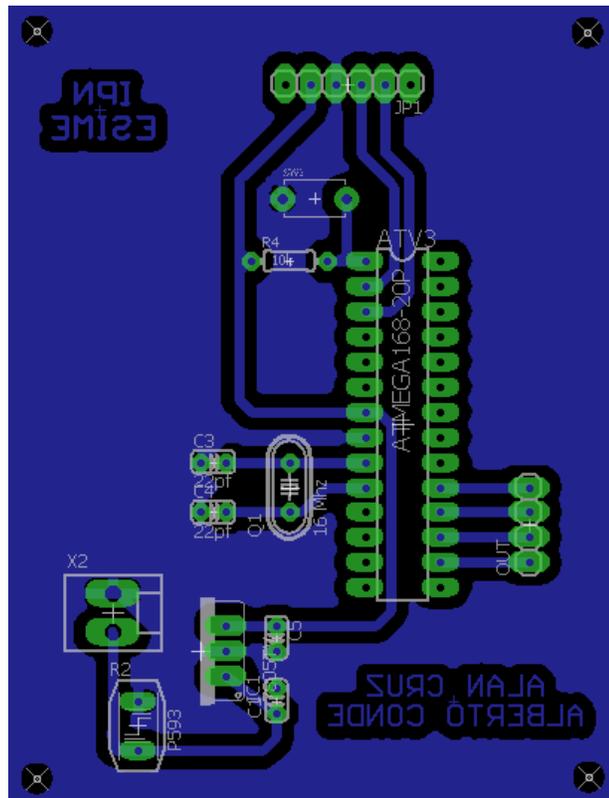
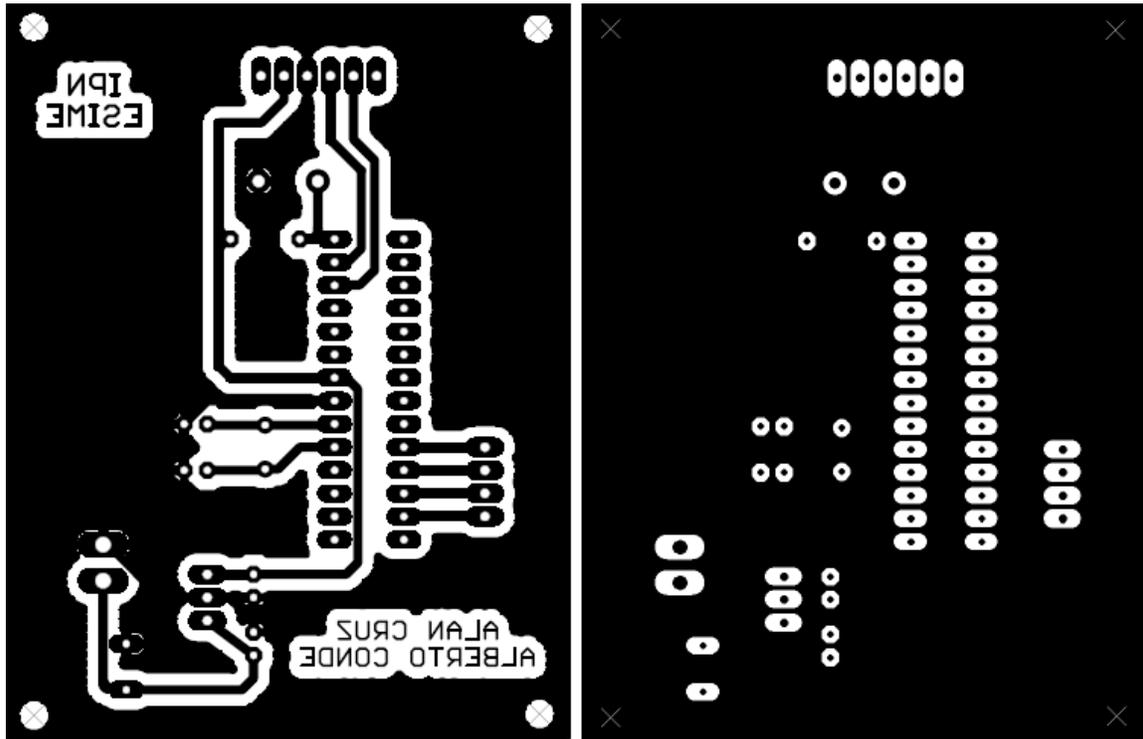


Figura 3.6 Diseño de Anubis 2.0 en Eagle.

