

RAE

1. **TIPO DE DOCUMENTO:** Trabajo de grado para optar por el título de INGENIERO DE SONIDO.
2. **TÍTULO:** DISPOSITIVO INALÁMBRICO CONTROLADOR DE SISTEMAS DE ENTRETENIMIENTO CASERO DE AUDIO Y VIDEO POR MEDIO DE COMANDOS DE VOZ.
3. **AUTOR:** Nicolás Reyes Cristancho.
4. **LUGAR:** Bogotá, D.C.
5. **FECHA:** Junio de 2014
6. **PALABRAS CLAVE:** EasyVR, Comandos dependientes de locutor, Comando independientes de locutor, dBFS, Mensajes Infrarrojos, Radiación Infrarroja, Reconocimiento de Voz, Relación Señal Ruido, Ruido de Fondo, Voz humana.
7. **DESCRIPCIÓN DEL TRABAJO:** El objetivo principal de este proyecto es Construir y Analizar el comportamiento de un dispositivo controlador de sistemas de entretenimiento casero de audio y video por medio de comandos de voz, teniendo en cuenta el reconocimiento de comandos dependientes e independientes de locutor a través del módulo EasyVR, proceso realizado bajo diferentes condiciones iniciales, permitiendo adquirir los niveles de presión sonora para los cuales el dispositivo desarrollado presenta un mayor porcentaje de efectividad.
8. **LÍNEAS DE INVESTIGACION:** Universidad de San Buenaventura/Tecnologías Actuales y Sociedad. Sub línea de Facultad de Ingeniería: Procesamiento de señales digitales y/o analógicas. Campo temático del programa: Diseño de sistema de sonido.
9. **METODOLOGÍA:** Es de carácter empírico-analítico, con un enfoque metodológico con base en el estudio, diseño y construcción de un sistema de reconocimiento de voz implementado en un ámbito casero.
10. **CONCLUSIONES:** Según las mediciones de efectividad realizadas se concluye que las fases de entrenamiento y reconocimiento de comandos de voz dependientes de locutor realizadas por el módulo EasyVR, no se ven afectadas de ninguna manera al variar el locutor, el idioma o el comando de voz, evidenciando los mismos resultados para las pruebas realizadas con los tres locutores involucrados en el estudio. Una vez realizado el análisis de los datos adquiridos para cada micrófono, en la tabla 34 se puede observar que los procesos involucrados en el reconocimiento de comandos voz se ven directamente afectados por la directividad y la sensibilidad del transductor utilizado, obteniendo los mejores resultados con un Shure SM58, micrófono de baja sensibilidad con patrón polar cardioide. Se evidenció que al programar el reconocimiento de comandos dependientes de locutor en el nivel 1 de dificultad, se obtienen resultados similares a un proceso de reconocimiento de comandos de voz independientes del locutor, permitiendo usar cualquier comando que el usuario desee en esta configuración. Se observa que los comandos de voz independientes de locutor predefinidos en el módulo EasyVR no mantienen los mismos niveles de efectividad, presenciando mejores resultados para los comandos “CERO”, “DOS”, “CUATRO” y “NUEVE”. Teniendo cuenta que se desconoce el algoritmo desarrollado por el fabricante, se reemplazó este proceso por el método que reconoce comandos dependientes de locutor en el nivel 1 de dificultad, el cual permite llevar a cabo un proceso satisfactorio de reconocimiento de comandos de voz independientes de locutor para cualquier palabra entrenada. El reconocimiento de comandos independientes por parte del módulo EasyVR, permite unos niveles de ruido de fondo mayores respecto a los comandos dependientes, sin que esto signifique un porcentaje de reconocimiento mayor. Para obtener una efectividad mayor al momento de controlar distintos sistemas de entretenimiento por medio de comandos de voz, se debe usar un micrófono cardioide de baja sensibilidad, ubicado a 15 centímetros de la boca del usuario. Es posible evidenciar que al momento de controlar simultáneamente los tres sistemas de entretenimiento casero a través del dispositivo desarrollado, se optimiza el proceso habitual de control sobre los mismos, permitiéndole al usuario manipular distintas de funciones a partir de la pronunciación de un solo comando de voz. Comparando los resultados obtenidos en la prueba estándar y en la prueba real, se observa que el proceso de reconocimiento de comandos de voz es más efectivo al tener un nivel de ruido de fondo constante. En la tablas 31 se observa claramente que el rango de niveles de presión sonora referentes al nivel de ruido de fondo presente al momento de realizar un proceso de reconocimiento de voz, se mejora o aumenta significativamente al reemplazar el micrófono HORN por los micrófonos Shure SM 58 y AT8015, notando que se puede realizar un proceso de reconocimiento de comandos de voz satisfactorio teniendo un nivel de presión sonora de ruido de fondo máximo de 80 dB que puede estar hasta 5dB por encima del nivel de presión sonora generado al pronunciar un comando de voz.

**DISPOSITIVO INALÁMBRICO CONTROLADOR DE SISTEMAS DE
ENTRETENIMIENTO CASERO DE AUDIO Y VIDEO POR MEDIO
DE COMANDOS DE VOZ**

NICOLÁS REYES CRISTANCHO

**PROYECTO DE GRADO PARA OPTAR AL TÍTULO DE
INGENIERO DE SONIDO.**

**ASESOR
ING. MIGUEL PÉREZ**

**UNIVERSIDAD DE SAN BUENAVENTURA, SEDE BOGOTÁ.
FACULTAD DE INGENIERÍA
INGENIERÍA DE SONIDO
BOGOTÁ
2014**

NOTA DE ACEPTACIÓN.

DIRECTOR

JURADO

JURADO

A mis padres que me apoyaron siempre.

AGRADECIMIENTOS.

El autor expresa sus agradecimientos a:

Miguel Pérez y Darío Páez, asesores durante todo el proceso de desarrollo del proyecto.

CONTENIDO.

	Pág.
INTRODUCCIÓN.....	20
1. DESCRIPCIÓN DEL PROYECTO.....	21
1.1. DESCRIPCIÓN DEL PROBLEMA.....	21
1.2. DELIMITACIÓN DEL PROYECTO.....	21
1.3. JUSTIFICACIÓN.....	22
1.4. OBJETIVOS.....	23
1.4.1. OBJETIVO GENERAL.....	23
1.4.2. OBJETIVOS ESPECÍFICOS.....	23
1.5. ANTECEDENTES.....	24
1.5.1. RECONOCIMIENTO DE VOZ.....	24
1.5.2. APLICACIONES.....	26
1.5.2.1. CONTROL REMOTO ACCENDA.....	27

1.5.2.2.	CONTROL REMOTO SURFBOARD.....	28
2.	MARCO TEÓRICO	29
2.1.	RADIACIÓN INFRARROJA	29
2.2.	SENSORES INFRARROJOS.....	30
2.3.	PROTOCOLO DE COMUNICACIÓN SONY SIRC PARA MANDOS INFRARROJOS.....	31
2.4.	MÓDULO DE RECONOCIMIENTO DE VOZ EASYVR.....	34
2.4.1.	CARACTERÍSTICAS.....	35
2.4.2.	MICRÓFONO.....	35
2.4.3.	EASYVR ARDUINO SHIELD.....	36
2.4.3.1.	DIAGRAMA DE PINES EASYVR ARDUINO SHIELD.....	37
2.4.3.2.	MODOS DE OPERACIÓN.....	39
2.4.3.3.	EASYVR COMMANDER.....	40
2.5.	VOZ HUMANA.....	41
2.6.	RUIDO.....	44

2.6.1.	RUIDO ROSA.....	44
2.7.	RELACIÓN SEÑAL RUIDO.....	45
3.	DESARROLLO INGENIERIL.....	46
3.1.	DIAGRAMA DE BLOQUES.....	46
3.2.	FUNCIONES A CONTROLAR.....	47
3.3.	ADQUISICIÓN DE MENSAJES INFRARROJOS.....	48
3.3.1.	CIRCUITO RECEPTOR DE MENSAJES INFRARROJOS.....	48
3.3.2.	MENSAJES INFRARROJOS DIGITALIZADOS.....	49
3.3.2.1.	CODIFICACIÓN DE MENSAJES INFRARROJOS.....	52
3.4.	CIRCUITO EMISOR DE MENSAJES INFRARROJOS.....	53
3.5.	COMANDOS DE VOZ.....	55
3.5.1.	COMANDOS INDEPENDIENTES DE LOCUTOR.....	55
3.5.2.	COMANDOS DEPENDIENTES DE LOCUTOR.....	56
3.6.	CIRCUITO DEL DISPOSITIVO CONTROLADOR.....	58
3.6.1.	ESQUEMÁTICO.....	59

3.7.	CÓDIGO DE PROGRAMACIÓN BASE.....	61
3.8.	CUBIERTA.....	63
3.9.	MEDICIONES DE EFECTIVIDAD.....	65
3.9.1.	EQUIPOS DE MEDICIÓN.....	65
3.9.1.1.	MONITORES ADAM A8X.....	66
3.9.1.2.	SONÓMETRO SVANTEK 943A.....	67
3.9.1.3.	MICRÓFONOS.....	68
3.9.2.	PRUEBA ESTÁNDAR.....	69
3.9.2.1.	MEDICIÓN DE RUIDO DE FONDO.....	70
3.9.2.2.	NORMALIZACIÓN DE NIVELES DE RUIDO DE FONDO.....	72
3.9.2.3.	NORMALIZACIÓN DE NIVELES DE COMANDOS DE VOZ.....	74
3.9.2.4.	MONTAJE.....	76
3.9.2.5.	ETAPA DE CALIBRACIÓN.....	77
3.9.2.6.	PRUEBA DE NIVELES DE PRESIÓN SONORA EN LA FASE DE ENTRENAMIENTO.....	78

3.9.2.7.	PRUEBA DE NIVELES DE DIFICULTAD PARA COMANDOS DEPENDIENTES.....	82
3.9.2.8.	PRUEBA DE NIVELES DE DIFICULTAD PARA COMANDOS INDEPENDIENTES.....	86
3.9.3.	PRUEBA REAL.....	89
3.9.3.1.	MEDICIÓN DE RUIDO DE FONDO.....	90
3.9.3.2.	MONTAJE.....	92
3.9.3.3.	MEDICION DE NIVELES DE PRESIÓN SONORA GENERADOS POR EL EQUIPO DE SONIDO.....	94
3.9.3.4.	NIVELES DE VOZ DE LOCUTORES.....	95
3.9.3.5.	PRUEBA DECOMANDOS DEPENDIENTES.....	96
4.	ANÁLISIS DE DATOS.....	100
4.1.	PRUEBA ESTÁNDAR.....	100
4.2.	PRUEBA REAL.....	103
5.	OPTIMIZACIÓN DEL DISPOSITIVO CONTROLADOR.....	105
6.	LIMITACIONES.....	111

7.	RECOMENDACIONES DE USO.....	114
8.	RECOMENDACIONES PARA TRABAJOS A FUTURO.....	115
9.	CONCLUSIONES.....	116
10.	BIBLIOGRAFÍA.....	117
ANEXO A. CÓDIGOS DE PROGRAMACIÓN.....		121
ANEXO B. RESULTADOS PRUEBA ESTÁNDAR.....		136
ANEXO C. RESULTADOS PRUEBA REAL.....		167
ANEXO D. ANÁLISIS DE LAS CARACTERÍSTICAS GENERALES DE LAS VOCES DE LOS LOCUTORES.....		177

LISTA DE TABLAS.

	Pág.
Tabla 1. Descripción pines y conectores EasyVR Arduino Shield.....	38
Tabla 2. Lista de fonemas del español.....	43
Tabla 3. Funciones a controlar.....	47
Tabla 4. Mensajes infrarrojos para el televisor.....	49
Tabla 5. Mensajes infrarrojos para el equipo de sonido.....	49
Tabla 6. Mensajes infrarrojos para el DVD.....	50
Tabla 7. Mensajes infrarrojos para el DVD.....	50
Tabla 8. Codificación lógica de mensajes infrarrojos.....	52
Tabla 9. Listado de conexiones.....	57
Tabla 10. Características principales del monitor ADAM A8X.....	66
Tabla 11. Características principales del sonómetro SVANTEK 943 A.....	67
Tabla 12. Características de los micrófonos usados en el estudio.....	68

Tabla 13.	Promedios aritméticos de medición de ruido de fondo prueba estándar.....	71
Tabla 14.	Niveles dBFS vs dB SPL.....	72
Tabla 15.	Nivel de presión sonora generado por comandos grabados.....	75
Tabla 16.	Características generales de la voz de los locutores al pronunciar el comando PRENDER.....	75
Tabla 17.	Niveles de presión sonora óptimos para la fase de entrenamiento....	79
Tabla 18.	Valores representativos de relación señal/ruido en fase de entrenamiento.....	80
Tabla 19.	Variación de niveles de presión sonora con nivel de relación señal ruido constante.....	80
Tabla 20.	Medición por niveles de dificultad con micrófono HORN.....	82
Tabla 21.	Medición por niveles de dificultad con micrófono SM58.....	83
Tabla 22.	Medición por niveles de dificultad con micrófono AT 8015.....	83
Tabla 23.	Variación de niveles de presión sonora distintos a la fase de entrenamiento.....	84
Tabla 24.	Comparación de efectividad para comandos independientes.....	86

Tabla 25.	Niveles de prueba para comandos independientes.....	87
Tabla 26.	Promedios aritméticos de medición de ruido de fondo prueba real...	90
Tabla 27.	Niveles de presión sonora generados por el equipo de sonido.....	94
Tabla 28.	Comprobación de niveles de dificultad para comandos dependientes.....	96
Tabla 29.	Datos de efectividad por nivel de ganancia del equipo de sonido con nivel 1 de dificultad.....	97
Tabla 30.	Datos de efectividad por nivel de ganancia del equipo de sonido con nivel de dificultad 5.....	98
Tabla 31.	Datos de efectividad en prueba a 15 centímetros.....	99
Tabla 32.	Funciones controladas por el dispositivo desarrollado.....	106
Tabla 33.	Limitaciones de funcionamiento.....	111

LISTA DE FIGURAS.

	Pág.
Figura 1. Control remoto Accenda.....	27
Figura 2. Control remoto Surfboard.....	28
Figura 3. Temperatura y longitud de ondas de luz.....	29
Figura 4. Representación lógica del protocolo SIRC.....	31
Figura 5. Mensaje infrarrojo en el protocolo SIRC.....	32
Figura 6. Series de bits en el protocolo SIRC.....	33
Figura 7. Repetición de envío de mensajes en el protocolo SIRC.....	33
Figura 8. Módulo EasyVR.....	34
Figura 9. Acople entre EasyVR Arduino Shield y Arduino UNO.....	36
Figura 10. Diagrama de pines EasyVR Arduino shield.....	37
Figura 11. Sección J12 EasyVR Arduino Shield.....	39
Figura 12. Pantalla inicial software EasyVR Commander.....	40

Figura 13.	Tracto vocal.....	41
Figura 14.	Espectro ruido rosa.....	44
Figura 15.	Relación señal/ruido.....	45
Figura 16.	Diagrama de bloques.....	46
Figura 17.	Circuito receptor de mensajes infrarrojos.....	48
Figura 18.	Circuito transmisor de mensajes infrarrojos.....	53
Figura 19.	Diagrama de flujo código 1.....	54
Figura 20.	Comandos independientes de locutor Wordset 3.....	55
Figura 21.	Entrenamiento de comandos dependientes de locutor.....	56
Figura 22.	Comandos dependientes.....	57
Figura 23.	Esquemático.....	59
Figura 24.	Conexión EasyVR – Arduino UNO.....	60
Figura 25.	Diagrama de flujo código 2 y código 3.....	61
Figura 26.	Diseño de la cubierta.....	63
Figura 27.	Cubierta y terminales.....	64

Figura 28.	Live Room estudio de grabación 5.1.....	69
Figura 29.	Puntos de medición de ruido de fondo para la prueba estándar.....	70
Figura 30.	Montaje prueba estándar.....	76
Figura 31.	Recinto prueba real (sala de estar).....	89
Figura 32.	Puntos de medición de ruido de fondo para la prueba real.....	90
Figura 33.	Montaje prueba real.....	92
Figura 34.	Diagrama de flujo Código final.....	107
Figura 35.	Esquemático final.....	108
Figura 36.	Dispositivo controlador.....	109
Figura 37.	Parte interna del dispositivo controlador.....	110

RESUMEN.

En este documento se presenta el desarrollo y el estudio de un dispositivo inalámbrico controlador por medio de comandos de voz, de sistemas de entretenimiento casero de audio y video fabricados por la marca SONY, desarrollado con el fin de obtener los niveles de presión sonora óptimos para el correcto funcionamiento del módulo de reconocimiento de voz EasyVR, el cual fue sometido a diferentes estudios de efectividad realizados bajo diferentes condiciones iniciales.

En el capítulo 1 se describe detalladamente el proyecto junto con sus alcances y objetivos, en el capítulo 2 se presenta la información teórica necesaria para llevar a cabo su cumplimiento, exponiendo temas referentes a la radiación infrarroja y al reconocimiento de voz. Posteriormente en el capítulo 3 se muestra el desarrollo ingenieril, haciendo referencia al uso de las diferentes herramientas empleadas a lo largo del desarrollo del proyecto; adicionalmente se explica detalladamente la metodología de las mediciones de efectividad realizadas donde se emplearon tres micrófonos, tres locutores y distintos niveles de relación señal/ruido, cuyo análisis posterior permitió concluir las condiciones para las cuales el dispositivo desarrollado presenta un mayor porcentaje de efectividad, información presente en el capítulo 4. Seguidamente se llevó a cabo una optimización del funcionamiento del dispositivo controlador haciendo referencia al código de programación implementado y a la cubierta de protección, dando paso a la etapa final del proyecto en donde se realizaron ajustes que permitieron obtener un producto que optimiza el proceso de control simultáneo de varios sistemas de entretenimiento casero.

INTRODUCCIÓN.

Durante del siglo XX se empezaron a desarrollar diferentes tecnologías que tenían como objetivo optimizar la interacción entre humanos y máquinas, época durante la cual se crearon herramientas que mejoraron significativamente dicha comunicación, como el reconocimiento de la voz humana y los protocolos de comunicación inalámbrica a través de rayos infrarrojos, tecnologías que se unificaron para el desarrollo de este proyecto.

Con el fin de implementar las herramientas más prácticas, se llevó a cabo una investigación previa que sirvió como fundamento para la implementación de un Arduino UNO y un módulo de reconocimiento de voz EasyVR, dispositivos que permiten desarrollar un sistema de emisión de rayos controlado por comandos de voz.

Una vez desarrollado el dispositivo controlador se realizó un estudio con el fin de mejorar la experiencia de las personas al controlar varios sistemas de entretenimiento caseros de audio y video de la marca SONY, que para diferentes condiciones iniciales permitiese optimizar el control sobre dichos dispositivos.

Dicho estudio permitió exponer las ventajas y desventajas de la implementación de comandos de voz para la realización de cualquier actividad cotidiana, laboral o industrial, dando a conocer las condiciones óptimas de funcionamiento del módulo de reconocimiento de voz EasyVR, referentes a los niveles de presión sonora para los cuales se puede llevar a cabo un proceso de reconocimiento de voz satisfactorio.

1. DESCRIPCIÓN DEL PROYECTO.

1.1. DESCRIPCIÓN DEL PROBLEMA.

El gran número de marcas y modelos de sistemas de entretenimiento caseros de audio y video presentes en el mercado actual, ha dado paso a que en los hogares se encuentren dos o más mandos a distancia, generando que los usuarios de estos dispositivos tengan que hacer uso de un control remoto distinto para cada sistema de entretenimiento. Teniendo en cuenta que este proceso de control es engorroso y poco práctico al momento de usar varios sistemas simultáneamente, se genera la siguiente pregunta:

¿Es posible mejorar la experiencia del usuario al controlar varios sistemas caseros de entretenimiento de audio y video del mismo fabricante, empleando un único dispositivo controlador inalámbrico que implemente reconocimiento de comandos de voz y un sistema de emisión de rayos infrarrojos?

1.2. DELIMITACIÓN DEL PROYECTO.

Se plantea desarrollar un dispositivo único de control inalámbrico controlado por comandos de voz, dependientes e independientes del locutor, que permita controlar hasta diez funciones de tres dispositivos de entretenimiento de audio y video de la marca SONY. Una vez desarrollado el producto, se pretende llevar a cabo un estudio de efectividad del módulo de reconocimiento de voz EasyVR que permita evidenciar las condiciones óptimas de su funcionamiento.

1.3. JUSTIFICACIÓN.

Cada fabricante de sistemas de entretenimiento casero implementa un protocolo de comunicación infrarrojo distinto para su respectiva variedad de productos, además de usar distintos mensajes y códigos para distintos sistemas de la misma marca (DVD, televisor, equipo de sonido, etc.) obligando a los usuarios a hacer uso de un control remoto distinto para controlar cada dispositivo por separado.

Con la implementación de un sistema de reconocimiento de comandos de voz y un sistema de emisión de mensajes vía infrarrojo capaz de enviar la información necesaria para controlar varios sistemas de la misma marca simultáneamente con un solo dispositivo controlador, se pretende mejorar y optimizar significativamente el control, la interacción y la experiencia de los usuarios de este tipo de sistemas.

Una vez desarrollado el dispositivo y realizado el estudio de las condiciones óptimas de funcionamiento referentes a los niveles de presión sonora permitidos al momento de llevar a cabo el proceso de reconocimiento de comandos voz, se pueden llegar a concluir los campos de acción para los cuales será más apto, práctico y efectivo, aportando significativamente a la optimización de procesos de control en diferentes espacios, permitiendo a cualquier persona controlar diversos dispositivos de forma inalámbrica por medio su voz.