**b) Impresión y grabado en placa de fibra de vidrio.**

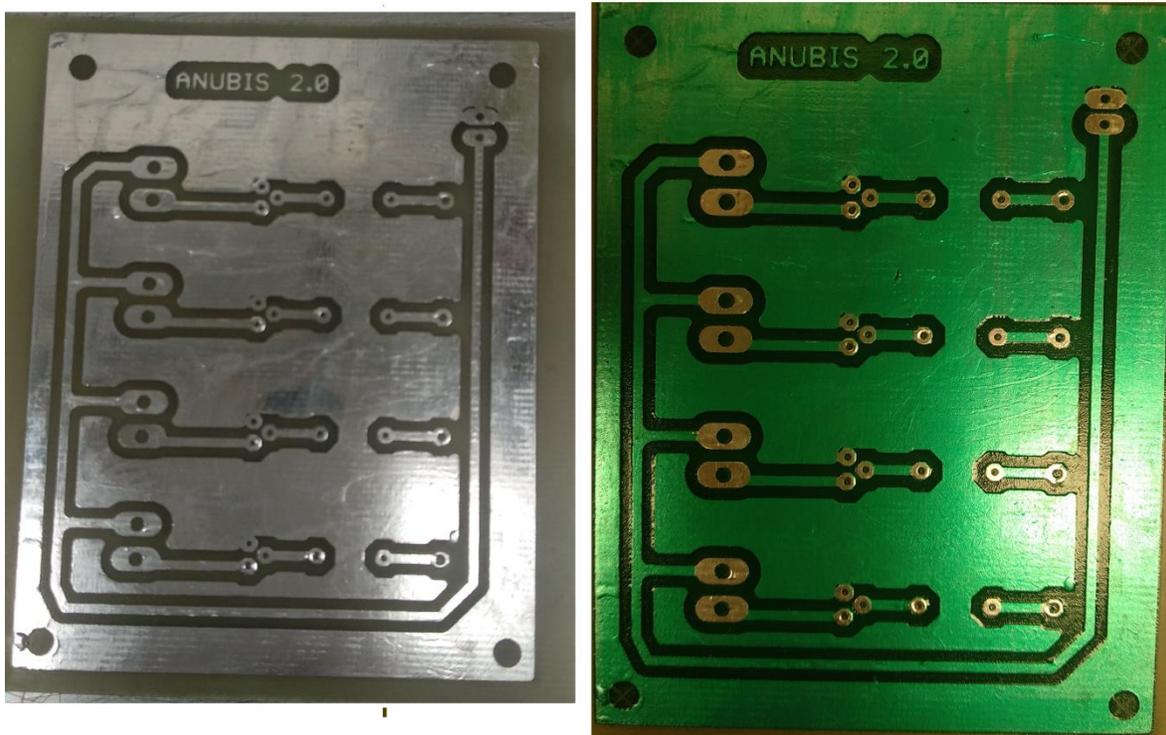
La técnica utilizada para la impresión del circuito fue serigrafía, la cual consiste en plasmar en un esténcil de 120 hilos el positivo del circuito y exponerlo a luz ultravioleta para su secado. Después se realiza la impresión en la placa de cobre con tinta industrialac, que es una tinta especial que soporta de buena manera el cloruro férrico. (Figura 3.7).