1.4. OBJETIVOS.

1.4.1. OBJETIVO GENERAL:

Construir un dispositivo electrónico inalámbrico capaz de controlar hasta 10 funciones de tres sistemas de entretenimiento caseros de audio y video de la marca SONY por medio de comandos de voz, con el fin de optimizar, innovar y mejorar la experiencia del usuario.

1.4.2. OBJETIVOS ESPECÍFICOS:

- Implementar un sistema de emisión de rayos infrarrojos que permita emplear el protocolo de comunicación respectivo.
- Utilizar un módulo de reconocimiento de voz acoplado al sistema de emisión de rayos infrarrojos que permita ejercer control sobre los sistemas de entretenimiento.
- Comparar la efectividad del dispositivo controlador a desarrollar, empleando tres micrófonos diferentes con tres locutores distintos.
- Identificar las condiciones óptimas de funcionamiento del módulo de reconocimiento de voz con el fin de optimizar el porcentaje de efectividad para comandos dependientes e independientes de locutor.

1.5. ANTECEDENTES.

1.5.1. RECONOCIMIENTO DE VOZ.

Es un proceso que permite identificar y procesar la voz humana, reconociendo la información allí contenida para convertirla en una salida útil que puede llegar a actuar sobre un proceso. Su objetivo principal es permitir la comunicación hombre-máquina a través de la voz, en donde se realizan distintos análisis acústicos en presencia de variables dependientes e independientes de cada persona.

Entre los primeros intentos por transcribir secuencias de sonidos de la voz humana, alrededor de 1880, Tihamir Nemes fue tachado de irrealista al intentar gestionar una patente para crear un sistema que reconociera sonidos de la voz humana y los pudiese imprimir como texto. Fue allí donde se empezó a vislumbrar la complejidad que podría conllevar un proceso de esta índole, en donde se tendrían que tener en cuenta diversas variables propias de la voz humana que no se habían estudiado anteriormente.

Treinta años después, se dio el primer gran paso hacia el reconocimiento de voz automático, cuando la compañía AT&T Bell Laboratories presentó la primera máquina capaz de reconocer la voz humana, logrando identificar y transcribir la pronunciación de los 10 dígitos con un 99% de efectividad, con lo cual se creyó que el reconocimiento automático de voz podría ser un proceso simple que permitiría identificar las palabras pronunciadas por cualquier persona en cualquier idioma con gran facilidad.

A partir de los años 60's, la mayoría de los investigadores interesados en el tema admitieron que el reconocimiento de voz era un proceso mucho más intrincado y sutil de lo que habían pensado anteriormente, por lo que empezaron a reducir los alcances de sus proyectos y se enfocaron en diseñar sistemas más específicos que dependieran directamente de un locutor, que reconocieran un máximo de 50 palabras y que registraran las pausas o espacios entre palabras pronunciadas [1].

Con el paso del tiempo se fueron perfeccionando este tipo de sistemas, hasta que a principios de los 70's se creó el VIP100 de Threshold Technology Inc, el primer sistema de reconocimiento de voz que utilizaba un vocabulario pequeño que dependía del locutor, dando paso al surgimiento del interés de las entidades de seguridad, incentivando la investigación y el mejoramiento de esta tecnología.

Una vez se adquirió la atención y el interés adecuado, se incluyeron otros tipos de análisis para el reconocimiento automático de la voz (léxico, sintáctico, semántico y pragmático), los cuales permitieron que en los 90's surgieran los sistemas de vocabulario amplio (más de 1000 palabras), haciendo que empresas reconocidas como Philips, Sensory Circuits, Dragon Systems, Speechworks, Apple, Microsoft y muchas más se interesaran por esta tecnología y la implementaran posteriormente en sus productos [2].

1.5.2. APLICACIONES.

El reconocimiento de voz ha tomado un papel importante en el ámbito tecnológico gracias a que es posible su desarrollo en software y en hardware, dando paso a que actualmente se use la voz como herramienta de control en diferentes campos de acción.

Entre las herramientas más comunes se encuentran convertidores de voz a texto, sistemas de seguridad, aplicaciones telefónicas en Smartphone y computadoras, donde también se destaca el control de diversos dispositivos mecánicos y electrónicos por medio de comandos de voz que facilitan y optimizan la interacción hombre-máquina, incluyendo aplicaciones caseras como la regulación de la calefacción y el aire acondicionado, apertura y cierre de cortinas o persianas, o el control sobre el encendido, apagado e intensidad de las luces [4].

Esta tecnología ha sido el tema central de desarrollo para varias compañías, incluyendo a INNOTECH SYSTEMS INC, marca que ha venido desarrollando diversos sistemas interactivos de reconocimiento de comandos de voz, diseñados para controlar funciones de diversos dispositivos modernos como relojes de alarma, televisores, iPod, DVD y decodificadores.

Entre sus principales productos se encuentran dos mandos a distancia universales llamados Accenda y SurfBoard, dispositivos que le permiten a cualquier persona controlar diversos sistemas de entretenimiento casero a través de comandos de voz configurables.

1.5.2.1. CONTROL REMOTO ACCENDA.

Este dispositivo es un control remoto universal que permite controlar cualquier televisor, decodificador, VCR y DVD mediante el protocolo habitual de oprimir botones o por medio de comandos de voz.

Accenda le permite al usuario grabar hasta 50 comandos de voz dependientes del locutor para los cuales se puede asignar una función diferente, convirtiendo un comando de voz en un mensaje infrarrojo.

Figura 1. Control remoto Accenda.



[20]. <http://www.innotechsystems.com/accenda/>

1.5.2.2. CONTROL REMOTO SURFBOARD.

Este es un dispositivo con el que se puede hacer uso de 12 comandos dependientes o 5 comandos independientes de locutor para controlar cualquier televisor o decodificador, permitiendo emplear una serie de comandos predefinidos por el fabricante o definidos por el usuario.

A diferencia del control remoto Accenda, el SurfBoard tiene una función de aprendizaje de mensajes infrarrojos útil en el caso de que algún mensaje enviado por los mandos a distancia convencionales no esté incluido en la base de datos del dispositivo.

Figura 2. Control remoto Surfboard.



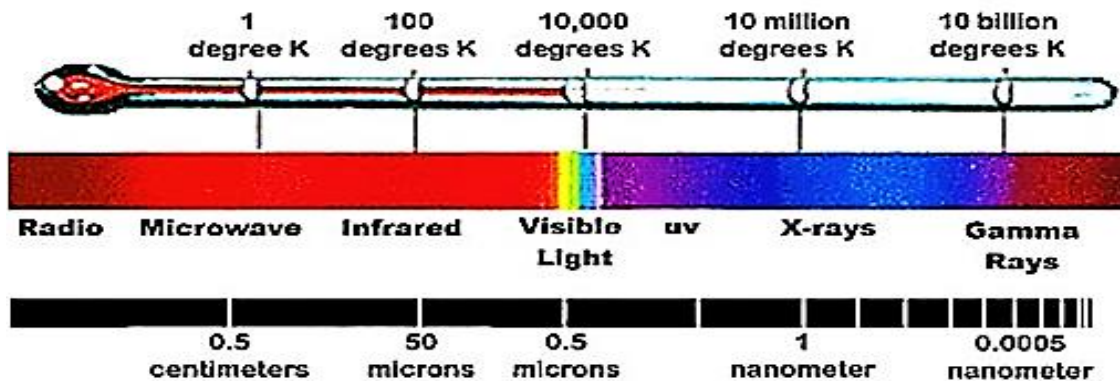
[21]. <http://www.innotech.com/surfboard/>

2. MARCO TEÓRICO.

2.1. RADIACIÓN INFRARROJA.

Es un tipo de radiación electromagnética cuya longitud de onda es menor a la de la luz visible, ubicada en la zona límite inferior del color rojo del espectro. Su descubrimiento se remonta a 1800, período en el que William Herschel utilizó un termómetro para medir el calor irradiado por cada color del espectro de la luz visible reflejado por un prisma de cristal. Con esta medición, Herschel descubrió que la temperatura aumentaba en una zona careciente de luz adyacente al color rojo del espectro, comprobando que el calor podía ser transmitido por una forma invisible de luz [8].

Figura 3. Temperatura y longitud de ondas de luz.



[22]. <http://diccionario.babylon.com/infrarrojo?&tl=/>

2.2. SENSORES INFRARROJOS.

Son dispositivos que detectan radiación infrarroja para posteriormente transformarla en una señal eléctrica. En la actualidad, para la transmisión y emisión de estos rayos se emplean principalmente sensores de luz infrarroja IRLED y fototransistores.

Un IRLED es un diodo emisor de rayos infrarrojos con el mismo aspecto de un led, cuya cápsula envolvente es de color azul o gris, en donde la pata más larga representa el ánodo y la pata más corta al cátodo.

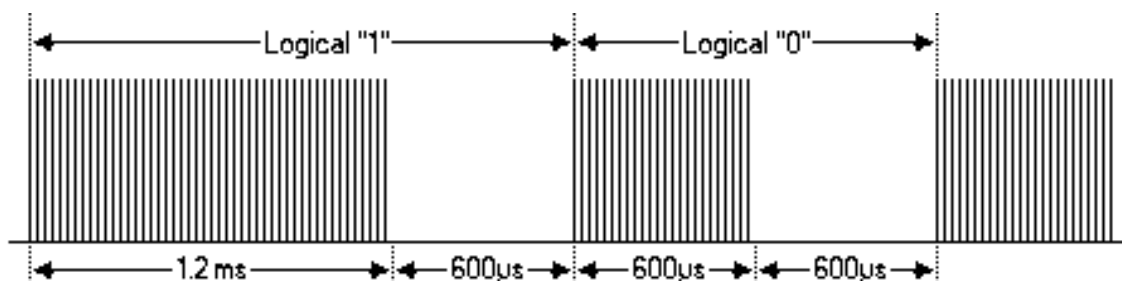
Por otro lado, un foto-transistor es un foto-detector cuyo funcionamiento es prácticamente igual al de un transistor clásico, variando principalmente en la carencia de una conexión base que es reemplazada con un cristal fotosensible que desbloquea el transistor al generar una corriente producida al recibir luz. Su aspecto es muy similar al de un led donde la pata larga es el negativo.

2.3. PROTOCOLO DE COMUNICACIÓN SONY SIRC PARA MANDOS INFRARROJOS.

Debido a la gran variedad de dispositivos de entretenimiento que emplean sistemas de comunicación vía infrarrojo, se han desarrollado varios protocolos de comunicación entre emisor y receptor, lo que ha permitido tener un control eficaz a distancia sobre diversas funciones de diferentes dispositivos. En su mayoría los protocolos vía infrarrojo de este tipo de sistemas de entretenimiento casero emplean una comunicación serial bit a bit que transmite información binaria con unos y ceros lógicos en forma de pulsos luminosos codificados por longitudes de ancho de pulso específicos, donde un impulso luminoso largo puede interpretarse como un "1" y uno corto como un "0". Este tipo de codificación es la que permite al receptor infrarrojo determinar la información recibida y la orden a ejecutar [12].

En este caso particular, el protocolo de comunicación vía infrarrojo SIRC emplea una longitud fundamental llamada periodo "T" de $600\mu\text{s}$ en sus mensajes, donde se usan múltiplos de esta longitud para diferenciar los bits enviados por el emisor.

Figura 4. Representación lógica del protocolo SIRC.



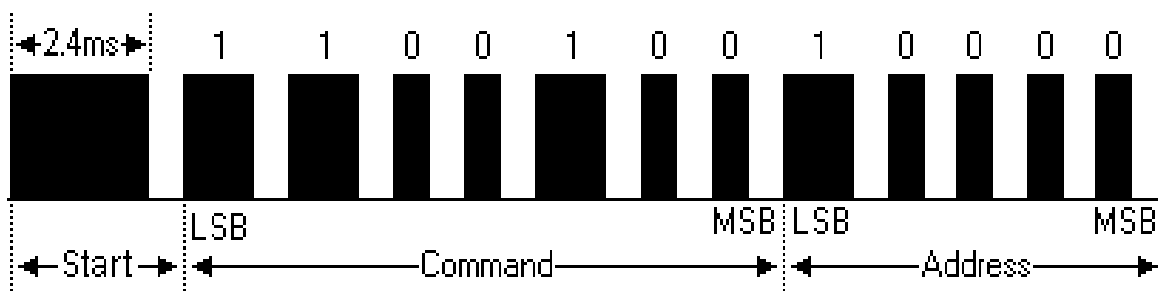
[23]. <http://www.sbprojects.com/knowledge/ir/sirc.php>

Como en todo protocolo, se emplea un bit de inicio o start cuya función es indicar al receptor que se disponga a recibir los datos emitidos por el emisor, siendo en este caso un pulso de longitud de $4T$ (2.4ms) a partir del cual los pulsos son más cortos, indicando información útil referente a unos y ceros lógicos [12].

Adyacente al bit de inicio, aparece una serie de siete bits enumerados del 0 al 6 correspondientes al comando de la tecla oprimida en un mando a distancia, conformado por los bit uno y los bit cero, donde el bit uno se representa mediante un pulso de $2T$ (1.2ms) y de un espacio sin señal adyacente de longitud T , y el bit cero se representa por un pulso de longitud T seguido por un espacio sin señal de igual longitud. [12].

Por último se envía una serie de 5 u 8 bits que hace referencia a la dirección de envío, o al dispositivo que se desea controlar (TV, DVD, etc.), el cual también viene codificado con bits uno y bits cero, completando así un mensaje de 12 bits que permite controlar una función del dispositivo que contiene el receptor respectivo.

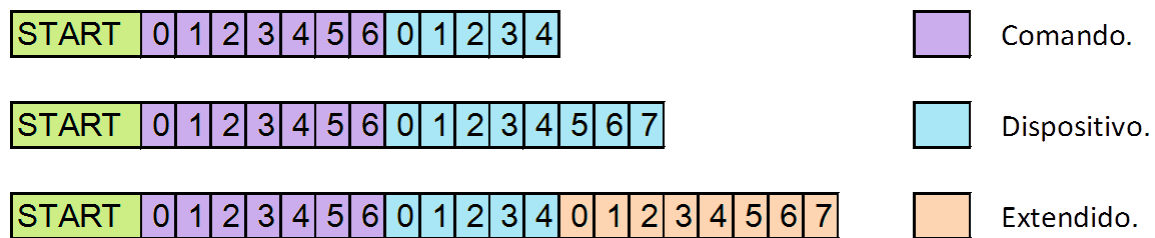
Figura 5. Mensaje infrarrojo en el protocolo SIRC.



[24]. <http://www.sbprojects.com/knowledge/ir/sirc.php>

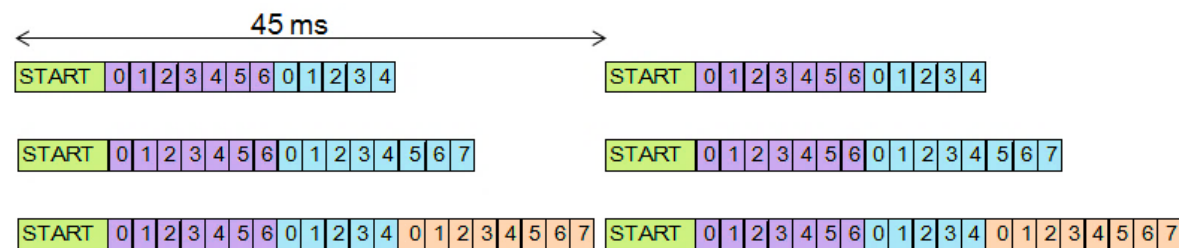
Adicionalmente, en algunos casos se emplea una serie extendida de 8 bits posterior a la dirección de envío, posibilitando el envío de tres tipos de mensajes diferentes en cuanto al número de bits.

Figura 6. Series de bits en el protocolo SIRC.



Dado el caso en el que el usuario mantenga oprimida cualquier tecla de un mando a distancia, automáticamente se reenvía la misma serie de bits en un intervalo de 45ms, enviando 3 series como mínimo.

Figura 7. Repetición de envío de mensajes en el protocolo SIRC.



2.4. MÓDULO DE RECONOCIMIENTO DE VOZ EASYVR.

Es un módulo multipropósito de reconocimiento de voz de bajo costo que permite usar 32 comandos dependientes de locutor en cualquier idioma y 26 comandos independientes de locutor pronunciados en Inglés, Japonés, Italiano, Alemán, Español y Francés, permitiéndole a cualquier usuario hacer uso del dispositivo en mención. [13].

Figura 8. Módulo EasyVR.



[25]. <http://www.veear.eu/products/easyvr/>

2.4.1. CARACTERÍSTICAS.

- 26 comandos predefinidos en modo independiente que pueden ser ejecutados en cualquiera de los 6 idiomas soportados [13].
- 32 comandos en modo dependiente para cualquier idioma.
- Dispone de una plataforma virtual para Windows que permite una mejor familiarización con el producto y su respectivo uso.
- Dispone de tres pines de entrada/salida configurables.
- Posee una salida de audio que permite conectar un altavoz de 8Ohms.
- 5 niveles de dificultad de reconocimiento de comandos de voz.

2.4.2. MICRÓFONO.

El módulo incluye un micrófono electret omnidireccional de referencia HORN-EM9745P-382, con las siguientes características:

- Sensibilidad de -38dB (0dB=1V/Pa @1kHz).
- Impedancia de 2.2 k Ω .
- Respuesta prácticamente plana entre 100Hz – 20kHz.

2.4.3. EASYVR ARDUINO SHIELD.

Con el fin de acoplar el módulo de reconocimiento de voz y el Arduino UNO de una manera sencilla, se implementó un Shield o escudo que permite un acople entre las dos tarjetas sin necesidad de usar cable alguno.

Figura 9. Acople de EasyVR Arduino Shield y Arduino UNO.

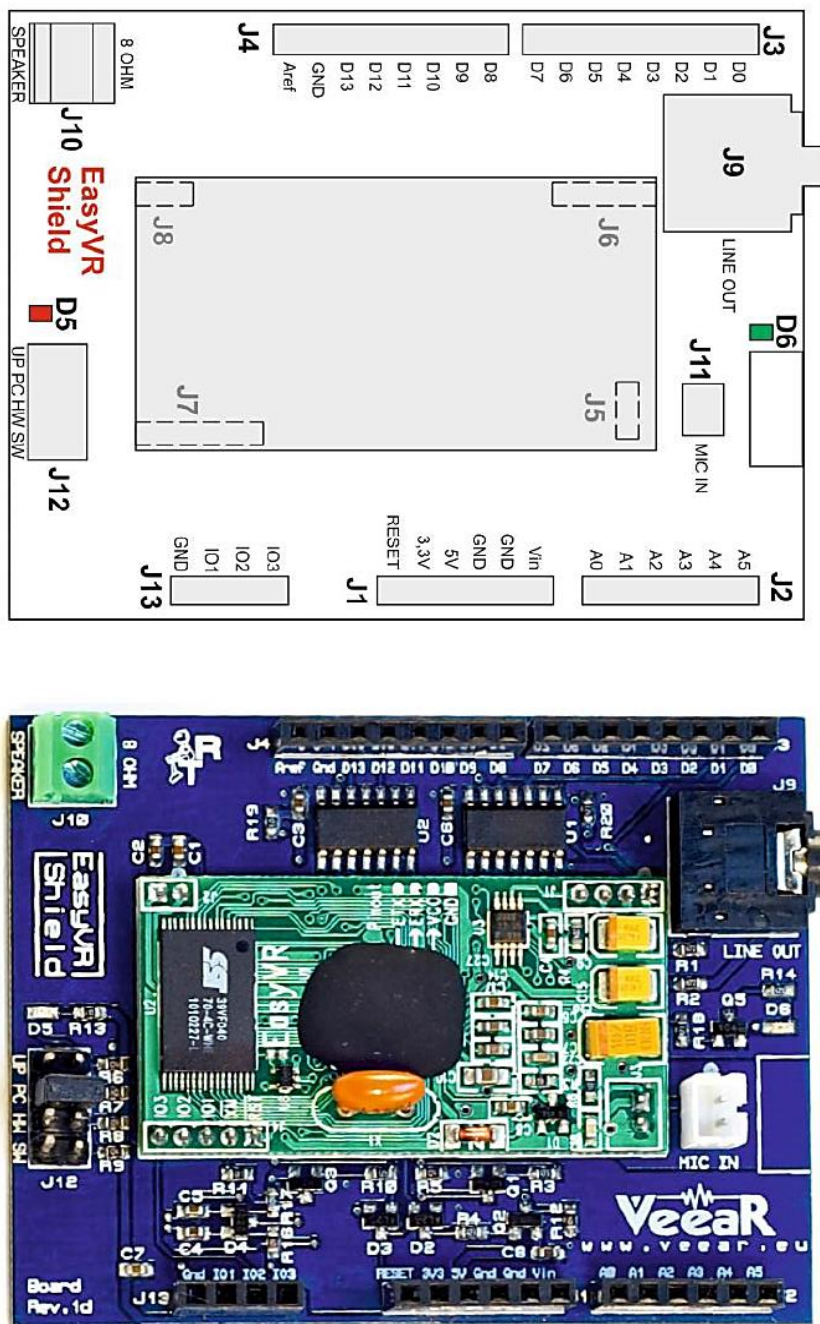


[26]. <http://www.zagrosrobotics.com/shop/item.aspx?itemid=865>

Con este acople es posible programar ambos dispositivos a través de la plataforma virtual de Arduino, haciendo uso la librería EasyVR. Para tal fin, el shield interconecta internamente los dos dispositivos, conectando los puertos ERX y ETX del módulo de reconocimiento de voz a cualquier par de pines multipropósito del Arduino [14].