*Figura 3.7 Vista inferior del esténcil de 120 hilos,*

**c) Estañado, perforación de placa y aplicación de máscara antisoldante.**

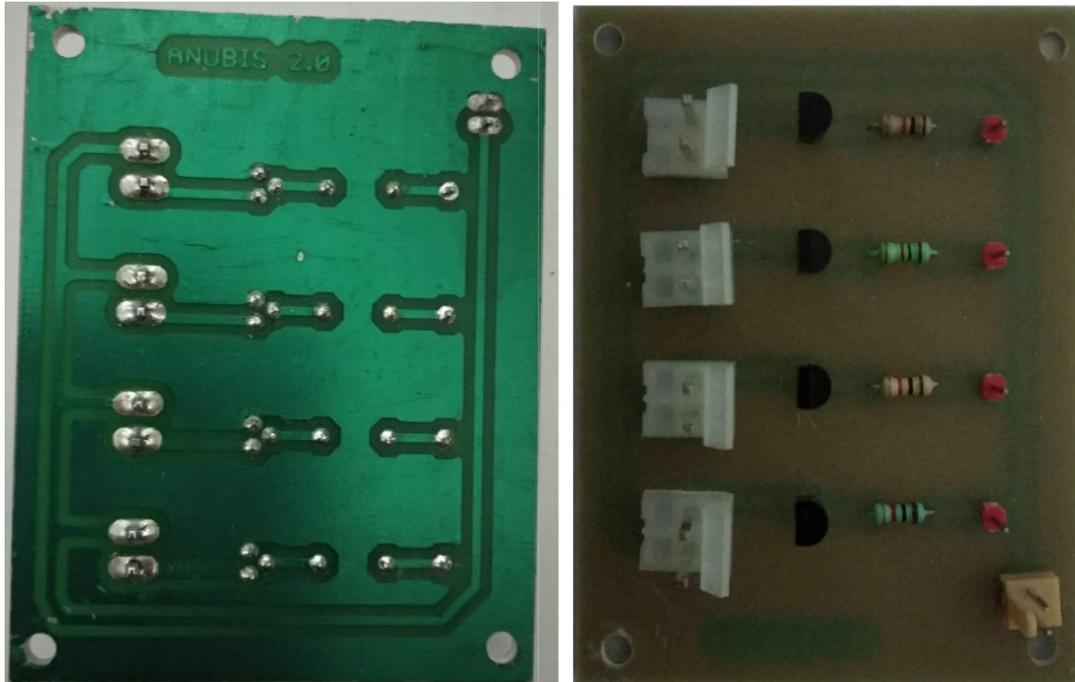
Para que la placa de cobre resista la intemperie y los estragos del clima a los cuales pueda ser sometida, es necesario aplicar una capa de estaño y posteriormente una capa de máscara antisoldante que sirve para dar más protección a la placa y a su vez impedir que se impregne estaño en lugares dónde no haya pads. (Figura 3.8).