2.4.3.1. DIAGRAMA DE PINES EASYVR ARDUINO SHIELD.

Figura 10. Diagrama de pines EasyVr Arduino Shield.



[27]. <http://www.robotshop.com/media/files/pdf/easyvr-user-manual-vrb-001.pdf>

Tabla 1. Descripción de pines y conectores del EasyVR Arduino Shield.

DESCRIPCIÓN PINES Y CONECTORES EASYVR ARDUINO SHIELD		
SECCIÓN	NOMBRE	DESCRIPCIÓN
J1		Secciones de pines bajo la misma configuración de los pines del Arduino UNO.
J2		
J3		
J4		
J9	LINE OUT	Salida para audífonos 3.5mm (16Ω - 32Ω).
J10	SPEAKER	Salida diferencial de audio (8Ω).
J11	MIC IN	Entrada de micrófono.
J12	MODE JUMPER	Selección de modo de operación.
J13	I/O 1	Pines multipropósito configurables IN/OUT.
	I/O 2	
	I/O 3	
	I/O 4	

Como se observa en la figura 10, el módulo de reconocimiento de voz EasyVR viene acoplado al Shield, cuyos pines vienen organizados de manera similar a los pines presentes en un Arduino UNO, específicamente en las secciones J1, J2, J3 y J4, referenciadas en la tabla anterior.

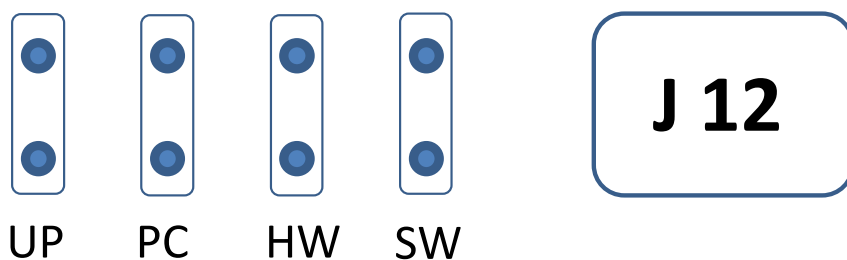
Para la configuración y uso de la sección J10, se hace necesario el uso del software libre Sensory Quick Synthesis 5, el cual permite almacenar en la memoria del dispositivo audios mono en formato .WAV exportados a 22050 Hz y 16 bits, los cuales pueden ser reproducidos a través de esta salida.

Por último, en la sección J12 se encuentran 4 posibles conexiones para un jumper, el cual según su posición le permite al módulo de reconocimiento de voz entrar en distintos modos de operación [14].

2.4.3.2. MODOS DE OPERACIÓN.

Los cuatro modos de operación en los que puede entrar el EasyVR son UP mode, PC mode, HW mode y SW mode, configurables en la sección J12 del shield.

Figura 11. Sección J12 EasyVR Arduino Shield.



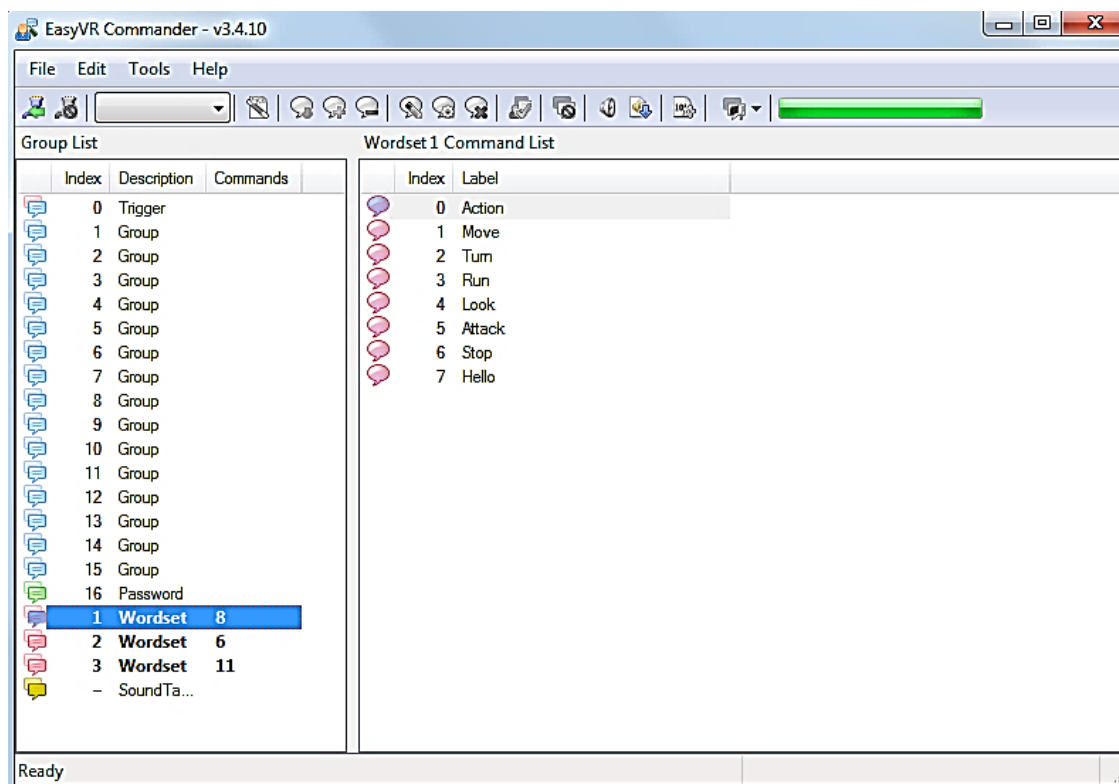
- **PC Mode:** En esta posición se permite la comunicación serial entre el EasyVR y el software EasyVR Commander, el cual permite realizar la fase de entrenamiento de comandos de voz que se acomoden a las necesidades del usuario.
- **SW Mode:** En esta configuración se puede programar y monitorear el módulo de reconocimiento de voz a través de la plataforma virtual de Arduino, permitiendo cargar cualquier código que permita la operación del Arduino y el EasyVR en conjunto.
- **UP Mode:** Se emplea para descargar archivos de audio en la memoria del dispositivo, a través del software Quick Synthesis 5.

2.4.3.3. EASYVR COMMANDER.

Es el software que permite configurar, ordenar y entrenar los comandos de voz, para lo cual solo es necesario conectar el acople del Shield con el Arduino al computador por medio de un cable USB, poniendo el jumper de la sección J12 en la posición PC.

En esta plataforma es donde el usuario dispone de los comandos independientes presentes en cualquiera de los 3 grupos de comandos predefinidos llamados “Wordset”, además de tener 15 grupos que permiten la organización de los 32 comandos dependientes que se deseen.

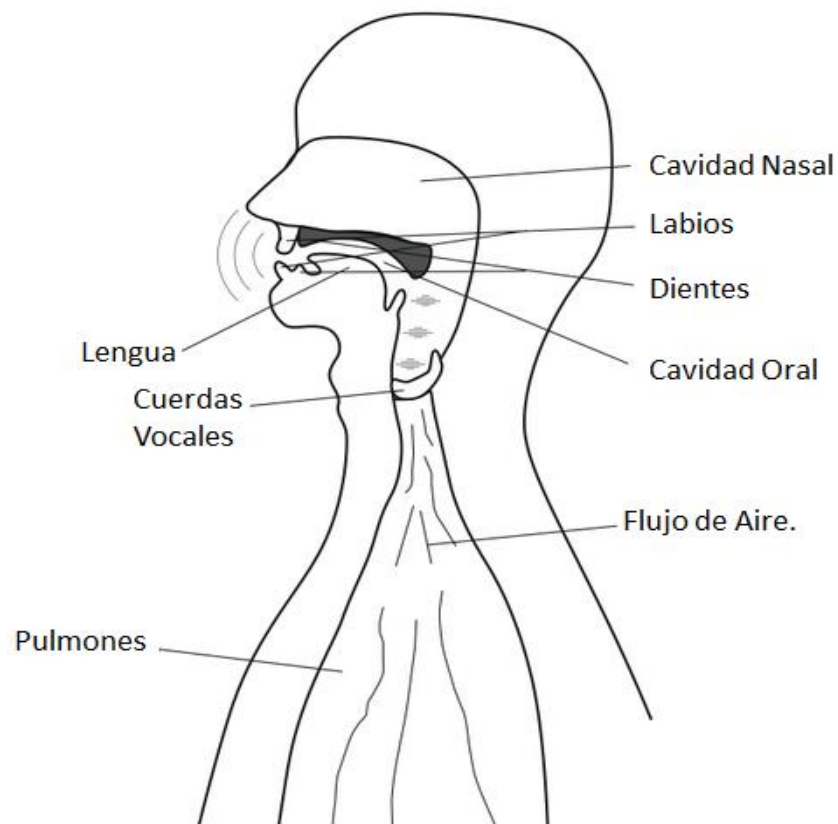
Figura 12. Pantalla inicial software EasyVR Commander.



2.5. VOZ HUMANA.

La voz es el sonido que produce la vibración de las cuerdas vocales en los humanos, las cuales son excitadas por el flujo de aire proveniente de los pulmones. Una vez dicho aire pasa las cuerdas vocales, se encuentra con el tracto vocal donde se le da a la voz de cada persona diferentes características particulares como el tono, el timbre y el acento [15].

Figura 13. Tracto Vocal.



[28]. <http://cnx.org/content/m0049/latest/>

Dado que la fisionomía de cada persona es diferente, dichas características varían según el género y la edad de las personas, presenciando principalmente un mayor tamaño de las cuerdas vocales en hombres adultos, generando un tono de voz más bajo que en las mujeres y en los niños.

Entendiendo la voz humana como una señal acústica, es posible su representación en función del tiempo en donde se puede determinar el dominio de la frecuencia, su frecuencia fundamental y sus respectivas frecuencias formantes, las cuales varían según la palabra pronunciada y la persona que la pronuncia.

Estas propiedades de la voz humana son las características principales que se emplean en las tecnologías que permiten el reconocimiento de locutor o de comandos de voz, en donde adicionalmente se pueden tener en cuenta las características sonoras propias de un lenguaje en particular denominadas fonemas, que al ser conjugados forman morfemas o palabras.

En el caso del idioma español, sin tener en cuenta sus variaciones, existen 22 fonemas específicos que se dividen en consonánticos y vocálicos, para los cuales se tienen en cuenta los sonidos que emite una persona al pronunciar letras del idioma en cuestión.

Tabla 2. Lista de Fonemas en Español.

LISTA DE FONEMAS DEL ESPAÑOL	
VOCÁLICOS	
SÍMBOLO	DESCRIPCIÓN
/a/	Fonema vocálico de apertura máxima
/e/	Fonema vocálico palatal de apertura media
/i/	Fonema vocálico palatal y apertura mínima
/o/	Fonema vocálico velar de apertura media
/u/	Fonema vocálico velar de apertura mínima
CONSONÁNTICOS	
SÍMBOLO	DESCRIPCIÓN
/B/	Fonema obstruyente bilabial sonoro (grafías: <i>b</i> , <i>v</i> y <i>w</i>).
/č/	Fonema africado palatal (grafía <i>ch</i>).
/D/	Fonema obstruyente coronal-alveolar sonoro
/f/	Fonema labial, fricativo, sordo, oral, en muchas zonas fricativo bilabial
/G/	Fonema obstruyente velar sonoro (grafías <i>g</i> y <i>gu</i> , como en gato y guerra)
/x/	Fonema fricativo velar (grafías <i>g</i> y <i>j</i> , como en giro y jirafa)
/k/	Fonema oclusivo velar sordo (grafías <i>c</i> , <i>qu</i> y <i>k</i>).
/l/	Fonema lateral (coronal-)alveolar.
/m/	Fonema nasal labial
/n/	Fonema nasal (coronal-)alveolar
/ñ/	Fonema nasal palatal
/p/	Fonema oclusivo (bi)labial sordo.
/r/	Fonema vibrante simple (grafía <i>-r-</i> , como en arado)
/r/(rr)/	Fonema vibrante múltiple (grafía <i>-rr-</i> y <i>r-</i> , como en carro y rosa)
/s/	Fonema fricativo (coronal-)alveolar (grafía <i>s</i> , en algunas variedades <i>z</i> y <i>c</i>)
/t/	Fonema oclusivo (coronal-)alveolar sordo
/y/	Fonema sonorante palatal (grafía <i>y</i> , en las zonas yeístas también //)

[29]. <http://www.gramaticas.net/2011/05/fonemas-definicion-y-ejemplos.html>

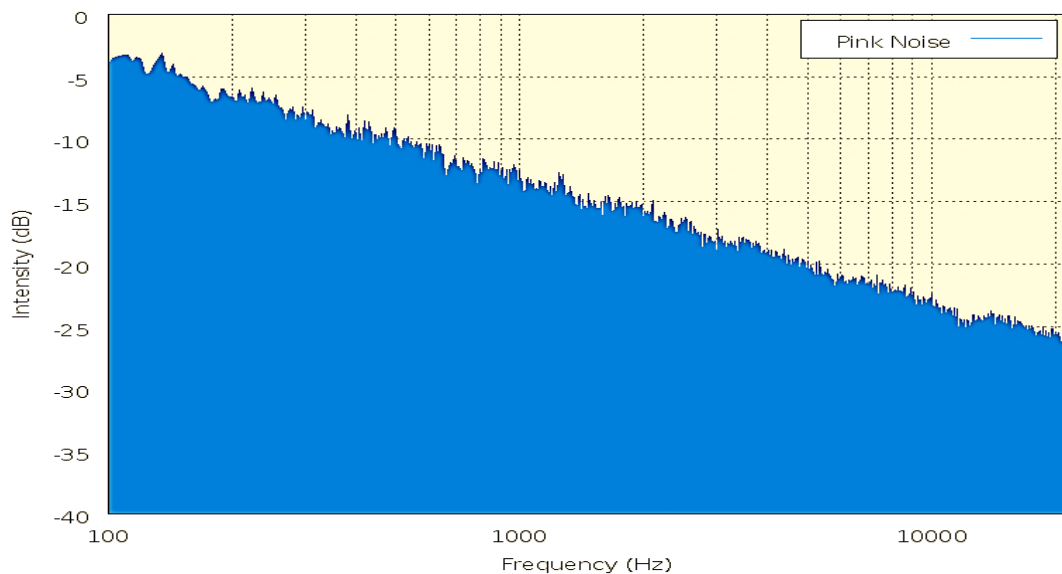
2.6. RUIDO.

Para este proyecto, el ruido se puede definir como todo sonido no deseado que interfiere con el reconocimiento de los comandos de voz por parte del módulo EasyVR. En el caso particular de un ambiente casero, debido a que estos sonidos no tienen un aspecto audible característico, es necesario realizar un análisis con una señal de ruido rosa que permita caracterizar cualquier tipo de sonido.

2.6.1. RUIDO ROSA.

Es un tipo de ruido en donde el nivel de presión sonora es inversamente proporcional a la frecuencia, cuyas bandas de octava mantienen el mismo nivel sonoro. Teniendo en cuenta que al aumentar de banda de octava se duplica el ancho de la misma, se tiene un decremento de 3dB por octava [16].

Figura 14. Espectro Ruido Rosa.

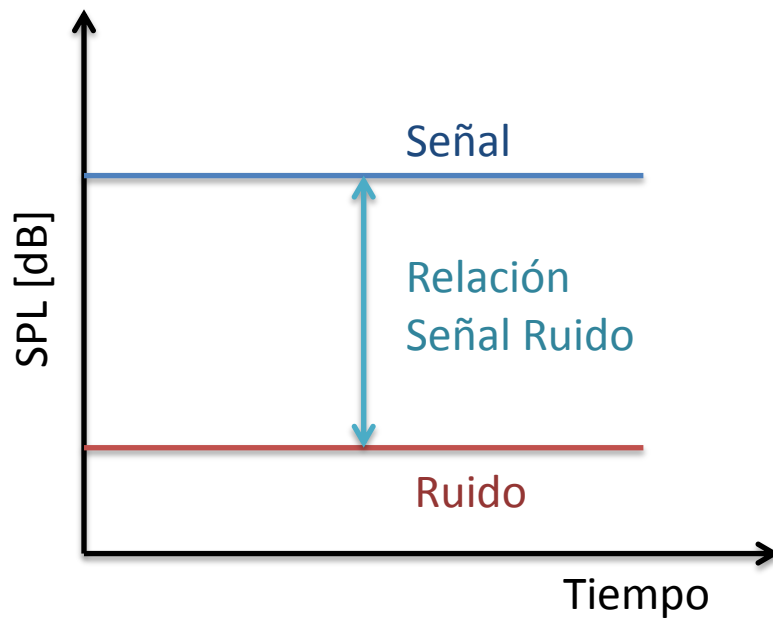


[30]. http://commons.wikimedia.org/wiki/File:Pink_noise_spectrum.png

2.7. RELACIÓN SEÑAL RUIDO.

Para este proyecto, la relación señal ruido hace referencia a la diferencia de presión sonora, representada en decibeles [dB], entre el ruido de fondo y un nivel de referencia, en este caso el ruido presente en un ambiente casero y el nivel de presión sonora generado por el usuario que desee controlar sus sistemas de entretenimiento por medio de comandos de voz.

Figura 15. Relación señal/ruido.

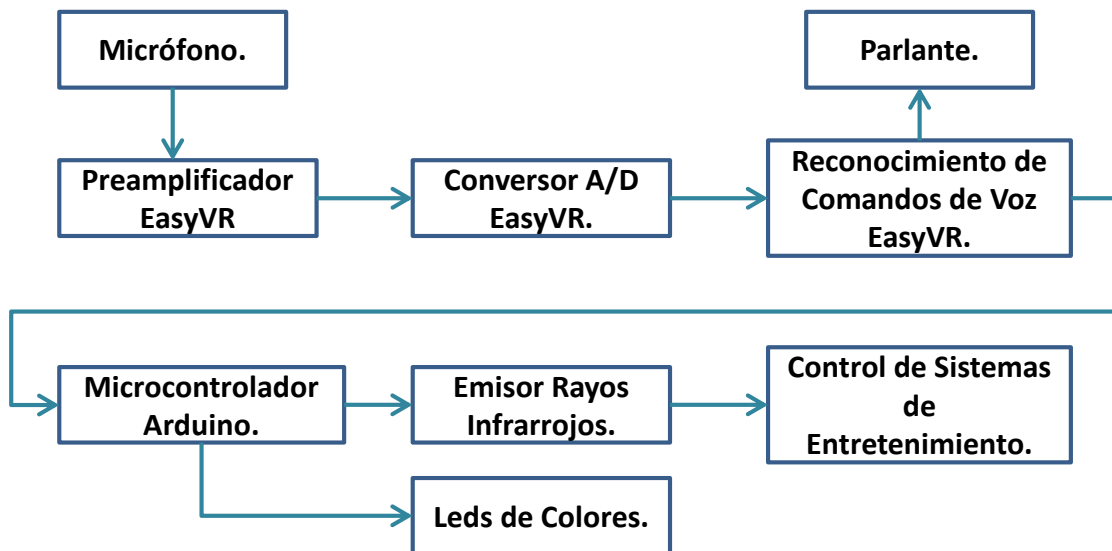


3. DESARROLLO INGENIERIL.

3.1. DIAGRAMA DE BLOQUES.

El siguiente diagrama representa el flujo de la información procesada por el dispositivo desarrollado, la cual está disponible a partir de la entrada de micrófono del módulo de reconocimiento de voz EasyVR, proceso que se finaliza con el envío de los mensajes infrarrojos destinados a controlar diversas funciones de los sistemas de entretenimiento casero.

Figura 16. Diagrama de Bloques.



Una vez el transductor de entrada convierte la energía acústica en energía eléctrica, se lleva a cabo la preamplificación de la señal y su posterior conversión análogo-digital en el módulo de reconocimiento de voz. Después de realizado el

proceso de reconocimiento, el Arduino lleva a cabo los procesos de control sobre los leds de colores y el IRLED emisor que posteriormente envía la información infrarroja a cada uno de los sistemas de entretenimiento casero a controlar.

3.2. FUNCIONES A CONTROLAR.

Inicialmente se dio paso al desarrollo de un modo de operación denominado “PELÍCULAS”, el cual permitiese controlar un televisor, un equipo de sonido y un DVD, haciendo uso de las funciones más usadas de cada sistema a la hora de ver una película en un centro de entretenimiento casero.

Tabla 3. Funciones a controlar.

FUNCIONES A CONTROLAR				
TELEVISOR	E. DE SONIDO	DVD		
Encendido	Encendido	Encendido	Fast Forward	Izquierda
Apagado	Apagado	Apagado	Rewind	Enter
Mute	Volumen +	Play	Arriba	Eject
	Volumen -	Pausa	Abajo	Menú
		Stop	Derecha	

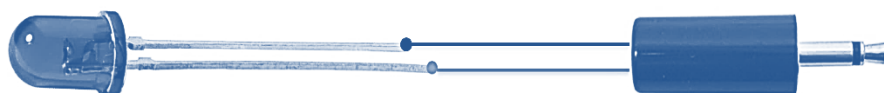
3.3. ADQUISICIÓN DE MENSAJES INFRARROJOS.

Seguidamente se inició la adquisición de los datos que envían los respectivos mandos a distancia para cada una de las funciones mencionadas anteriormente, empleando un circuito receptor de infrarrojos que permitiese digitalizar los datos en mención a través de un software de captura y edición de audio, visualizando la señal eléctrica generada por un receptor infrarrojo conectado a la entrada de micrófono de un computador, al momento de recibir información proveniente de los mandos a distancia.

3.3.1. CIRCUITO RECEPTOR DE MENSAJES INFRARROJOS.

El circuito implementado para adquirir la información necesaria incluye un receptor infrarrojo y un conector mono de 3.5mm, permitiendo transformar los mensajes enviados por un control remoto en una señal eléctrica que posteriormente puede ser digitalizada por medio de la entrada de micrófono de un computador.

Figura 17. Circuito receptor de mensajes Infrarrojos.



Una vez conectado el circuito, se configuró el software de captura de audio a un sample rate de 96000 Hz, permitiendo visualizar los mensajes mostrados en las tablas 4, 5, 6 y 7, cuyas representaciones lógicas se exponen en la tabla 8.

3.3.2. MENSAJES INFRARROJOS DIGITALIZADOS.

Definida la forma de adquirir los mensajes enviados por los mandos a distancia, se inició el proceso de captura y digitalización necesario para llegar a controlar las funciones mencionadas en la tabla 3.

Tabla 4. Mensajes infrarrojos para el televisor.

MENSAJES INFRARROJOS TELEVISOR SONY	
FUNCIÓN.	MENSAJE INFRARROJO DIGITALIZADO.
Prender Apagar	
Mute	

Tabla 5. Mensajes infrarrojos para el equipo de sonido.

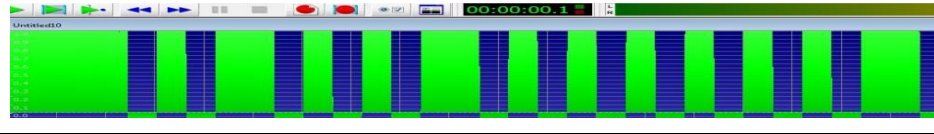
MENSAJES INFRARROJOS EQUIPO DE SONIDO SONY.	
FUNCIÓN.	MENSAJE INFRARROJO.
On/Off.	
Volumen +	
Volumen -	

Tabla 6. Mensajes infrarrojos para el DVD.

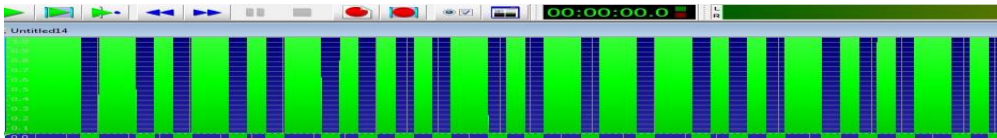
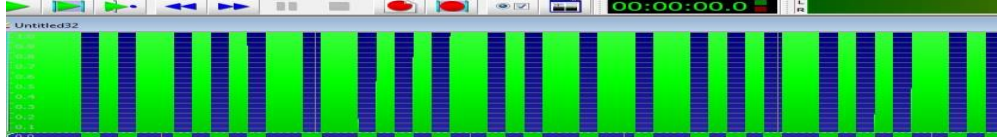
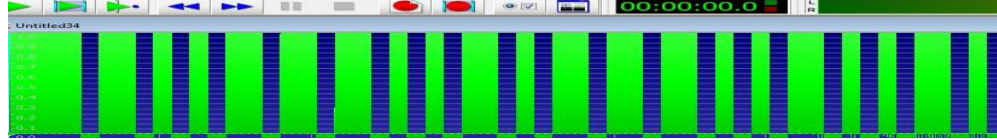
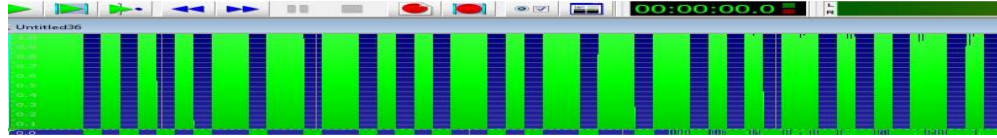
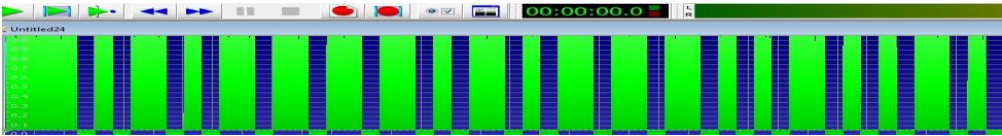
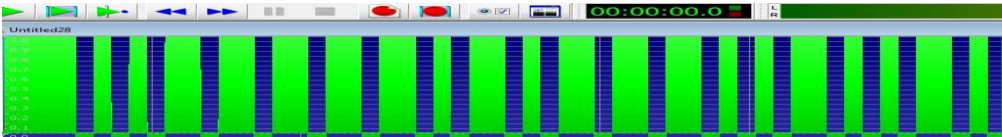
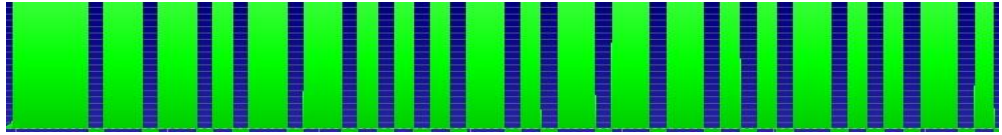
MENSAJES INFRARROJOS DVD SONY	
FUNCIÓN.	MENSAJE INFRARROJO.
Prender Apagar	 The image shows a software interface for creating an infrared remote control message. At the top, there is a toolbar with various icons for editing the message. Below the toolbar, the message is displayed as a series of vertical bars of varying heights, representing the pulse-width modulation of the infrared signal. The message is labeled 'Untitled14' and the time is set to '00:00:00.0'.
Play.	 The image shows a software interface for creating an infrared remote control message. At the top, there is a toolbar with various icons for editing the message. Below the toolbar, the message is displayed as a series of vertical bars of varying heights, representing the pulse-width modulation of the infrared signal. The message is labeled 'Untitled32' and the time is set to '00:00:00.0'.
Pause.	 The image shows a software interface for creating an infrared remote control message. At the top, there is a toolbar with various icons for editing the message. Below the toolbar, the message is displayed as a series of vertical bars of varying heights, representing the pulse-width modulation of the infrared signal. The message is labeled 'Untitled34' and the time is set to '00:00:00.0'.
Stop.	 The image shows a software interface for creating an infrared remote control message. At the top, there is a toolbar with various icons for editing the message. Below the toolbar, the message is displayed as a series of vertical bars of varying heights, representing the pulse-width modulation of the infrared signal. The message is labeled 'Untitled36' and the time is set to '00:00:00.0'.
Fast Forward.	 The image shows a software interface for creating an infrared remote control message. At the top, there is a toolbar with various icons for editing the message. Below the toolbar, the message is displayed as a series of vertical bars of varying heights, representing the pulse-width modulation of the infrared signal. The message is labeled 'Untitled42' and the time is set to '00:00:00.0'.
Rewind.	 The image shows a software interface for creating an infrared remote control message. At the top, there is a toolbar with various icons for editing the message. Below the toolbar, the message is displayed as a series of vertical bars of varying heights, representing the pulse-width modulation of the infrared signal. The message is labeled 'Untitled42' and the time is set to '00:00:00.0'.
Arriba.	 The image shows a software interface for creating an infrared remote control message. At the top, there is a toolbar with various icons for editing the message. Below the toolbar, the message is displayed as a series of vertical bars of varying heights, representing the pulse-width modulation of the infrared signal. The message is labeled 'Untitled22' and the time is set to '00:00:00.0'.

Tabla 7. Mensajes infrarrojos para el DVD.

MENSAJES INFRARROJOS DVD SONY	
FUNCIÓN.	MENSAJE INFRARROJO.
Abajo.	
Derecha.	
Izquierda.	
Enter.	
Eject/Open	
Menu.	

3.3.2.1. CODIFICACIÓN MENSAJES INFRARROJOS.

Una vez realizada la digitalización de los comandos infrarrojos de 12 y 20 bits usados en el envío de información por parte de los mandos a distancia de la marca SONY, se llevó a cabo su respectiva transcripción a los valores lógicos de cada mensaje a enviar, los cuales hacen referencia a los diferentes tiempos durante los cuales el IRLED transmisor infrarrojo debe prenderse y apagarse sucesivamente.

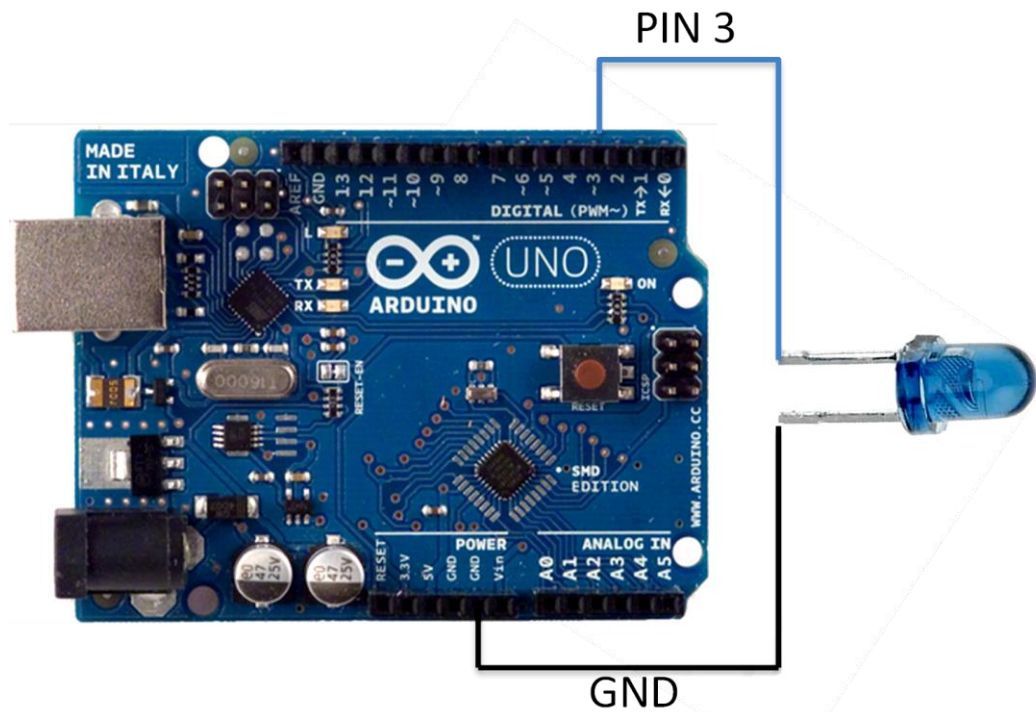
Tabla 8. Codificación lógica de mensajes infrarrojos.

MENSAJES CODIFICADOS			
DISPOSITIVO	FUNCIÓN	MENSAJE	# BITS
Televisor	On/Off	101010010000	12
	Mute	001010010000	12
E. de Sonido	On/Off	101010000001	12
	Volumen +	010010000001	12
	Volumen -	110010000001	12
Dvd	ON/Off	10101000101110010010	20
	Play	01001100101110010010	20
	Pause	10011100101110010010	20
	Stop	00011100101110010010	20
	Fast Forward	11000100101110010010	20
	Rewind	01000100101110010010	20
	Arriba	10011110101110010010	20
	Abajo	01011110101110010010	20
	Derecha	00111110101110010010	20
	Izquierda	11011110101110010010	20
	Enter	11010000101110010010	20
	Eject	01101000101110010010	20
	Menu	11011000101110010010	20

3.4. CIRCUITO EMISOR DE MENSAJES INFRARROJOS.

Una vez se adquirieron los mensajes que el dispositivo controlador en desarrollo debía enviar a cada uno de los sistemas de entretenimiento, se implementó un circuito emisor de rayos infrarrojos que consta de un Arduino UNO y un IRLED transmisor.

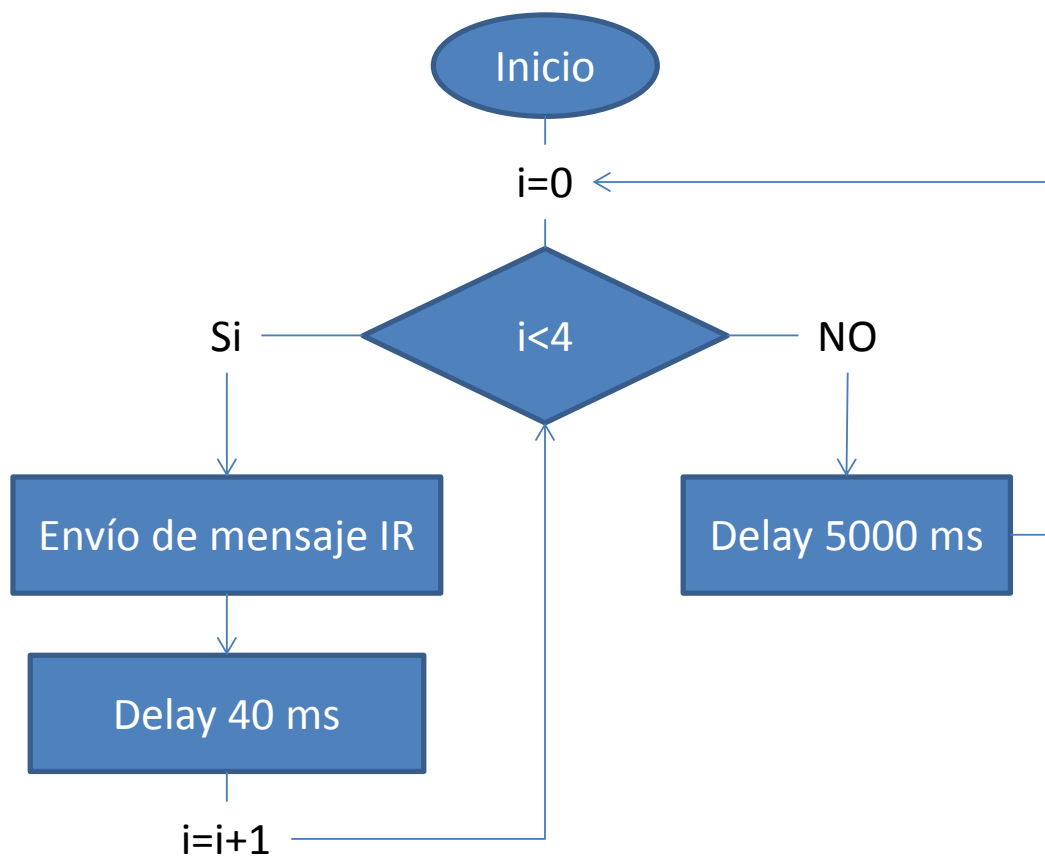
Figura 18. Circuito emisor de mensajes infrarrojos.



El IRLED transmisor se conectó a tierra y al pin 3 del Arduino, para lo cual se desarrolló un código que incluyera la librería IRremote, la cual permite enviar mensajes infrarrojos basados en los protocolos SIRC, NEC, RC5, RC6 y RAW.

Dicho código permite enviar 3 series de un mismo comando infrarrojo cada 5 segundos, con el fin de comprobar si los mensajes adquiridos anteriormente correspondían a las funciones a controlar en los tres sistemas de entretenimiento, cuyo funcionamiento se ve reflejado en la figura 19. [Ver Anexo A, Código 1].

Figura 19. Diagrama de flujo Código 1.



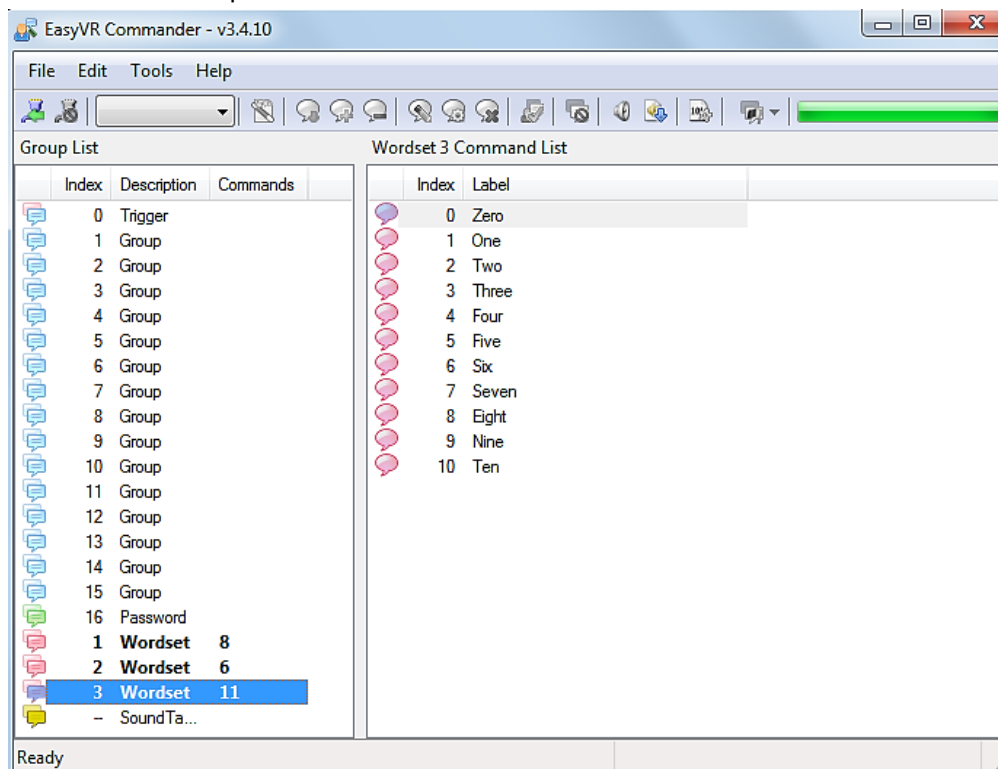
Una vez comprobado que todos los mensajes codificados en la tabla 8 correspondieran, se llevó a cabo la selección de los comandos dependientes e independientes de locutor que posteriormente permitiesen el envío de los mensajes infrarrojos respectivos.

3.5. COMANDOS DE VOZ.

3.5.1. COMANDOS INDEPENDIENTES DE LOCUTOR.

Como se expuso anteriormente, el módulo de reconocimiento de voz incorpora 26 comandos independientes de locutor predefinidos que no requieren una fase de entrenamiento, lo cual le permite a cualquier persona hacer uso de los mismos por medio de la voz. Para este caso se estudió la efectividad de los comandos independientes presentes en el Wordset 3 del software EasyVR Commander, donde se encuentran asignados los comandos alusivos a los números del 0 al 10, los cuales se muestran en la siguiente figura.

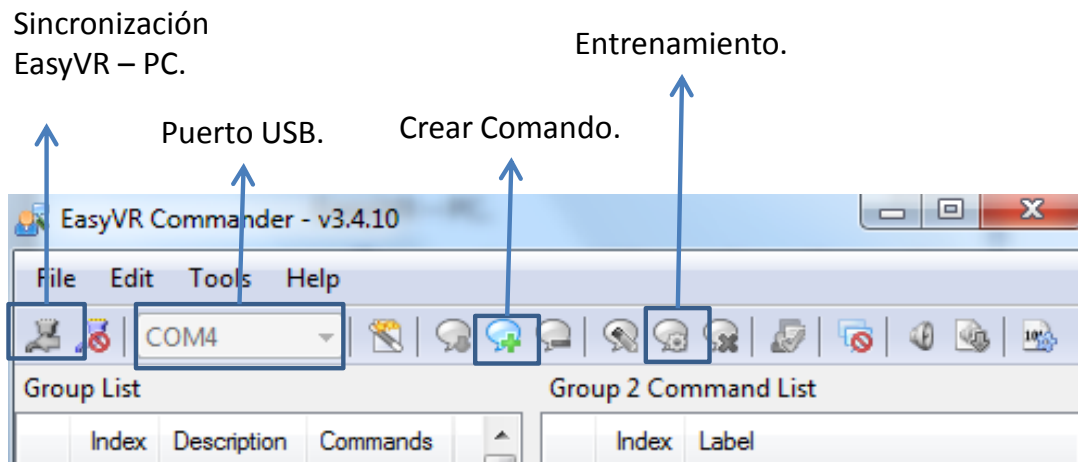
Figura 20. Comandos independientes de locutor Wordset 3.