*Figura 3.8 Vista inferior del estañado y de la máscara antisoldante.*

**d) Soldadura de componentes.**

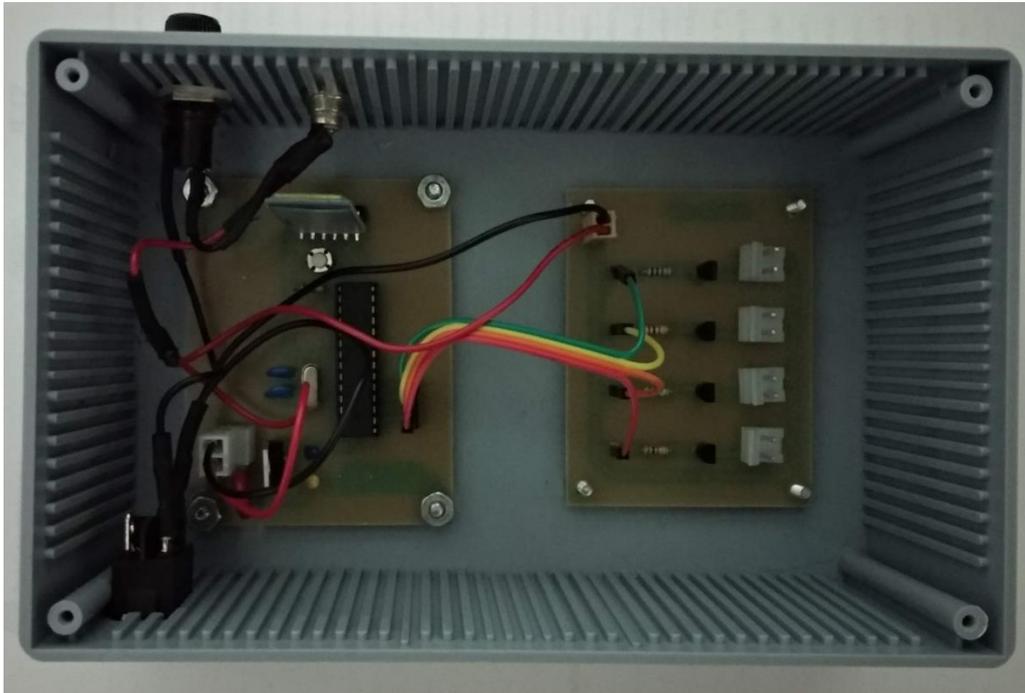
La soldadura de componentes se realiza con estaño en caliente para lograr una perfecta sujeción del componente con el pad y evitar algún tipo de movimiento que pueda impedir el correcto funcionamiento del sistema. (Figura 3.9).



*Figura 3.9 Vista inferior de la soldadura y superior del acomodo de componentes de la etapa de potencia.*

**e) Armado**

Para el armado del dispositivo se utilizaron conectores tipo molex paso 100 y paso 156, así como cable y headers macho y hembra. (Figura 3.10).



*Figura 3.10 Vista interior del ensamblado de Anubis 2.0.*

**CONSUMO ELÉCTRICO**

El prototipo Anubis 2.0 tiene un consumo de corriente en estado activo de 800 mA, por lo tanto la potencia a la que opera es de 4 W. A su vez en estado de espera consume 50 mA por lo que su potencia es de 250 mW. Por ende cumple y mejora los criterios de diseño establecidos para el presente proyecto.

## PRUEBAS DE FUNCIONAMIENTO

Las pruebas realizadas para comprobar el buen funcionamiento de *Anubis 2.0* fueron el dictar al sistema mediante el uso de dos teléfonos inteligentes (uno de la marca *Samsung* modelo *J7 PRO* y el segundo de la marca *Xiaomi* modelo *Mi A1*), las palabras clave previamente programadas para activar los sistemas manuales del automóvil en 10 ocasiones cada uno con 60 dB de ruido de fondo. Siguiendo las instrucciones que se muestran en el *Anexo A* y siendo las palabras clave:

- **abrir**
- **cerrar**
- **encender**
- **parar**
- **luz**
- **ventana**

Arrojando los siguientes resultados para el teléfono *Samsung* modelo *J7 PRO* en la Tabla 3.6.

*Tabla 3.6 Efectividad del prototipo Anubis 2.0 con Samsung J7 PRO.*

<b>Palabra</b>	<b>Total pronunciamientos</b>	<b>Pronunciamientos exitosos</b>	<b>Efectividad (%)</b>
<i>abrir</i>	10	9	90
<i>cerrar</i>	10	9	90
<i>encender</i>	10	9	90
<i>apagar</i>	10	9	90
<i>luz</i>	10	8	80
<i>ventana</i>	10	10	100

Y obteniendo la Tabla 3.7 para el teléfono *Xiaomi* modelo *Mi A1*.

Tabla 3.7 Efectividad del prototipo *Anubis 2.0* con *Xiaomi* modelo *Mi A1*.