3.5.2. COMANDOS DEPENDIENTES DE LOCUTOR.

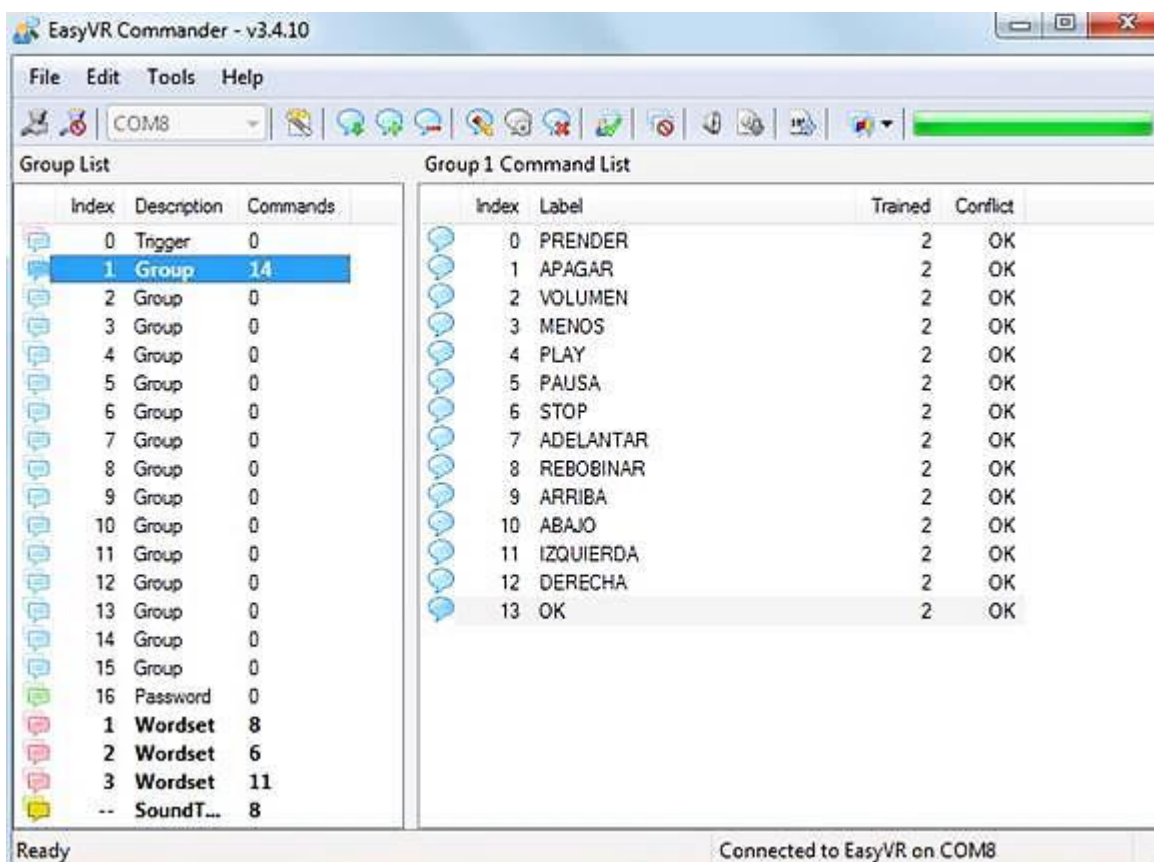
A diferencia de los comandos independientes, es necesario llevar a cabo una fase de entrenamiento para usar comandos dependientes de locutor, para lo cual se debe establecer la comunicación entre el computador y el EasyVR a través del puerto USB del Arduino, después de haber puesto el jumper de la sección J12 en la posición PC. Seguidamente se procede a entrenar los comandos por medio del software EasyVR Commander, haciendo uso de los botones mostrados en la siguiente figura.

Figura 21. Entrenamiento de comandos dependientes de locutor.



Una vez los comandos de voz son creados y entrenados, el software muestra una lista con el orden y los nombres asignados por el usuario.

Figura 22. Comandos dependientes.



Como se mencionó anteriormente, se puede entrenar al módulo de reconocimiento de voz para que sea capaz de reconocer 32 comandos dependientes, los cuales tienen un índice y un grupo asignado que sirven de referencia en la realización del código de programación respectivo.

3.6. CIRCUITO DEL DISPOSITIVO CONTROLADOR.

Una vez acoplado el EasyVR Arduino Shiled al Arduino, se definieron las siguientes conexiones.

Tabla 9. Listado de Conexiones.

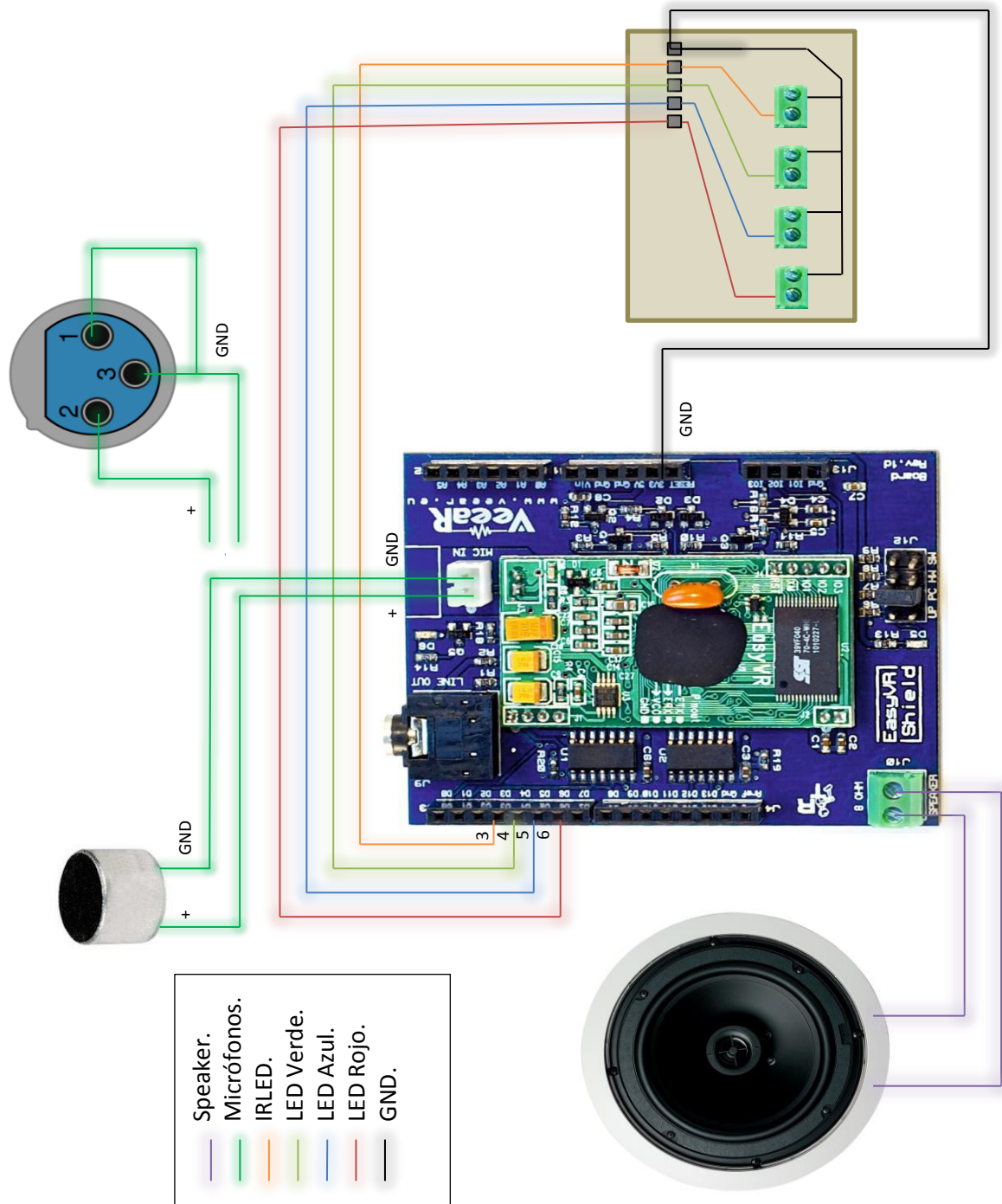
CONEXIONES	
NOMBRE	DESCRIPCIÓN
Speaker	J10 a Speaker de 8Ω .
Micrófonos.	J11 a Conector hembra XLR de chasis o mic. electret.
IRLED	Pin 3 a IRLED
LED VERDE	Pin 4 a LED verde.
LED AZUL	Pin 5 a LED azul.
LED ROJO	Pin 6 a LED rojo.
GND	Tierra a Tierra.

El circuito consta esencialmente del IRLED emisor, tres leds de colores, un conector XLR hembra de chasis, un parlante de 2" de 8Ω, un Arduino UNO y el módulo de reconocimiento de voz con su respectivo micrófono de referencia HORN.

Los tres LEDS de distinto color conectados a los pines 4, 5, y 6, se implementaron con el fin de permitirle al usuario una mayor comprensión del funcionamiento del dispositivo inalámbrico de control al momento de usarlo. Dichas conexiones se realizaron por medio de borneras que interconectasen los pines con las patas de los LEDS en cuestión.

3.6.1. ESQUEMÁTICO.

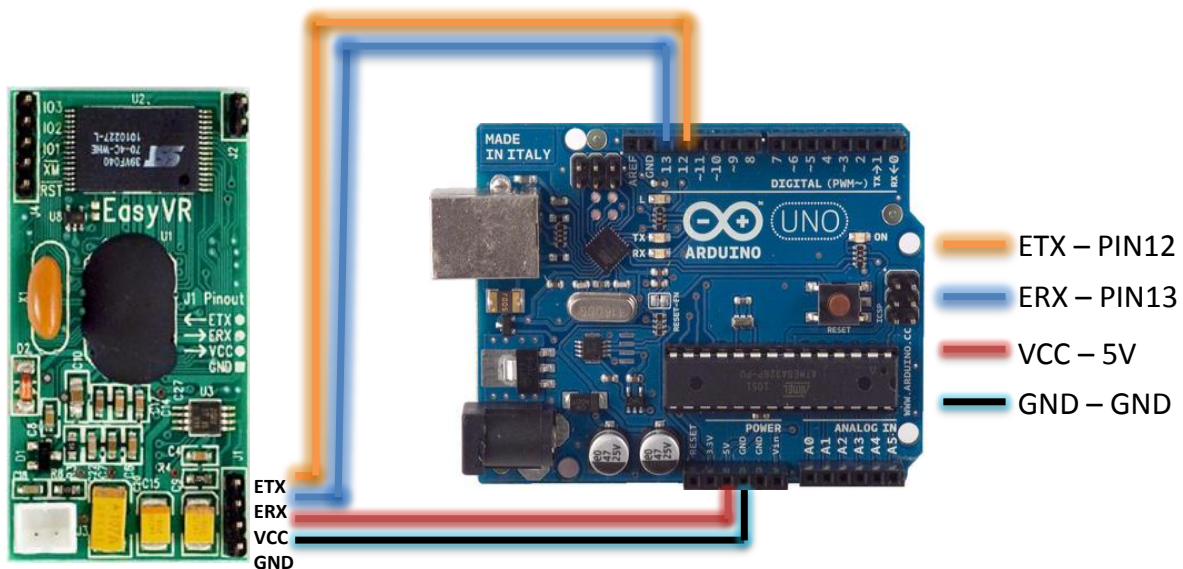
Figura 23. Esquemático.



Como se muestra en la figura anterior, el conector XLR hembra de chasis permite la inserción de la señal eléctrica generada por micrófonos ajenos al módulo de reconocimiento de voz, los cuales se emplearon posteriormente para realizar distintos análisis de efectividad. Al conector XLR se soldó un conector idéntico al conector presente en el micrófono HORN, permitiendo la conexión entre dichos micrófonos y la entrada respectiva del módulo EasyVR. Adicionalmente se acopló un speaker de 2 pulgadas que facilitara la reproducción de los sonidos presentes en la memoria del microcontrolador, los cuales indicarían un proceso de reconocimiento de comandos de voz satisfactorio.

En la siguiente figura se muestra la conexión interna que permite la comunicación serial entre el EasyVR y el Arduino UNO, la cual se realiza automáticamente al realizar el acople mostrado en la figura 9.

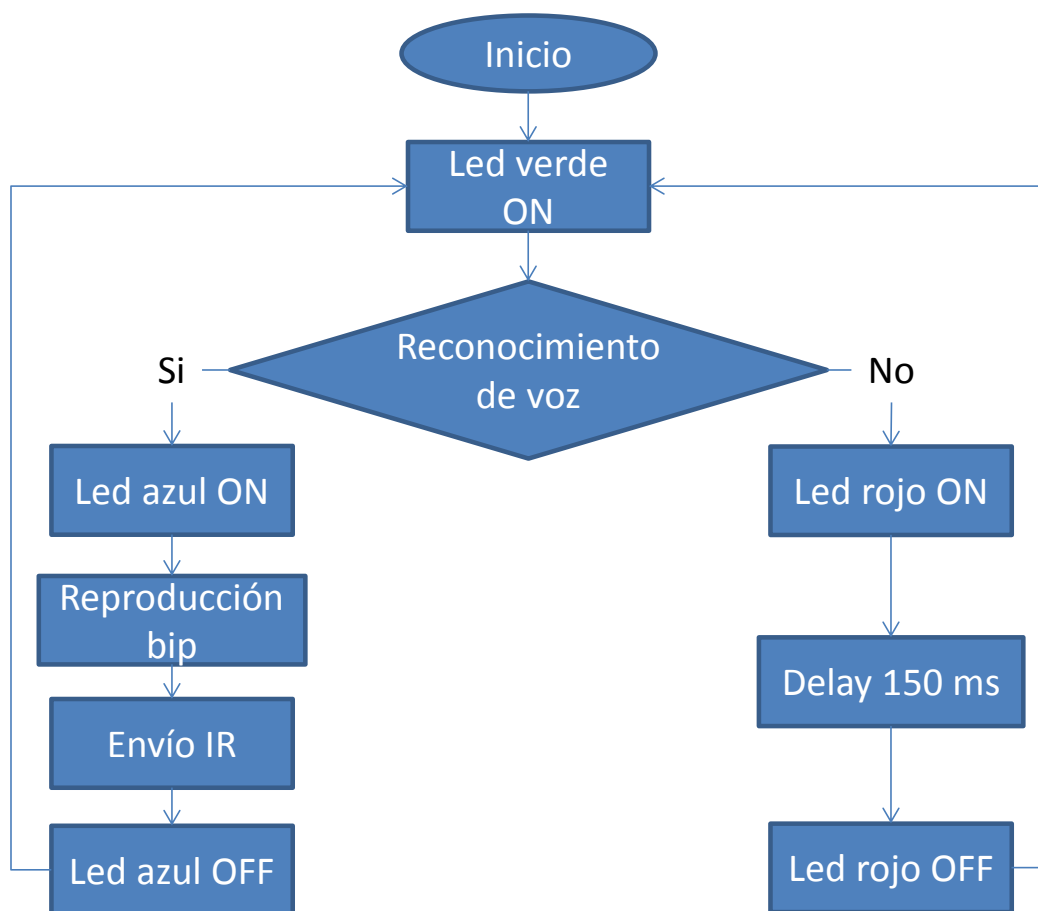
Figura 24. Conexión EasyVR – Arduino UNO.



3.7. CÓDIGO DE PROGRAMACIÓN BASE.

Una vez se realizó el montaje del circuito mostrado, se desarrolló un algoritmo para realizar las pruebas de efectividad. Con el fin de tener control de manera dependiente e independiente del locutor, se realizaron distintas versiones del mismo código que se acoplaron a las necesidades de las mediciones expuestas en la sección 3.9, versiones cuyo funcionamiento se ve representado en el siguiente diagrama de flujo. [Ver Anexo A, Código 2 y 3].

Figura 25. Diagrama de flujo código 2 y código 3.

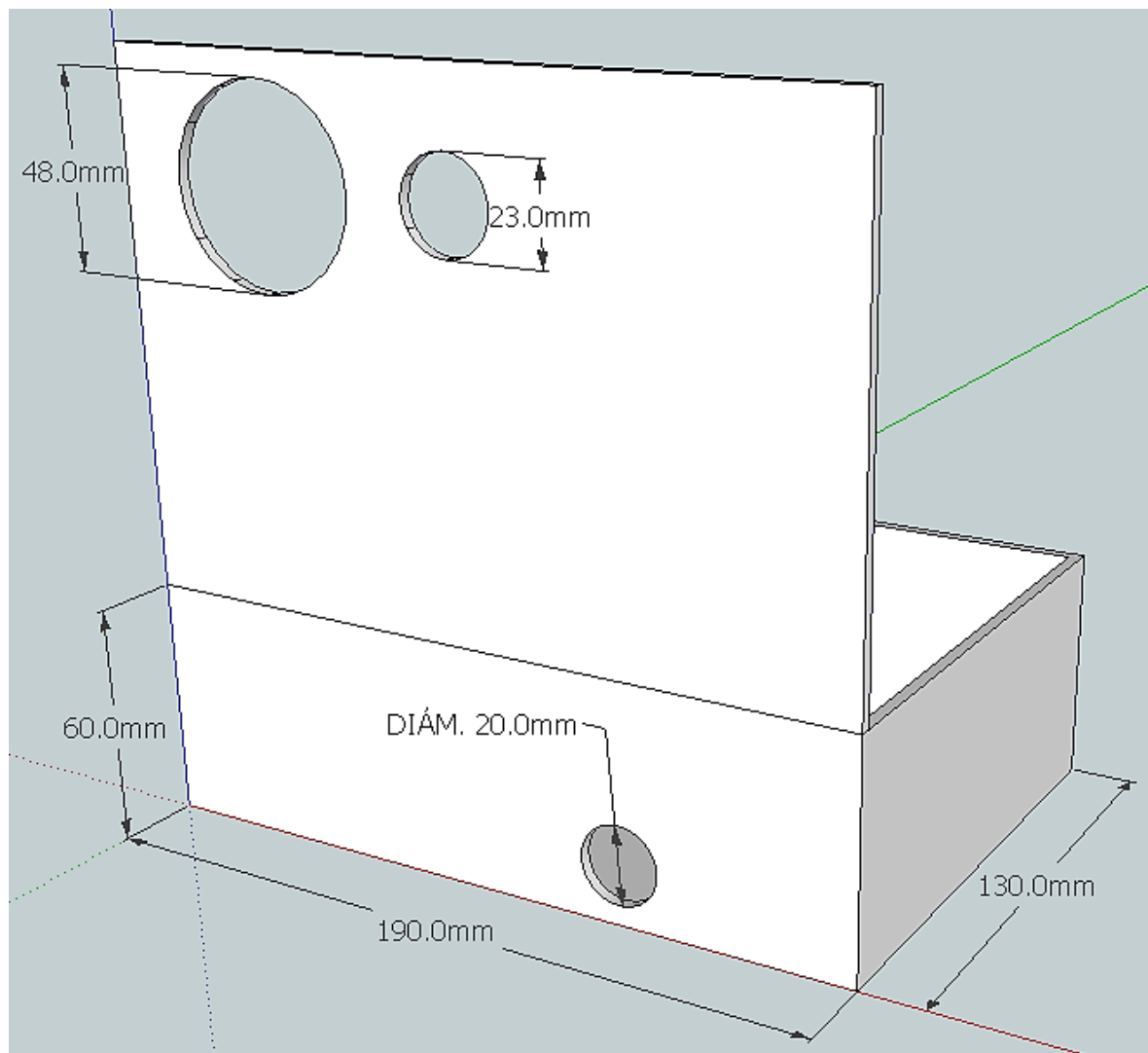


Una vez el dispositivo entra en periodo de escucha se enciende el led verde, indicando que el usuario debe pronunciar uno de los comandos de voz seleccionados previamente. Seguidamente se lleva a cabo el análisis que permite identificar un reconocimiento de voz satisfactorio, mostrando el led azul encendido acompañado de un bip reproducido a través del parlante de 2". En el caso contrario en el que el proceso de reconocimiento de voz no es satisfactorio, se muestra el led rojo encendido durante 150 milisegundos, después de lo cual el módulo de reconocimiento vuelve a entrar en el periodo de escucha.

3.8. CUBIERTA.

Una vez desarrollado el circuito, se diseñó una cubierta en el software SketchUp que permitiese una mayor organización y protección de los componentes a la hora de realizar las pruebas de efectividad.

Figura 26. Diseño de la cubierta.



En este diseño se aprecia una caja rectangular con tres orificios circulares diseñados para acoplar el parlante de 2", el conector XLR hembra de chasis y la inserción de un cable USB que permite monitorear el funcionamiento y la respuesta del dispositivo a través del puerto respectivo del Arduino UNO.

Posteriormente se incluyó en la cubierta el micrófono HORN, los tres LEDS de distinto color y el IRLED emisor.

Figura 27. Cubierta y terminales.



3.9. MEDICIONES DE EFECTIVIDAD.

Teniendo el circuito acoplado a la cubierta, se dio inicio a dos grandes grupos de mediciones de efectividad denominadas prueba estándar y prueba real, las cuales se realizaron bajo condiciones controladas y condiciones aleatorias respectivamente, haciéndose necesaria la implementación de diferentes equipos que permitieran su ejecución.

Adicionalmente se escogieron tres locutores de distintas edades (Mujer 50 años, hombre 23 años y niño de 12 años) que permitiesen llevar a cabo un análisis respecto a la respuesta del módulo de reconocimiento para voces con características distintas, las cuales son mostradas en la sección 3.9.2.3 NORMALIZACIÓN DE NIVELES DE COMANDOS DE VOZ.

3.9.1. EQUIPOS DE MEDICIÓN.

Para esta prueba se tuvieron en cuenta dos monitores de referencia ADAM A8X, empleados para reproducir distintos comandos de voz y simular distintos niveles de ruido de fondo empleando una señal de ruido rosa. Adicionalmente se usó un sonómetro SVANTEK 943A con su respectivo pistófono, un micrófono Shure SM58, un Audiotecnica AT 8015 y el micrófono HORN EM9745P-382 mencionado anteriormente.

3.9.1.1. MONITORES ADAM A8X.

Tabla 10. Características principales del monitor ADAM A8X.

MONITOR ADAM A8X	
	Monitor activo de campo cercano.
	Respuesta en frecuencia de 38Hz-50KHz
	Crossover en 2800Hz.
	Entrada balanceada XLR y no balanceada RCA.
	Control de Ganancia.
	Dieño de caja abierta, Bass Reflex.
	Impedancia de entrada de 30KΩ

Estos monitores se emplearon en la prueba estándar para reproducir simultáneamente una señal de ruido rosa generada en un secuenciador y varios comandos de voz grabados previamente por los locutores involucrados en el estudio.

En todas las mediciones en las que fue necesario su uso, fueron ubicados a 1 metro sobre el nivel del suelo, ubicando el nivel de ganancia en la posición nominal.

3.9.1.2. SONÓMETRO SVANTEK 943A.

Tabla 11. Características principales del sonómetro SVANTEK 943 A.

SVANTEK 943A	
	CARACTERÍSTICAS
	Sonómetro tipo 2.
	Análisis en tiempo real por 1/1 y 1/3 de octava.
	Software de análisis SVAN PC.
	Respuesta en frecuencia de 20Hz - 20000Hz.
	Micrófono electret omnidireccional.
	Ponderación A, C y lineal.
	Autonomía de 8 horas.
	Frecuencia de muestreo de 48 KHz.