<b>Palabra</b>	<b>Total pronunciamentos</b>	<b>Pronunciamentos exitosos</b>	<b>Efectividad (%)</b>
<i>abrir</i>	10	5	50
<i>cerrar</i>	10	5	50
<i>encender</i>	10	4	40
<i>apagar</i>	10	7	70
<i>luz</i>	10	3	30
<i>ventana</i>	10	5	50

Como se pudo constatar el prototipo *Anubis 2.0* funciona de manera muy eficiente, incluso en ambientes con 70 dB de ruido de fondo con el modelo *Samsung*, a diferencia del *Xiaomi* que presenta poca efectividad en cuanto a funcionamiento. Después de analizar las anomalías por parte de *Xiaomi*, se encontró que el sistema operativo del mismo es la causa de que no se estable una comunicación eficiente entre el teléfono y el sistema. Dando los mejores resultados en *Samsung* por tener otra versión del mismo sistema operativo.

## **COSTOS.**

Para la realización del presente trabajo se previeron varios aspectos, uno de ellos es que no se toma en cuenta el precio del teléfono inteligente, ya que en nuestra sociedad la mayoría de la gente dispone de uno para realizar sus actividades diarias, lo que lo hace necesario en sus vidas.

Todos y cada uno de los costos se presentan en la Tabla 3.8.

*Tabla 3.8 Costos para el desarrollo del proyecto.*

<b>Descripción</b>	<b>Cantidad</b>	<b>Precio</b>	<b>Total</b>
Módulo bluetooth HC-05	1	\$150,00	\$150,00
Placa de pruebas	1	\$83,00	\$83,00
Microcontrolador ATmega328	1	\$70,00	\$70,00
Programador	1	\$150,00	\$150,00
Oscilador 16 MHz	1	\$5,00	\$5,00
Diodos LED	5	\$1,50	\$7,50
Conectores*	1	\$50,00	\$50,00
Regulador de Voltaje LM7905	1	\$10,00	\$10,00
Capacitores de tantalio	2	\$2,00	\$4,00
Capacitores cerámicos	2	\$1,00	\$2,00
Transistores BC547A	2	\$3,00	\$6,00
Transistor TIC31	1	\$15,00	\$15,00
Botón NA	1	\$1,00	\$1,00
Resistores	5	\$1,00	\$5,00
Placa para circuito impreso	1	\$120,00	\$120,00
Estaño**	1	\$10,00	\$10,00
Máscara Antisoldante**	1	\$90,00	\$90,00
Gabinete	1	\$97,00	\$97,00
Molex	10	\$4,00	\$40,00
Maquila	NA	\$250,00	\$250,00
Depreciación de herramienta	NA	\$173,50	\$173,50
Instalación	NA	\$800,00	\$800,00
Ingeniería	NA	\$17.000,00	\$17.000,00
*Paquete, **Tubo.		<b>Gran Total</b>	<b>\$19.139,00</b>

## **CONCLUSIONES**

El reconocimiento de voz es una técnica poco explorada en nuestro país, esto dificulta incursionar a fondo en el tema, ya que la literatura escasea y mucha parte de la que existe se encuentra en otros idiomas, por ende los proyectos que pretenden hacer uso de las diferentes técnicas de reconocimiento de voz presentan ambigüedades e inconvenientes al momento de la investigación para lograr obtener una interpretación aceptable del mensaje recibido y por consecuencia hacen confuso el proceso de realización del mismo.

Debido a la poca información sobre esta tecnología en libros, se tuvo que basar sobre algunos documentos electrónicos, artículos y trabajos de grado, esto coadyuvó a indagar más sobre el tema y desarrollar la investigación.