Este dispositivo es empleado para realizar diversos análisis respecto al nivel de presión sonora presente en un lugar o recinto por un espacio de tiempo determinado. En este caso se utilizó para medir el nivel de ruido de fondo presente en los recintos donde se realizaron las mediciones de efectividad, además de medir el nivel de presión sonora generado por las fuentes sonoras involucradas.

3.1.9.3. MICRÓFONOS.

Con el fin de estudiar y comparar la efectividad del módulo de reconocimiento de voz, se tomaron en cuenta dos transductores distintos al micrófono HORN, los cuales tuviesen distintos valores de sensibilidad, impedancia y patrón polar.

Tabla 12. Características de micrófonos usados en el estudio.

CARACTERÍSTICAS MICRÓFONOS			
FABRICANTE	HORN	SHURE	AUDIOTECHNICA
	EM9745P-382	SM 58	AT 8015 (batería)
Referencia			
Tipo	Electret	Dinámico	Condensador
Patrón Polar	Omnidireccional	Cardioide	Shotgun
Respuesta en Frecuencia	100 - 20000 Hz	50 - 15000 Hz	40 - 20000Hz
Sensibilidad	-38 dBV (12.58 mV)	-54.5 dBV/Pa (1.85 mV)	-39 dBV/Pa (11.2mV)
Impedancia	2200 Ω	300 Ω	300 Ω

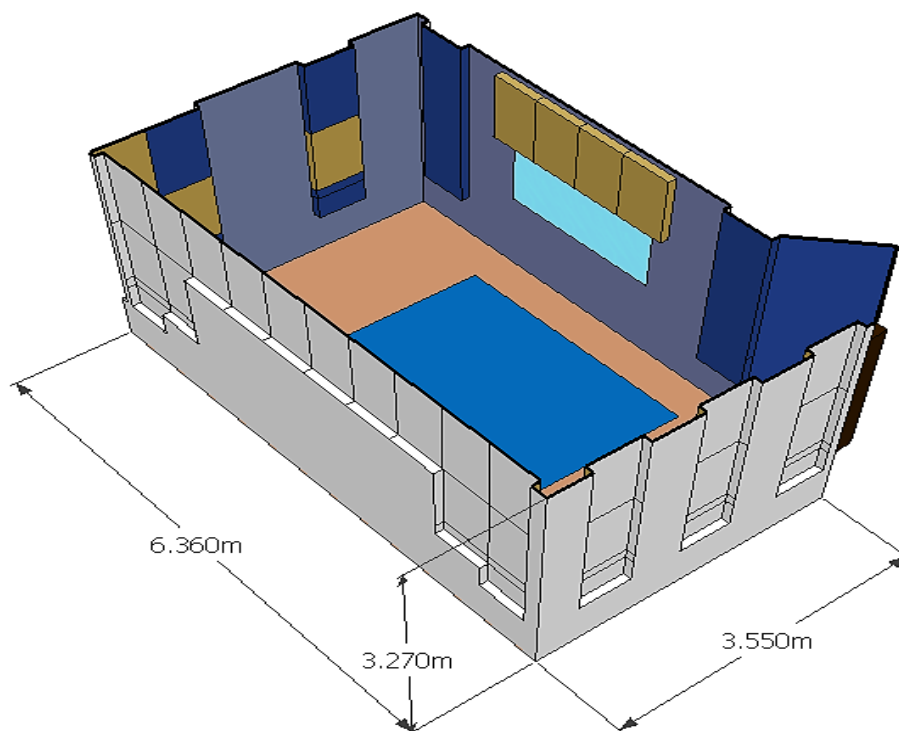
El voltaje de polarización requerido por el micrófono HORN es generado por el módulo de reconocimiento de voz EasyVR al igual que el proceso de preamplificación de la señal, mientras que el phantom power requerido para el funcionamiento del micrófono AT 8015, no se hace necesario debido a que este transductor puede funcionar con una batería AA de 1.5 voltios.

3.9.2. PRUEBA ESTÁNDAR.

Este grupo de pruebas tuvo como objetivo principal verificar que el módulo de reconocimiento de voz tuviera el mismo porcentaje de efectividad para cualquier comando de voz, para lo cual se realizó una grabación previa en donde se registrara a cada uno de los locutores pronunciando los comandos de voz necesarios para tal fin.

Para este caso se escogió la sala del estudio 5.1 de la Universidad de San Buenaventura en Bogotá, donde se llevó a cabo una medición de ruido de fondo con el fin de obtener un punto de partida referente a los niveles de presión sonora a generar en dicho espacio.

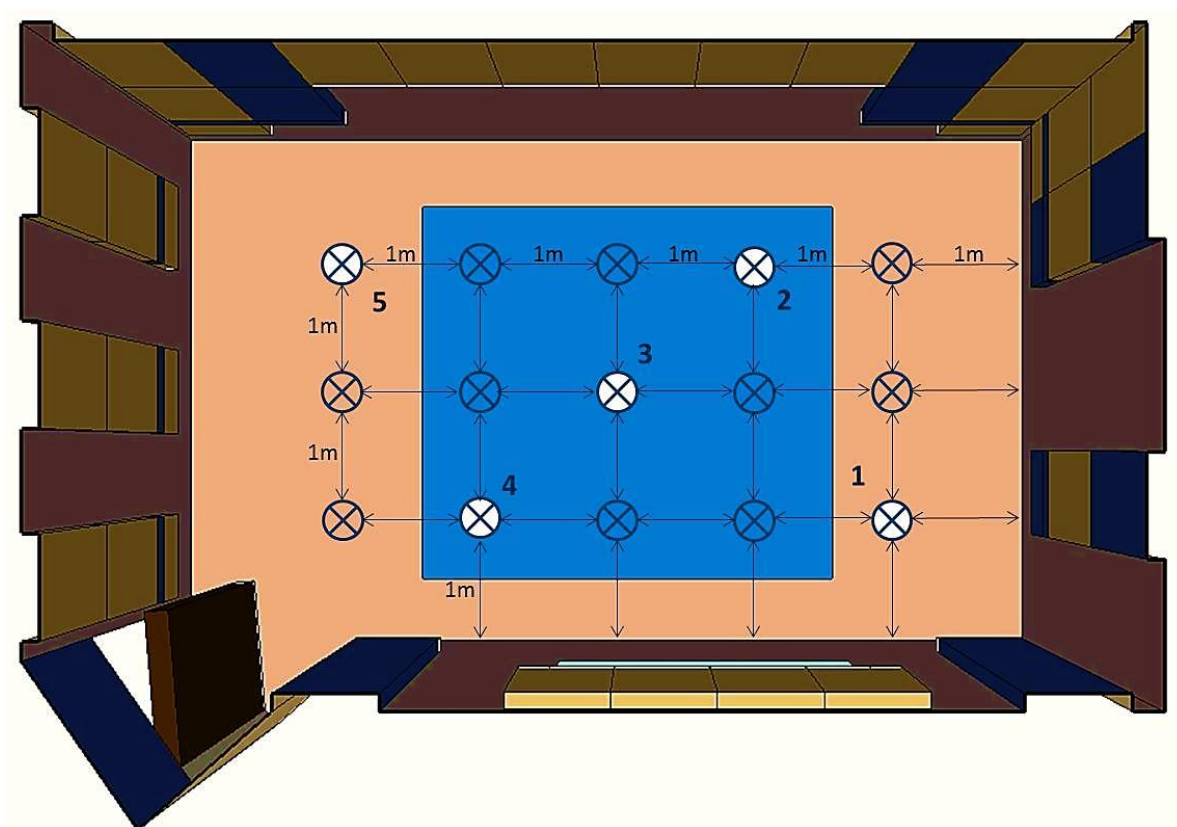
Figura 28. Live Room estudio de grabación 5.1.



3.9.2.1. MEDICIÓN DE RUIDO DE FONDO.

Debido a la no existencia de una norma o resolución que especifique un método general para la medición de ruido de fondo en un recinto cerrado, se llevó a cabo un método alternativo que involucrase parámetros generales de documentos relacionados, tomando 5 posiciones de micrófono separadas a una distancia mínima de 1 metro, realizando 2 repeticiones por posición de micrófono con un tiempo de integración de 5 minutos y un buffer step de 1 segundo, definiendo los 5 puntos de micrófono como se muestra en la siguiente figura.

Figura 29. Puntos de medición de ruido de fondo prueba estándar.



Definidos los puntos de medición, se llevó a cabo la medición de ruido de fondo obteniendo los niveles equivalentes de presión sonora totales y por 1/3 de octava por cada toma realizada. [Ver Anexo B, Medición de ruido de fondo].

Una vez promediados aritméticamente los niveles de presión sonora adquiridos en las mediciones realizadas, se obtuvo la siguiente tabla donde se muestran los niveles de ruido de fondo presente en el recinto.

Tabla 13. Promedios aritméticos medición de ruido de fondo prueba estándar.

NIVELES DE RUIDO DE FONDO					
F [Hz]	Lev[dB]	F [Hz]	Lev[dB]	F [Hz]	Lev[dB]
25	40,016	315	22,123	4000	22,466
31,5	35,096	400	19,345	5000	24,483
40	31,426	500	20,263	6300	23,854
50	33,157	630	23,111	8000	25,933
63	28,230	800	22,080	10000	26,393
80	26,673	1000	23,590	12500	27,263
100	26,478	1250	25,970	16000	27,643
125	24,376	1600	21,862	20000	33,640
160	23,631	2000	19,825	TOT_A	34,313
200	20,766	2500	20,784	TOT_C	40,717
250	19,822	3150	21,487	TOT_Lin	44,284

Después de conocer las características acústicas del ruido de fondo presente en el recinto, se realizaron las mediciones posteriores partiendo de un nivel equivalente total de ruido de fondo de 45dB, valor a partir del cual se simularon diferentes niveles de ruido de fondo, teniendo como fuente un monitor ADAM A8X.

3.9.2.2. NORMALIZACIÓN DE NIVELES DE RUIDO DE FONDO.

Teniendo la necesidad de realizar una prueba de efectividad en donde se utilizaran distintos niveles de ruido de fondo y distintos nivel de presión sonora para comandos de voz reproducidos en los altavoces ADAM A8X, se llevó a cabo una medición que permitiera verificar que al estar a 60 centímetros de los mismos y al variar un valor específico en la escala digital (dBFS), se podría llegar a obtener la misma variación de presión sonora sobre cada uno de los micrófonos a implementar.

Para lo anterior, se utilizó una señal de ruido rosa generada por un secuenciador, en donde su pudiese tener un control exacto en la escala digital, llevando a cabo una medición por cada variación realizada, tomando un tiempo de integración de un minuto con un buffer step de 2 milisegundos, ubicando el sonómetro a 60 centímetros de la fuente. Una vez adquiridos, se resumieron los datos de las mediciones realizadas como se muestra en la siguiente tabla. [Ver Anexo B, Normalización de niveles de ruido de fondo].

Tabla 14. Niveles dBFS vs dBSPL.

NIVEL EN DBFS	NIVEL EQUIVALENTE DE PRESIÓN SONORA GENERADO
0	83.8 dB
-1	82.8 dB
-2	81.8 dB
-3	80.8 dB
-4	79.7 dB
-5	78.8 dB
-6	77.8 dB
-7	76.8 dB
-8	75.8 dB
-9	74.8 dB

Con los resultados de las mediciones realizadas, se pudo verificar que la variación de la presión sonora respecto a la variación digital es proporcional para el montaje mencionado anteriormente, en donde al bajar distintas cantidades de dBFS se obtuvo una variación muy parecida en la escala dBSPL a lo largo del espectro audible.

3.9.2.3. NORMALIZACIÓN DE NIVELES DE COMANDOS DE VOZ.

Teniendo en cuenta que la prueba estándar consistía en estudiar la respuesta del módulo EasyVR bajo condiciones iniciales controladas donde se tuviesen niveles de voz precisos, se realizó un ajuste a los comandos de voz grabados previamente por cada uno de los tres locutores.

Una vez organizados los audios, se llevó a cabo una normalización de los niveles de amplitud digital por medio del plugin nativo de AVID Phase Scope, el cual permite monitorear el valor RMS de una señal y el nivel digital equivalente en ponderación A generado por un espacio de tiempo determinado. Con este proceso se pretendía generar el mismo nivel de presión sonora al reproducir cualquier comando de voz en un monitor ADAM A8X a 60 centímetros del mismo.

Con el fin de comprobar que la normalización fuese efectiva, se reprodujo continuamente el comando “PRENDER” registrado por cada uno de los locutores, hecho que permitiera llevar a cabo una comparación de los niveles de presión sonora generados, realizando una medición de 30 segundos con un buffer step de 2 milisegundos para cada locutor, ubicando el sonómetro a 60 centímetros de la fuente. [Ver Anexo B, Normalización de niveles de comandos de voz].

Una vez adquiridos los datos, se tomaron los niveles equivalentes de cada medición, valores mostrados en la siguiente tabla.

Tabla 15. Nivel de presión sonora generado por los comandos grabados.

LOCUTOR	NIVEL EQUIVALENTE DE PRESIÓN SONORA GENERADO
1	84.5 [dB]
2	84.4 [dB]
3	84.6 [dB]

Una vez normalizados los niveles de los comandos de voz grabados por los locutores, se analizaron los audios generados con el fin de exponer las características generales de las voces de cada uno de los locutores involucrados en el estudio.

Para el estudio mencionado se tuvieron en cuenta los audios normalizados referentes al comando de voz “PRENDER”, cuyo espectro fue analizado a través de diferentes plugins de audio, adquiriendo los datos mostrados en la tabla 17. [Ver Anexo D].

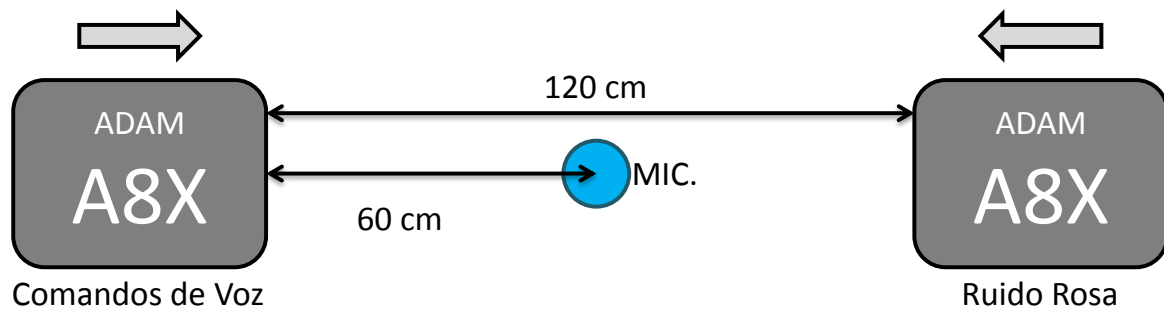
Tabla 17. Características generales de la voz de los locutores al pronunciar el comando PRENDER.

CARACTERÍSTICAS DEL COMANDO DE VOZ PRENDER			
LOCUTOR	FRECUENCIA FUNDAMENTAL [Hz]	FORMANTES REPRESENTATIVAS [Hz]	RANGO DE FRECUENCIAS REPRESENTATIVO [Hz]
LOCUTOR 1 NIÑO	250	500 750 1000	150 - 16000
LOCUTOR 2 MUJER	200	400 600 800 2000 2100	150 - 8000
LOCUTOR 3 HOMBRE	115	360 480 550 1800 1900	80 - 10000

3.9.2.3.1. MONTAJE.

Teniendo los niveles normalizados, se realizó un montaje para realizar las pruebas posteriores donde se mantuvieran las distancias empleadas en la normalización de niveles mencionada anteriormente, ubicando los dos monitores enfrentados y separados a 120 centímetros, colocando los micrófonos involucrados a la mitad de esta distancia, apuntando sus respectivas cápsulas hacia el monitor en donde se reprodujeron los comandos de voz.

Figura 30. Montaje Prueba estándar.



3.9.2.4. ETAPA DE CALIBRACIÓN.

Con el montaje listo, se llevó a cabo la medición del nivel de presión sonora equivalente generado por cada uno de los monitores a 60cm de los mismos, reproduciendo una señal de ruido rosa generada por un secuenciador y un comando de voz reproducido en un ciclo de repetición infinito, teniendo un control de ganancia único, tomando un tiempo de integración de 30 segundos y un buffer step de 2 milisegundos.

Una vez medidos dichos niveles, se realizó el respectivo ajuste en la escala digital que permitiese obtener un nivel de presión sonora de 90 dB al reproducir ruido rosa y 70 dB al reproducir los comandos pregrabados por los locutores involucrados en el estudio.

3.9.2.5. PRUEBA DE NIVELES DE PRESIÓN SONORA EN LA FASE DE ENTRENAMIENTO.

Con los sistemas calibrados, se procedió a obtener los niveles de relación señal ruido óptimos para cada micrófono que permitieran realizar la fase de entrenamiento de los comandos de voz dependientes de locutor satisfactoriamente.

Como se mencionó anteriormente, se implementó un altavoz ADAM A8X, donde se reprodujo una señal de ruido rosa cuyo nivel de presión sonora variase de 90dB a 45dB con pasos de -5dB, y un segundo altavoz ADAM A8X donde se reprodujeron los comandos de voz “PRENDER” a un nivel de presión sonora de 70dB, los cuales fueron grabados por cada uno de los locutores, pretendiendo omitir las diferencias de pronunciación de una persona al enunciar una misma palabra.

Con el montaje listo, se llevaron a cabo las mediciones de efectividad que permitieran obtener los niveles de voz y de ruido de fondo para los cuales su pudiera llevar a cabo el entrenamiento de los comandos dependientes. [Ver Anexo B, Prueba de niveles en la fase de entrenamiento].

Posteriormente se resumieron los resultados adquiridos, mostrando en la tabla 17 los valores para los cuales la fase de entrenamiento de comandos dependientes fue satisfactoria.

Tabla 16. Niveles de presión sonora óptimos para la fase de entrenamiento.

NIVELES DE PRESIÓN SONORA ÓPTIMOS FASE DE ENTRENAMIENTO			RELACIÓN SEÑAL
MICRÓFONO	R. FONDO	NIVEL VOZ	RUIDO [dB]
HORN	45	70	25
	50	70	20
SM 58	45	70	25
	50	70	20
	55	70	15
	60	70	10
	65	70	5
	70	70	0
	75	70	-5
AT 8015	45	70	25
	50	70	20
	55	70	15
	60	70	10
	65	70	5

Con estos datos fue posible deducir que la diferencia entre las voces de los locutores no afecta significativamente en la etapa de entrenamiento de comandos dependientes de locutor, observando que se mantienen los mismos valores de relación señal ruido óptimos al variar de locutor sin importar el micrófono empleado.

Notando en la tabla 18 que los valores obtenidos no son regulares para los tres transductores, se tomaron los valores mínimos de relación señal/ruido para llevar a cabo las pruebas posteriores.

Tabla 17. Valores representativos de relación señal/ruido en fase de entrenamiento.

VALORES REPRESENTATIVOS			
MICRÓFONO	R. S/R. [dB]	NIVEL R. FONDO [dB]	NIVEL VOZ [dB]
HORN	20	50	70
SM58	-5	75	70
AT8015	5	65	70

Con estos datos se obtuvo un punto de partida para evidenciar si los valores representativos de relación señal/ruido se mantenían para diferentes niveles, pretendiendo evidenciar si la fase de entrenamiento de comandos de voz seguía siendo satisfactoria al cambiar los valores de ruido de fondo y de comandos de voz sin variar los niveles de relación señal/ruido seleccionados.

Una vez realizada la medición mencionada, se obtuvieron los niveles para los cuales la fase de entrenamiento fue satisfactoria. [Ver Anexo B, Variación de niveles de presión sonora con relación señal ruido constante].

Tabla 18. Variación de niveles de presión sonora con relación señal ruido constante.

NIVELES DE PRESIÓN SONORA ÓPTIMOS PARA FASE DE ENTRENAMIENTO			RELACIÓN SEÑAL RUIDO [dB]
MICRÓFONO	R. FONDO [dB]	NIVEL VOZ [dB]	
HORN	45	65	20
	50	70	20
	55	75	20
SM 58	75	70	-5
	80	75	-5
AT 8015	55	60	5
	60	65	5
	65	70	5

Adquiridos los niveles de fuentes aptos para realizar una fase de entrenamiento satisfactoria con la relación señal ruido mínima que permitía cada micrófono, se dio paso al análisis de diferentes comandos dependientes e independientes de locutor, teniendo en cuenta los diferentes niveles de dificultad de reconocimiento que permite configurar el módulo de reconocimiento de voz EasyVR.

3.9.2.6. PRUEBA DE NIVELES DE DIFICULTAD PARA COMANDOS DEPENDIENTES.

Para esta medición se tomaron los niveles de fuentes más altos para cada micrófono que mantuviesen el mismo nivel de relación señal/ruido adquirido en las mediciones anteriores, valores con los cuales se llevó a cabo la fase de entrenamiento de cuatro comandos dependientes con su respectiva medición de efectividad por niveles de dificultad, los cuales hacen referencia a una comparación entre el comando entrenado y el comando pronunciado, teniendo el nivel 5 como el proceso en donde dicha comparación es más estricta.

En esta prueba se realizaron 10 repeticiones por medición con el fin de obtener el número de veces en las que los comandos fuesen reconocidos satisfactoriamente (mostradas en verde), entrenados y reproducidos con los mismos audios bajo las mismos niveles de presión sonora con los que se realizaron las respectivas fases de entrenamiento.