El presente trabajo da muestra de los diferentes procesos que son necesarios para realizar un sistema de reconocimiento de voz, y a su vez constatar los problemas que presentó a lo largo de su realización, como fue encontrar la manera de procesar la voz en un dispositivo móvil que fuese de fácil acceso, es por eso que se optó por emplear la API de Google que es de uso libre y lo podemos encontrar en los teléfonos y tabletas electrónicas, que principalmente fueron creados para facilitar la comunicación entre personas con capacidades diferentes. La API permitió convertir la señal de voz a texto de manera fácil y sencilla incluso en ambientes con niveles de ruido moderado, permitiendo simular el entorno y las condiciones que se pueden presentar en el interior de un automóvil que circula por las calles de la Ciudad de México. El dispositivo, incluso es capaz de funcionar hasta en ambientes con 70 dB de ruido de fondo.

Asimismo, se corroboró que la transmisión bluetooth es una opción eficaz, segura y económica de transmitir los datos provenientes del dispositivo móvil hacia el microcontrolador por un medio inalámbrico, que facilita la comunicación entre dispositivos a distancia sin que se presenten pérdidas considerables, y así obtener un mejor desempeño en el sistema. De esta manera el usuario puede activar los sistemas manuales del automóvil como encender el motor, apagar el motor, activar o desactivar las luces internas y subir o bajar las ventanillas de una forma sencilla por medio del uso de comandos de voz.

En la parte de seguridad, la vinculación de los dispositivos, la aplicación móvil y los comandos de voz resultaron ser un gran apoyo para la terminación del proyecto, ya que contribuyó a proporcionar una alta tasa de confianza y seguridad al acceder al mismo, haciéndolo casi infalible en caso de que terceras personas no autorizadas intenten ingresar al sistema aun cuando el usuario haya extraviado o

le hayan hurtado su dispositivo móvil.

El sistema puede contar con mejoras, tanto en la instalación en el automóvil como mejoras técnicas, agregar más comandos de voz, mejorar el diseño, así como sus aplicaciones. La tecnología sigue avanzando y puede que la información empleada no sea suficiente en algunos años, por lo cual se recomienda indagar más en el tema con el fin de aportar y complementar más sobre esta técnica.

El reconocimiento de voz puede tener varias y diferentes aplicaciones ya que es de gran conveniencia al ahorrarnos tiempo y esfuerzo en la realización de las tareas que realizamos a diario.

## **ANEXO A.**

### **INSTRUCCIONES Y CONSEJOS.**

#### **ANTES DEL USO.**

- 2) *Encienda el bluetooth de su dispositivo móvil.*
- 3) *Vincule el módulo bluetooth con su dispositivo móvil ingresando el código de vinculación correspondiente.*
- 4) *Descargue el reconocimiento de voz sin conexión en el idioma que se utilizará.*

#### **DESPUÉS DEL USO.**

*En su teléfono o Tablet Android, abra la aplicación “Anubis”.*

- 1) *Ingrese el usuario y contraseña correspondiente para poder ingresar a la aplicación.*
- 2) *Presione el botón de “conectar” para poder hacer una conexión entre el módulo de bluetooth y el dispositivo móvil.*
- 3) *Presione el botón de “reconocimiento voz” y emita las palabras correspondientes que ejecutarán una orden en el automóvil. Hable despacio y con claridad. Deje espacio entre sus palabras y ponga énfasis en una dicción correcta para obtener mejores resultados.*
- 4) *Si ha terminado de utilizar el sistema de reconocimiento de voz, presione el botón “desconectar” para deshacer la conexión entre el módulo bluetooth y el dispositivo móvil y así evitar problemas en las conexiones futuras.*

#### **SOLUCIÓN DE PROBLEMAS.**

*Las causas más comunes de errores son:*

- 1) *Un problema de hardware; en el micrófono del dispositivo móvil.*
- 2) *Un dispositivo móvil con versiones muy antiguas de Android.*
- 3) *Problemas de entendimiento al ingresar los comandos de voz.*
- 4) *Olvidar deshacer la conexión del bluetooth después de haber usado la aplicación.*