Tabla 19. Medición por niveles de dificultad con micrófono HORN.

MEDICIÓN EFECTIVIDAD POR NIVELES DE DIFICULTAD.									
MIC.	RUIDO DE FONDO [dB]	NIVEL VOZ [dB]	LOCUTOR	COMANDO	NIVEL DIFICULTAD				
					1	2	3	4	5
HORN	55	75	1	PRENDER	10	10	10	10	10
				REPRODUCIR	10	10	10	10	10
				POWER	10	10	10	10	10
				STOP	10	10	10	10	10
			2	PRENDER	10	10	10	10	10
				REPRODUCIR	10	10	10	10	10
				POWER	10	10	10	10	10
				STOP	10	10	10	10	10
			3	PRENDER	10	10	10	10	10
				REPRODUCIR	10	10	10	10	10
				POWER	10	10	10	10	10
				STOP	10	10	10	10	10

Tabla 20. Medición por Niveles de Dificultad con micrófono SM 58.

MEDICIÓN EFECTIVIDAD POR NIVELES DE DIFICULTAD.									
MIC.	RUIDO DE FONDO [dB]	NIVEL VOZ [dB]	LOCUTOR	COMANDO	NIVEL DIFICULTAD				
					1	2	3	4	5
SM58	80	75	1	PRENDER	10	10	10	10	10
				REPRODUCIR	10	10	10	10	10
				POWER	10	10	10	10	10
				STOP	10	10	10	10	10
			2	PRENDER	10	10	10	10	10
				REPRODUCIR	10	10	10	10	10
				POWER	10	10	10	10	10
				STOP	10	10	10	10	10
			3	PRENDER	10	10	10	10	10
				REPRODUCIR	10	10	10	10	10
				POWER	10	10	10	10	10
				STOP	10	10	10	10	10

Tabla 21. Medición por niveles de dificultad con micrófono AT8015.

MEDICIÓN EFECTIVIDAD POR NIVELES DE DIFICULTAD.									
MIC.	RUIDO DE FONDO [dB]	NIVEL VOZ [dB]	LOCUTOR	COMANDO	NIVEL DIFICULTAD				
					1	2	3	4	5
AT8015	65	70	1	PRENDER	10	10	10	10	10
				REPRODUCIR	10	10	10	10	10
				POWER	10	10	10	10	10
				STOP	10	10	10	10	10
			2	PRENDER	10	10	10	10	10
				REPRODUCIR	10	10	10	10	10
				POWER	10	10	10	10	10
				STOP	10	10	10	10	10
			3	PRENDER	10	10	10	10	10
				REPRODUCIR	10	10	10	10	10
				POWER	10	10	10	10	10
				STOP	10	10	10	10	10

Una vez comprobado que bajo las mismas condiciones iniciales el módulo de reconocimiento de voz posee un 100% de efectividad para todos los niveles de dificultad al pronunciar cualquiera de los cuatro comandos reproducidos, se llevaron a cabo diferentes fases de entrenamiento para cada locutor manteniendo los valores de relación señal/ruido anteriores, en donde posteriormente se fue variando el nivel de presión sonora de ruido de fondo presente en el recinto, esto con el fin de evidenciar el margen de niveles de presión sonora para el cual dicho porcentaje de efectividad no cambiase. [Ver Anexo B, Variación de niveles dificultad con niveles de presión sonora distintos a la fase entrenamiento].

Resumidos los datos adquiridos en esta medición, se tomaron los diferentes rangos de niveles para los cuales el módulo de reconocimiento tuvo un porcentaje de efectividad del 100% al variar las condiciones iniciales de ruido de fondo presente en la fase de entrenamiento.

Tabla 22. Variación de niveles dificultad con niveles de presión sonora distintos a la fase entrenamiento.

NIVELES DE PRUEBA COMANDOS DEPENDIENTES				
MICRÓFONO	NIVELES DE ENTRENAMIENTO		NIVELES DE PRUEBA	
	RUIDO DE FONDO [dB]	NIVEL DE VOZ [dB]	RUIDO DE FONDO [dB]	NIVEL DE VOZ [dB]
HORN	45	65	45 - 65	65
	50	70	45 - 65	70
	55	75	45 - 65	75
SM 58	75	70	45 - 85	70
	80	75	45 - 90	75
AT 8015	55	60	45 - 70	60
	60	65	45 - 75	65
	65	70	45 - 80	70

En la tabla 24 se muestra una columna de color verde donde aparecen los niveles de ruido de fondo con los que se realizó la fase de entrenamiento, y una columna de color azul que expone los niveles de ruido de fondo con los que se pudo llevar a cabo una fase de reconocimiento de comandos de voz satisfactoria para cada combinación de niveles realizada.

Obtenidos los datos de efectividad del módulo de reconocimiento de voz para comandos dependientes de locutor, se llevó a cabo un análisis similar para encontrar los niveles de presión sonora aptos para el reconocimiento de voz de comandos independientes de locutor, para los cuales no existe una fase de entrenamiento previa.

3.9.2.7. PRUEBA DE NIVELES DE DIFICULTAD PARA COMANDOS INDEPENDIENTES.

Inicialmente se realizó una prueba que permitiera verificar que el reconocimiento de los comandos de voz independientes presentes en el wordset 3 tuvieran el mismo porcentaje de efectividad para los tres locutores, para lo cual se reprodujeron los 11 comandos involucrados a un nivel de presión sonora de 70 dB una distancia de 60 centímetros del transductor de entrada respectivo, realizando 10 repeticiones por cada comando con el fin de obtener el número de veces en el que el proceso de reconocimiento fue satisfactorio para cada locutor, tomando en cuenta únicamente el nivel 1 de dificultad del módulo EasyVR.

Tabla 23. Comparación de efectividad para comandos independientes.

COMPARACIÓN DE EFECTIVIDAD DE COMANDOS INDEPENDIENTES			
COMANDO	LOCUTOR 1	LOCUTOR 2	LOCUTOR 3
CERO	10	10	10
UNO	9	10	8
DOS	10	10	10
TRES	10	10	9
CUATRO	10	10	10
CINCO	9	9	8
SEIS	10	4	9
SIETE	8	9	9
OCHO	3	5	2
NUEVE	10	10	10
DIEZ	9	4	2

Observando la efectividad no es constante para los comandos independientes en el nivel 1 de dificultad de reconocimiento, se realizó una medición de efectividad más estricta, teniendo en cuenta los 5 niveles de dificultad reproduciendo únicamente el comando “CERO”, suponiendo los mismos resultados para los comandos independientes que presentaron un 100% de efectividad. Para esto se tomaron valores de ruido de fondo de 45 dB a 90 dB y niveles de voz entre 55 dB y 90dB, prueba cuyos resultados se exponen en la tabla 26, mostrando los niveles para los cuales el reconocimiento de voz del comando independiente de locutor “CERO” fue satisfactorio. [Ver Anexo B, Niveles de prueba para comandos independientes].

Tabla 24. Niveles de prueba para comandos independientes.

NIVELES DE PRUEBA PARA COMANDOS INDEPENDIENTES							
MICRÓFONO	NIVELES DE PRUEBA		NIVELES DE DIFICULTAD				
	RUIDO DE FONDO [dB]	NIVEL DE VOZ [dB]	1	2	3	4	5
HORN	45 - 60	60	10	10	10	10	10
	45 - 65	65	10	10	10	10	10
	45 - 70	70	10	10	10	10	10
	45 - 75	75	10	10	10	10	10
	45 - 80	80	10	10	10	10	10
	45 - 85	85	10	10	10	10	10
	45 - 90	90	10	10	10	10	10
SM 58	45 - 85	75	10	10	10	10	10
	45 - 90	80	10	10	10	10	10
	45 - 95	85	10	10	10	10	10
	45 - 95	90	10	10	10	10	10
AT 8015	45 - 65	60	10	10	10	10	10
	45 - 70	65	10	10	10	10	10
	45 - 75	70	10	10	10	10	10
	45 - 80	75	10	10	10	10	10
	45 - 85	80	10	10	10	10	10
	45 - 90	85	10	10	10	10	10
	45 - 95	90	10	10	10	10	10

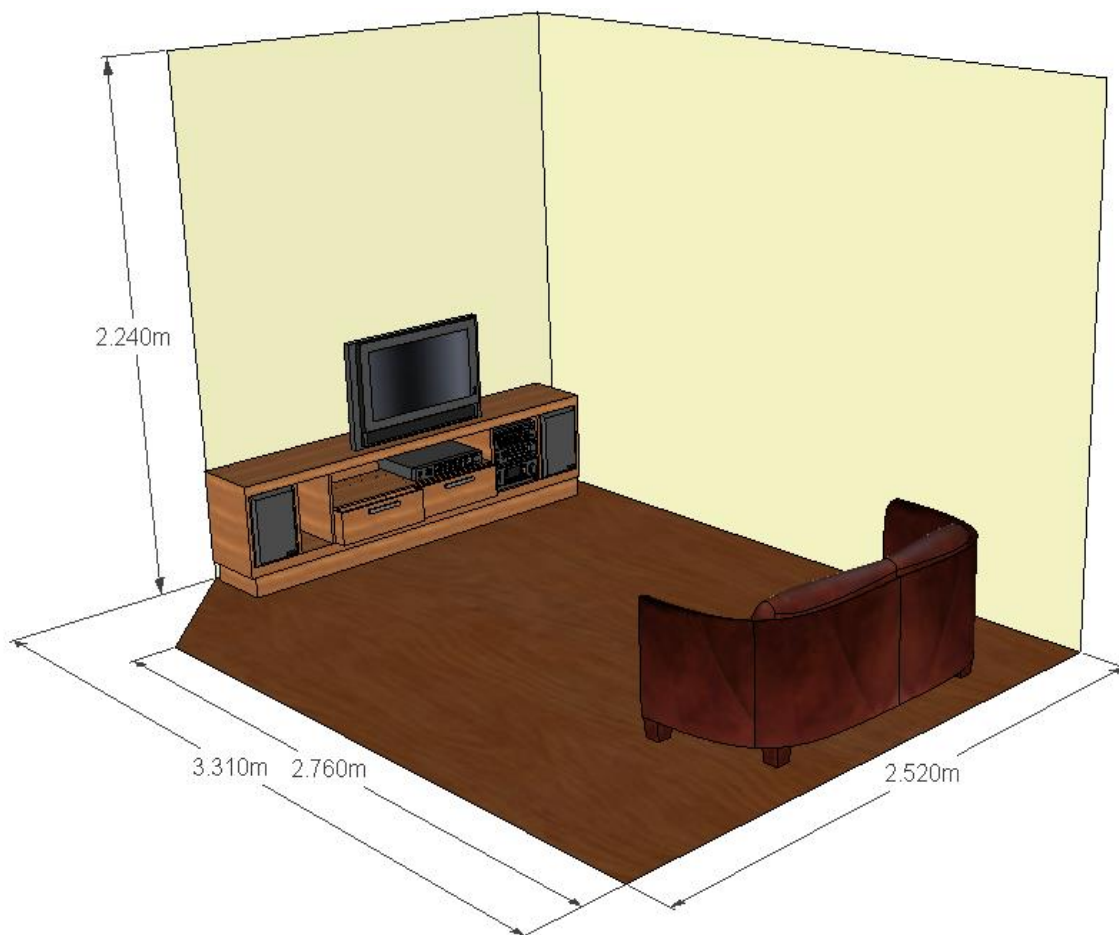
En la tabla 26 se muestra una columna azul que expone los niveles de presión sonora de ruido de fondo para los cuales se pudo realizar el proceso de reconocimiento de manera satisfactoria para cada nivel de presión sonora de comandos de voz.

Con la obtención y organización de estos datos, se culminó con la prueba estándar, permitiendo un posterior análisis respecto a la respuesta del módulo EasyVR frente a los cambios de locutor, micrófono y comandos de voz bajo condiciones aleatorias.

3.9.3. PRUEBA REAL.

Este grupo de mediciones se basó en el cambio de las condiciones iniciales de la prueba estándar a la hora de estudiar la respuesta del módulo EasyVR. Para este caso se realizaron las mediciones en una sala de estar real, con comandos de voz pronunciados por los mismos locutores, teniendo niveles de voz aleatorios y niveles de ruido de fondo variables.

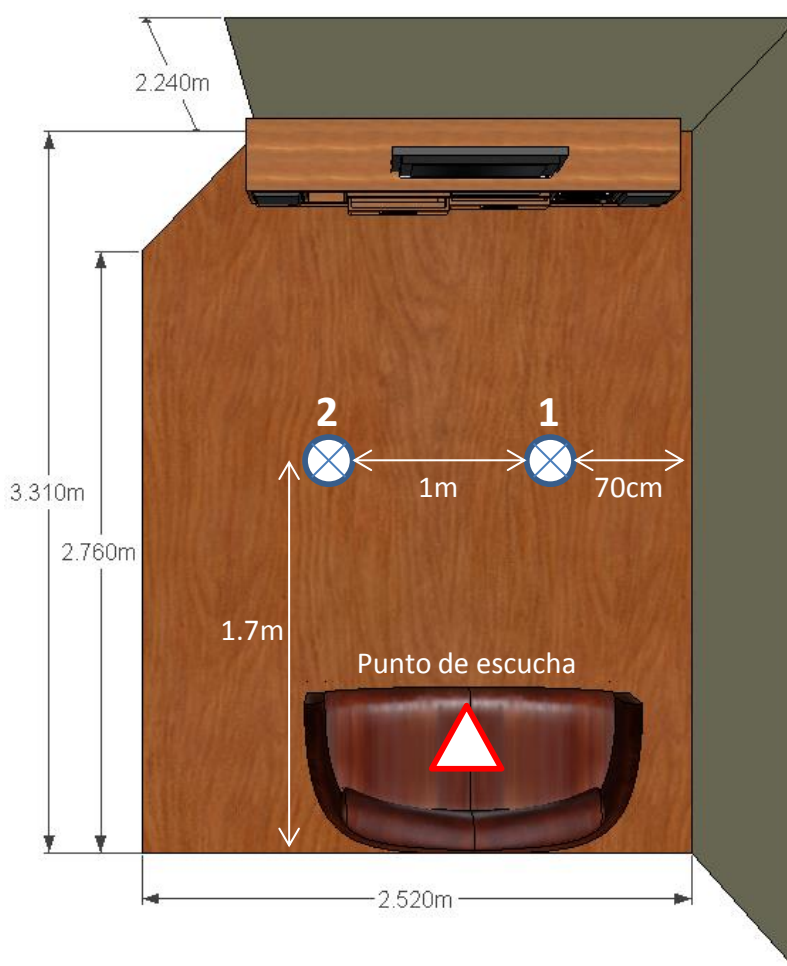
Figura 31. Recinto prueba real (sala de estar).



3.9.3.1. MEDICIÓN DE RUIDO DE FONDO.

Para llevar a cabo la medición de ruido de fondo para esta prueba se tomó en cuenta el método empleado en la medición de ruido de la prueba estándar mencionado en la sección 3.9.2.1. En este caso se tomaron en cuenta 2 posiciones de micrófono separadas a 1 metro de distancia, en donde se realizaron tres repeticiones por punto, tomando un tiempo de integración de 5 minutos y un buffer step de 500 milisegundos.

Figura 32. Puntos de medición de ruido de fondo prueba real.



Una vez promediados aritméticamente los valores adquiridos en esta medición, se generó la siguiente tabla, mostrando el nivel de presión sonora equivalente y por 1/3 de octava. [Ver Anexo C, Medición de ruido de fondo].

Tabla 25. Promedios aritméticos de mediciones ruido de fondo prueba real.

NIVELES PROMEDIADOS					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	55,446	315	41,476	4000	29,413
31,5	47,669	400	38,968	5000	27,636
40	46,371	500	47,141	6300	23,534
50	43,976	630	43,401	8000	24,499
63	37,631	800	36,729	10000	19,513
80	41,728	1000	37,782	12500	21,095
100	39,713	1250	35,319	16000	14,300
125	40,678	1600	34,810	20000	33,941
160	39,868	2000	33,391	TOT_A	48,418
200	39,843	2500	31,301	TOT_C	55,879
250	41,621	3150	29,714	TOT_Lin	58,669

3.9.3.2. MONTAJE.

Con el fin de mantener la comodidad del usuario al ver una película en su centro de entretenimiento, se realizaron las mediciones posteriores ubicando los micrófonos a 60 centímetros de la boca de los locutores, realizando el montaje mostrado en la siguiente figura.

Figura 33. Montaje prueba real.



El montaje se basó en la posición original en el que un usuario ubicaría su centro de entretenimiento, el cual se acomoda a las condiciones propias de cada espacio. En este caso se tenían los dos altavoces del equipo de sonido a la altura del suelo, separados entre sí a una distancia de 90 centímetros, posición distanciada a 2.3 metros de la posición de escucha del usuario. Finalmente se fueron ubicando los micrófonos en el brazo del sofá ubicado en la sala de estar, posición referenciada en la figura anterior.

3.9.3.3. MEDICIÓN DE NIVELES DE PRESIÓN SONOAR GENERADOS POR EL EQUIPO DE SONIDO.

Teniendo en cuenta que se venía desarrollando un modo de operación en donde funcionaran en conjunto los tres sistemas de entretenimiento, se conectaron las salidas de audio del DVD a las entradas auxiliares RCA del equipo de sonido. Adicionalmente, se escogió una película al azar de la cual se seleccionó uno de los capítulos como referencia para realizar las mediciones de efectividad. Ubicando el sonómetro en la posición de escucha mostrada en la figura 30, se obtuvieron las variaciones de niveles de presión sonora generados por el equipo de sonido durante 5 minutos, teniendo un buffer step de 200 milisegundos, variando el nivel de ganancia entre los valores 2 y 10, llegando a adquirir 1500 muestras por medición. [Ver Anexo C, Medición de niveles de presión sonora generados por el equipo de sonido].

En la siguiente tabla se muestra el rango de niveles generados por el equipo de sonido por cada nivel de ganancia estudiado, mostrando el nivel de presión sonora con mayor porcentaje de muestras capturadas.

Tabla 26. Niveles de presión sonora generados por el equipo de sonido.

NIVELES DE PRESIÓN SONORA GENERADOS POR EL EQUIPO DE SONIDO		
NIVEL DE GANANCIA EQUIPO DE SONIDO	NIVELES DE PRESIÓN SONORA [dB]	NIVEL DE PRESIÓN SONORA CON MAYOR PORCENTAJE [dB]
NIVEL 2	51 - 73	57
NIVEL 3	54 - 80	58
NIVEL 4	54 - 88	62
NIVEL 5	55 - 91	63
NIVEL 6	53 - 92	63
NIVEL 7	55 - 96	68
NIVEL 8	55 - 98	72
NIVEL 9	55 - 100	73
NIVEL 10	57 - 101	77

3.9.3.4. NIVELES DE VOZ DE LOCUTORES.

Teniendo en cuenta que los niveles de presión sonora generados por una persona al hablar son dependientes de variables como el ruido de fondo, el estado de ánimo y de salud, no se estableció un proceso de normalización para niveles de presión sonora de comandos de voz, por lo cual se le pidió a cada locutor que hablara al nivel que creyeran necesario para que los comandos pronunciados fuesen reconocidos satisfactoriamente.

3.9.3.5. PRUEBA DE COMANDOS DEPENDIENTES.

Teniendo en cuenta que no se había encontrado una utilidad real para los niveles de dificultad de reconocimiento de comandos de voz, se llevó a cabo una medición que permitiese evidenciar que los comandos entrenados realmente fuesen dependientes de locutor, para lo cual cada locutor entrenó el comando “PRENDER” pretendiendo evidenciar si los dos locutores restantes podían generar un proceso de reconocimiento satisfactorio al pronunciar el mismo comando, teniendo en cuenta los 5 niveles de dificultad.

Tabla 27. Comprobación de niveles de dificultad para comandos dependientes.

COMPROBACIÓN NIVELES DE DIFICULTAD PARA COMANDOS DEPENDIENTES							
MICRÓFONO	LOCUTOR ENTRENAMIENTO	LOCUTOR DE PRUEBA	NIVELES DE DIFICULTAD				
			NIVEL 1	NIVEL 2	NIVEL 3	NIVEL 4	NIVEL 5
HORN	1	1	10	10	10	10	10
		2	10	10	7	1	0
		3	10	10	8	1	0
SM 58	2	1	8	5	0	0	0
		2	10	10	10	10	10
		3	9	6	4	0	0
AT 8015	3	1	8	8	0	0	0
		2	10	10	9	6	4
		3	10	10	10	10	10

Observando la tabla 29 se pudo concluir que al tener un proceso de reconocimiento de comandos de voz dependientes de locutor para el nivel 1 de dificultad, se obtiene un proceso similar al reconocimiento de comandos independientes, teniendo como ventaja la posibilidad de que cualquier persona generase un proceso de reconocimiento satisfactorio para cualquier comando de voz sin tener

que acceder de los comandos predefinidos mostrados en la tabla 25, descartándolos definitivamente para tener un control de forma independiente de locutor.

Observando que la voz del locutor no influyó en el reconocimiento de comandos de voz en la prueba estándar, se realizó esta última prueba de efectividad con un solo locutor, probando con los niveles 1 y 5 de dificultad al momento de realizar pruebas con los niveles de ganancia del equipo medidos anteriormente.

Una vez resumidos los datos adquiridos en la medición, se generaron las siguientes tablas. [Ver Anexo C, Medición de efectividad prueba real].

Tabla 28. Datos de efectividad por nivel de ganancia del equipo de sonido con nivel 1 de dificultad.