*Antes de comenzar a usar la aplicación revisar el funcionamiento del micrófono y el reconocimiento de voz de Google mediante la aplicación de búsqueda por voz, emitiendo palabras al azar. Hablar claro para que la aplicación entienda los comandos de voz.*

## **ANEXO B.**

### **CÓDIGO EN C DEL PROGRAMA.**

```
String comandos;
```

```
int ledluz = 11;
```

```
int ledventana = 10;
```

```
void setup() {
```

```
    Serial.begin(9600);
```

```
    pinMode(13,OUTPUT);
```

```
    pinMode(12,OUTPUT);
```

```
    pinMode(11,OUTPUT);
```

```
    pinMode(10,OUTPUT);
```

```
    digitalWrite(ledluz, LOW);
```

```
    digitalWrite(ledventana, LOW);
```

```
}
```

```
void loop() {
```

```
    while(Serial.available()){
```

```
        delay(10);
```

```
        char a = Serial.read();
```

```
        comandos += a;
```

```
    }
```

```
    if(comandos.length(>0){
```

```
if(comandos=="*abrir")(digitalWrite(13,HIGH));

else if (comandos=="*cerrar")(digitalWrite(13,LOW));

if(comandos=="*encender")(digitalWrite(12,HIGH));

else if (comandos=="*parar")(digitalWrite(12,LOW));

if(comandos=="*luz")(digitalWrite(ledluz, !digitalRead(ledluz)));

if(comandos=="*ventanas")(digitalWrite(ledventana, !digitalRead(ledventana)));

}

comandos="";

}
```

## **REFERENCIAS.**

Varona Fernández, Amparo (1999) Historical evolution in the current Automatic Speech Recognition System , CICYT.

Bermúdez, Jesús Bernal, Bobadilla Sancho Jesús, Gómez Villa, Pedro (2000) Reconocimiento de voz y fonética acústica, Alfa omega, México.

Richard K. Miller, Terri C. Walker (1990) Natural Language and voice processing, The fairmont press, Inc, U.S.A.

Ahuactzin Larios (1999) Sistemas de reconocimiento y síntesis de voz, Tesis de licenciatura, Universidad de las Américas-Puebla, México.

José Francisco Giménez Albacete (2015) Seguridad de equipos informáticos, IC Editorial,

John G. Proakis, Dimitris G. Manolakis, (2007) Tratamiento digital de señales. Cuarta edición. Vuelapluma

A. Bruce Carlson, Paul B. Crilly, Janet C. Rutledge (2007) Sistemas de comunicación, una introducción a las señales y el ruido den las comunicaciones eléctricas. Cuarta edición.

Paul N. Cheremisinoff (1993) Industrial noise control. PTR Prentice Hall.

Etiquetado del corpus [Web en línea] Disponible desde internet en:  
[http://catarina.udlap.mx/u\\_dl\\_a/tales/documentos/lis/macias\\_k\\_r/capitulo4.pdf](http://catarina.udlap.mx/u_dl_a/tales/documentos/lis/macias_k_r/capitulo4.pdf) [con acceso 09 de noviembre de 2017]

[http://ww1.microchip.com/downloads/en/DeviceDoc/Atmel-42735-8-bit-AVR-Microcontroller-ATmega328-328P\\_Datasheet.pdf](http://ww1.microchip.com/downloads/en/DeviceDoc/Atmel-42735-8-bit-AVR-Microcontroller-ATmega328-328P_Datasheet.pdf) [con acceso 18 de febrero de 2018]

[https://cdn.makezine.com/uploads/2014/03/hc\\_hc-05-user-instructions-bluetooth.pdf](https://cdn.makezine.com/uploads/2014/03/hc_hc-05-user-instructions-bluetooth.pdf) [con acceso 11 de marzo de 2018]

<https://cloud.google.com/speech/?hl=es> [con acceso 25 de marzo de 2018]

<https://www.sparkfun.com/datasheets/Components/LM7805.pdf> [con acceso 15 de abril de 2018]