NIVEL 1						
MICRÓFONO	COMANDO DE VOZ	NIVELES DE GANANCIA				
		2	3	4	5	6
HORN	PRENDER	10	10	8	0	0
	APAGAR	10	10	9	0	0
	PLAY	10	10	4	0	0
	PAUSA	10	10	8	0	0
SM58	PRENDER	10	0	0	0	0
	APAGAR	10	0	0	0	0
	PLAY	10	0	0	0	0
	PAUSA	9	0	0	0	0
AT 8015	PRENDER	10	10	10	0	0
	APAGAR	10	10	10	0	0
	PLAY	10	10	6	0	0
	PAUSA	10	10	9	0	0

Tabla 29. Datos de efectividad por nivel de ganancia del equipo de sonido con nivel 5 de dificultad.

NIVEL 5						
MICRÓFONO	COMANDO DE VOZ	NIVELES DE GANANCIA				
		2	3	4	5	6
HORN	PRENDER	10	10	4	0	0
	APAGAR	10	10	6	0	0
	PLAY	10	10	2	0	0
	PAUSA	10	10	0	0	0
SM58	PRENDER	8	0	0	0	0
	APAGAR	6	0	0	0	0
	PLAY	9	0	0	0	0
	PAUSA	6	0	0	0	0
AT 8015	PRENDER	10	10	8	0	0
	APAGAR	10	10	8	0	0
	PLAY	10	8	6	0	0
	PAUSA	10	10	5	0	0

Observando en las tablas 30 y 31 que la efectividad del módulo de reconocimiento de voz solo presentó valores importantes de efectividad hasta el nivel 3 de ganancia del equipo de sonido, se llevó a cabo una nueva medición en donde se acortó a 15 centímetros la distancia entre los micrófonos y la boca del locutor involucrado, teniendo en cuenta únicamente el nivel 5 de dificultad de reconocimiento.

Tabla 30. Datos de efectividad en prueba a 15 centímetros.

EFECTIVIDAD A 15 CM											
LOCUTOR	MIC	COMANDO	NIVEL DEL EQUIPO DE SONIDO								
			2	3	4	5	6	7	8	9	10
1	HORN	PRENDER	10	10	9	9	3	1	0	0	1
	SM58		10	10	10	10	10	10	10	10	10
	AT8015		10	10	10	10	8	7	8	5	3
2	HORN	APAGAR	10	10	10	9	8	5	5	2	1
	SM58		10	10	10	10	10	10	10	10	10
	AT8015		10	10	10	10	10	8	7	7	5
3	HORN	PLAY	10	10	9	7	7	7	4	0	0
	SM58		10	10	10	10	10	10	10	10	10
	AT8015		10	10	10	9	10	9	4	5	2

Notando en los datos presentes en la tabla 32 una gran mejoría en los niveles de efectividad, se llevó a cabo el análisis de datos que permitiese optimizar el funcionamiento del dispositivo controlador respecto al rango de niveles de presión sonora que permiten realizar un reconocimiento de voz satisfactorio a la hora de ver una película en un centro de entretenimiento casero.

4. ANÁLISIS DE DATOS.

Con el análisis de los datos de las mediciones realizadas en la prueba estándar y en la prueba real, se da paso al desarrollo del código de programación en el que se implementen los comandos de voz con mayor efectividad y las optimizaciones referentes a los niveles de presión sonora soportados al momento de entrenar y usar el del dispositivo controlador.

4.1. PRUEBA ESTÁNDAR.

Iniciando con el análisis de los datos adquiridos en la prueba estándar, es necesario aclarar que los niveles de presión sonora de ruido de fondo simulados a lo largo de cada medición individual fueron constantes, motivo por el cual los niveles de presión sonora adquiridos solo sirven de referencia al momento de utilizar el dispositivo controlador en un ambiente en donde el ruido de fondo no presente mayores variaciones en el nivel de presión sonora ni en su espectro.

Dando paso al análisis de los datos adquiridos en las mediciones referentes a la fase de entrenamiento de comandos de voz dependientes de locutor mostrados en la sección 3.9.2.5, se pudo observar claramente que los niveles aptos para llevar a cabo una fase de entrenamiento satisfactoria para el comando de voz “PRENDER”, no se ven afectados por las diferencias presentes en las voces de los tres locutores involucrados en el estudio, permitiendo deducir que el módulo de reconocimiento de voz EasyVR responde de la misma manera para cualquier voz.

También se puede observar que el micrófono Shure SM58 permite llevar a cabo una fase de entrenamiento satisfactoria con niveles de presión sonora de ruido de fondo más altos respecto al micrófono HORN y al AT 8015, llegando a admitir que el nivel de la voz esté hasta 5 dB por debajo del nivel del ruido de fondo, teniendo una relación señal ruido de -5 dB.

Teniendo los valores mínimos de relación señal/ruido óptimos para cada micrófono, se llevó a cabo una prueba posterior en donde se variaron los niveles de ruido de fondo y de comandos de voz, sin alterar los valores de relación señal ruido mencionados. Con esto se adquirió una serie de niveles de presión sonora para los cuales se podía realizar una fase de entrenamiento de comandos de voz dependientes de manera satisfactoria, datos mostrados en la tabla 19.

Posteriormente en la sección 3.9.2.6 PRUEBA DE NIVELES DE DIFICULTAD PARA COMANDOS DEPENDIENTES, se realizó una prueba que permitiera verificar si existían comandos dependientes para los cuales el módulo de voz presentara un porcentaje de efectividad mayor respecto a otros, para lo cual se entrenaron y reprodujeron sucesivamente los comandos “PRENDER”, “REPRODUCIR”, “POWER” y “STOP”, evidenciando que para todos los comandos se presentaba el mismo porcentaje de efectividad sin importar el locutor, el micrófono empleado, ni el nivel de dificultad de reconocimiento de comandos dependientes que permite programar el módulo EasyVR, datos mostrados en las tablas 20, 21 y 22. Con estos valores se pudo deducir que se puede llevar a cabo una fase de entrenamiento y de reconocimiento satisfactoria para cualquier comando de voz pronunciado por cualquier persona.

El paso a seguir fue realizar una prueba en donde se tuvieran condiciones de reproducción de comandos distintas a la fase de entrenamiento, para lo cual se utilizaron los mismos comandos de voz mencionados anteriormente, pretendiendo

adquirir el rango de niveles de presión sonora que permitiera llevar a cabo una fase de reconocimiento de voz con condiciones distintas a la fase de entrenamiento, datos presentes en la tabla 23.

Una vez adquiridos las condiciones óptimas para llevar a cabo una fase de entrenamiento y reconocimiento de comandos de voz dependientes de locutor, se llevó a cabo una serie de pruebas similares para obtener las condiciones óptimas para el reconocimiento de comandos de voz independientes de locutor.

Inicialmente en la sección 3.9.2.7 se realizó una prueba que permitiese verificar que el porcentaje de efectividad de reconocimiento para comandos independientes fuera regular para los comandos predefinidos en el wordset 3 mostrados en la figura 19, obteniendo los datos presentes en la tabla 24 en donde se evidencia que los comandos para los que se obtuvo un porcentaje mayor fueron “CERO”, “DOS”, “CUATRO” y “NUEVE”.

Seguidamente se realizó una medición exclusiva para el comando “CERO” en donde se tuvieron en cuenta los 3 locutores, los 3 micrófonos y los 5 niveles de dificultad de reconocimiento para comandos independientes de voz que permite asignar el módulo EasyVR. El objetivo fue adquirir las condiciones óptimas referentes a los niveles de ruido de fondo y de comandos de voz que permitieran llevar a cabo un reconocimiento satisfactorio, datos mostrados en la figura 25.

Con esto se culminó con la prueba estándar, la cual permitió adquirir los niveles de presión sonora de ruido de fondo y de comandos de voz que permiten realizar un entrenamiento y reconocimiento de comandos de voz satisfactorio.

4.2. PRUEBA REAL.

Con el fin de comprobar los datos adquiridos en la prueba estándar, se realizó un estudio similar en una sala de estar que permitiera tener unas condiciones iniciales de ruido de fondo variable y aleatorio al momento de ver una película en un centro de entretenimiento que incluyera un televisor, un equipo de sonido y un DVD de la marca SONY.

Después de llevar a cabo la medición de ruido de fondo respectiva, se midieron los niveles que generaba una película cuyo audio fuese reproducido a través del equipo de sonido en mención, teniendo en cuenta los niveles de ganancia del mismo dispositivo. Posteriormente en la tabla 27 presente en la sección 3.9.3.3. se organizaron los datos adquiridos, mostrando el rango de niveles de presión sonora generados por cada nivel de ganancia estudiado, con el fin de obtener los niveles de ruido de fondo a los que sería sometido el módulo EasyVR en esta prueba real.

Ya que en las pruebas realizadas hasta el momento no se había encontrado una utilidad para los niveles de dificultad de reconocimiento de comandos de voz, se llevó a cabo una prueba con los mismos tres locutores pronunciando el comando “PRENDER”, en donde uno de ellos realizaría la fase de entrenamiento y los dos restantes la fase de reconocimiento. Con los datos organizados en la tabla 28 de la sección 3.9.3.5 se pudo evidenciar que al implementar el nivel 1 de dificultad de reconocimiento de comandos de voz dependientes, cualquier locutor podía generar una fase de reconocimiento satisfactoria, permitiendo llegar a usar esta configuración como reconocimiento de comandos independientes, en donde se podría evitar el uso de los comandos predefinidos en los wordset presentes en el software EasyVR Commander. Adicionalmente se evidenció que en el nivel 5 de

dificultad de reconocimiento de comandos de voz dependientes de locutor se obtienen los mejores resultados para comandos de esta índole.

Por tal razón se llevó a cabo una prueba en donde se ubicaron los micrófonos a 60 centímetros de la boca de los locutores, reproduciendo una película con la configuración de audio mencionada anteriormente, teniendo en cuenta los niveles de ganancia del equipo de sonido, y los niveles de dificultad 1 y 5 para el reconocimiento de comandos voz dependientes. Los datos adquiridos se muestran en las tablas 29 y 30, donde solo se tuvo en cuenta a un locutor.

Observando que en esta prueba la efectividad del proceso de reconocimiento de comandos de voz no fue totalmente satisfactoria a partir del nivel de ganancia 4 del equipo de sonido, se realizó una prueba similar en donde únicamente se cambió a 15 centímetros la distancia entre la boca del locutor y la cápsula de los micrófonos usados. En la tabla 31 se muestran los datos adquiridos, donde se observa una gran mejoría del proceso de reconocimiento, llegando a tener procesos satisfactorios hasta el nivel 10 de ganancia del equipo de sonido usando el micrófono Shure SM58.

5. OPTIMIZACIÓN DEL DISPOSITIVO CONTROLADOR.

Inicialmente se llevó a cabo la programación del código final cuyo funcionamiento varió muy poco respecto al diagrama de flujo mostrado en la figura 23, en donde se utilizó el código base referenciado en la sección 3.7 para el cual se hicieron dos versiones distintas. La primera versión implementa el nivel 1 de dificultad de reconocimiento de voz de comandos dependientes, usado para tener la posibilidad de usar el dispositivo controlador con comandos independientes configurables por el usuario. La segunda versión implementa el mismo código variando únicamente en la implementación del nivel 5 de dificultad, permitiendo tener un control dependiente del locutor.

Como modificaciones iniciales se eliminaron el parlante y dos de los tres leds de colores usados en las pruebas de efectividad, resultando en un proceso diferente a la hora de interactuar con el usuario, partiendo del hecho de que cuando el dispositivo controlador entra en periodo de escucha el led restante permanece encendido, en dado caso en que el proceso de reconocimiento sea satisfactorio, el led se apaga hasta realizar el envío de información respectivo a los sistemas de entretenimiento. En el caso contrario, si el proceso de reconocimiento no resulta exitoso el led parpadea cuatro veces, después de lo cual el dispositivo entra nuevamente en periodo de escucha.

Viendo que durante la reproducción de una película, en el centro de entretenimiento analizado, no se obtienen valores de efectividad del 100% para todos los micrófonos al involucrar los distintos niveles de ganancia del equipo de sonido, se han desarrollado los códigos mencionados con el fin de controlar la mayor cantidad de funciones de los sistemas de entretenimiento casero antes de

iniciar la reproducción de la película, controlando los sistemas de entretenimiento de la forma mencionada en la siguiente tabla.

Tabla 31. Funciones controladas por el dispositivo controlador.

COMANDO	FUNCIONES CONTROLADAS		
	TV	E. DE SONIDO	DVD
Prender	Prender	Prender	Prender
	Mute	Volumen -	Eject
	x	Volumen +	x
Play	x	x	Play
Menú	x	x	Menú
Volumen	x	Volumen +	x
Bajar	x	Volumen -	x
Pausa	x	x	Pausa
Apagar	Apagar	Apagar	Eject
	x	x	Apagar

El dispositivo controlador se reprogramó para que al pronunciar la palabra “PRENDER”, se enviase la información necesaria para encender los tres sistemas de entretenimiento, donde adicionalmente envía el comando respectivo para mutear la salida de audio del televisor y para abrir la bandeja del DVD. Mientras se realiza este proceso, también se controla el nivel de ganancia del equipo de sonido para situarlo en la posición 7.

Posteriormente el usuario debe pronunciar el comando “PLAY”, ordenando al DVD cerrar la bandeja e iniciar la reproducción de la película, en donde el usuario puede controlar el volumen de reproducción, el acceso al menú y la función pausa, haciendo uso de los comandos respectivos. En el momento en el que el usuario desee apagar los sistemas debe pronunciar el comando “APAGAR”,

desencadenando en la apertura de la bandeja del DVD donde posteriormente se procede al apagado de los sistemas de entretenimiento. [Ver Anexo A, Código final].

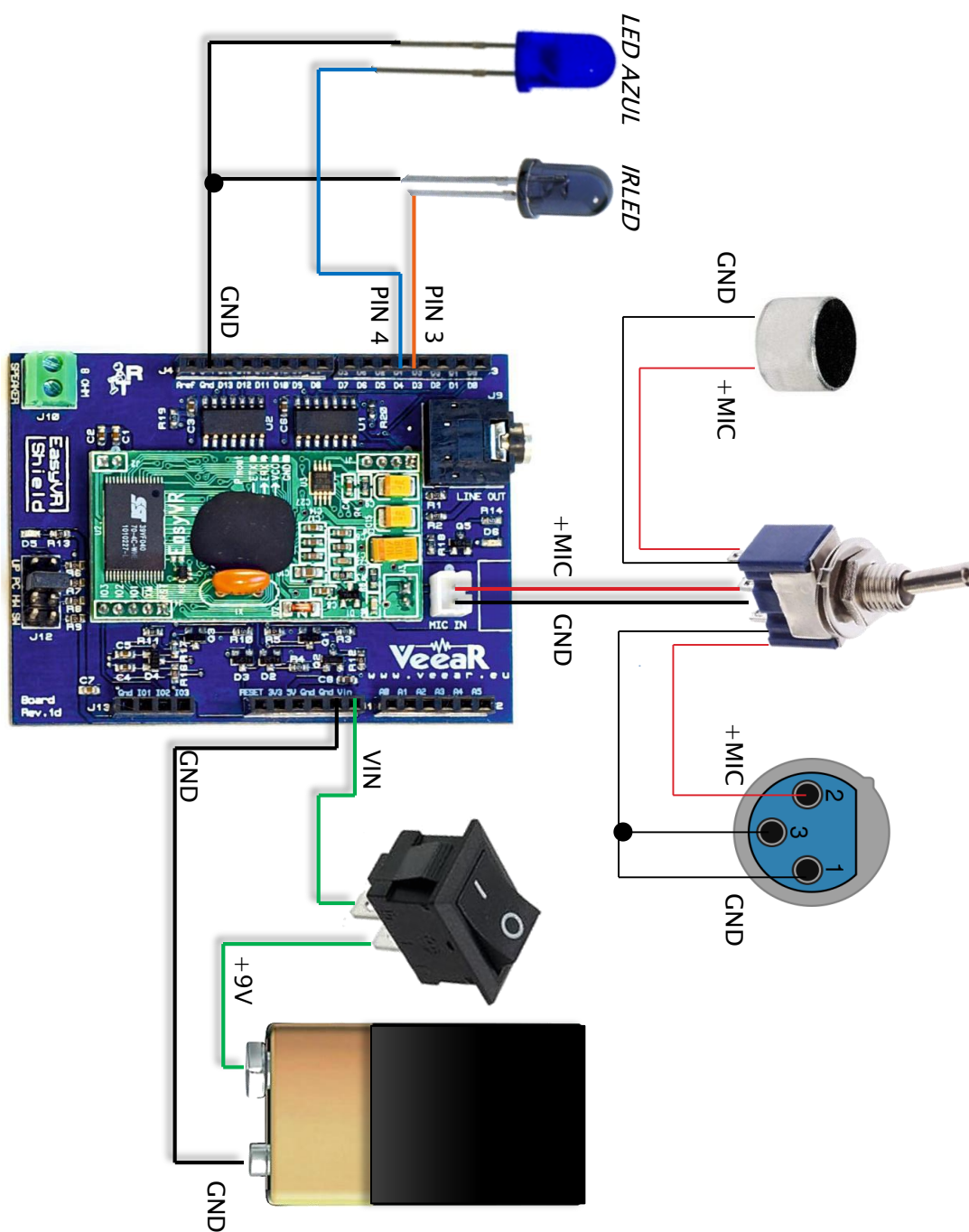
El diagrama de flujo presente en la figura 34 muestra el funcionamiento del código referenciando.

Figura 34. Diagrama de flujo Código Final.



Finalmente, en el circuito mostrado en el esquemático de la figura 22 se montó en una nueva cubierta mostrada en la figura 37 donde se implementó un switch de codillo de dos posiciones que permitiese la selección de la entrada de micrófono a utilizar. Adicionalmente se incluyó una batería de 9 voltios y un switch que controla el encendido y apagado del dispositivo controlador, mostrado en el esquemático mostrado en la figura 36.

Figura 35. Esquemático Final.



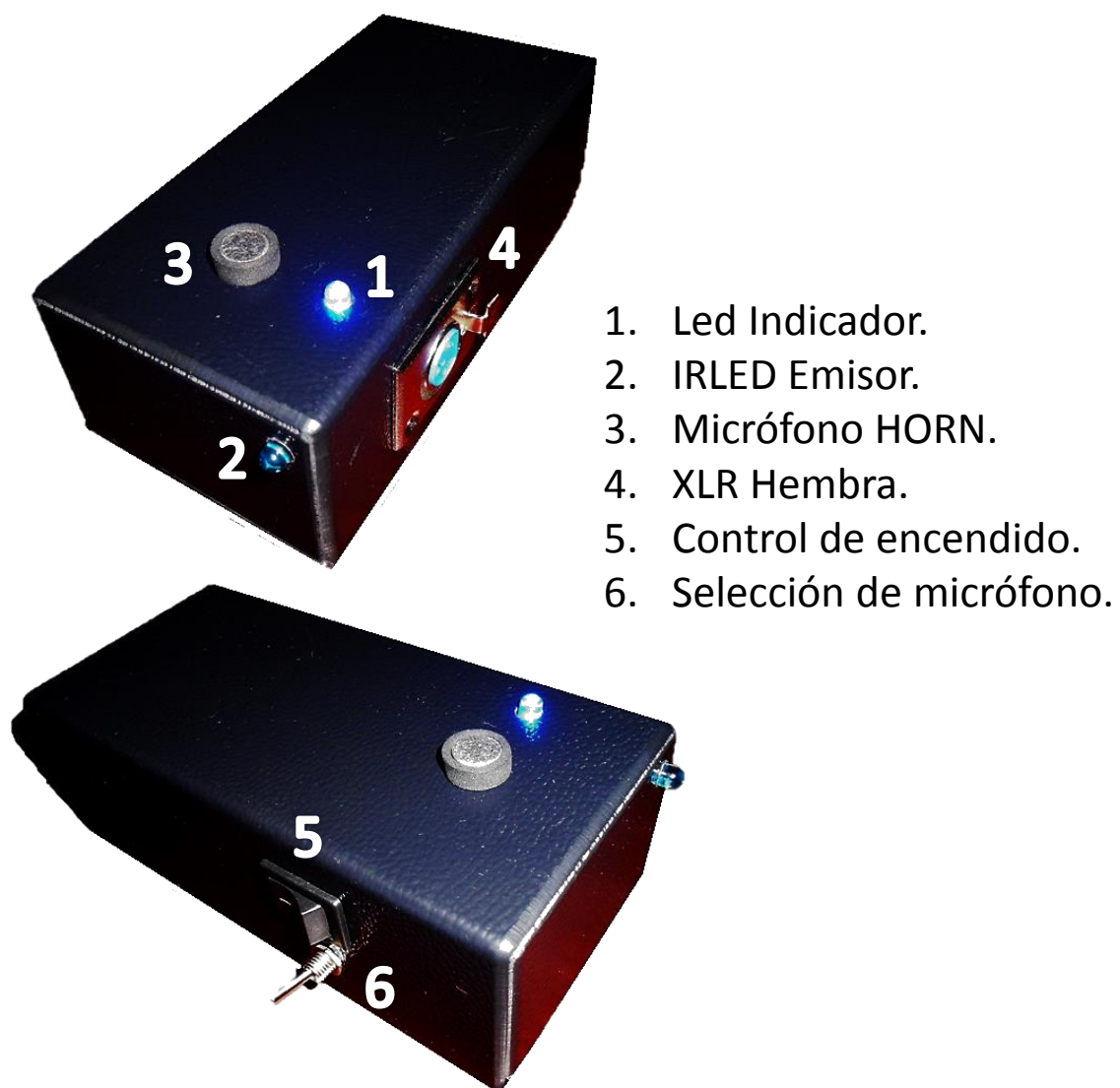
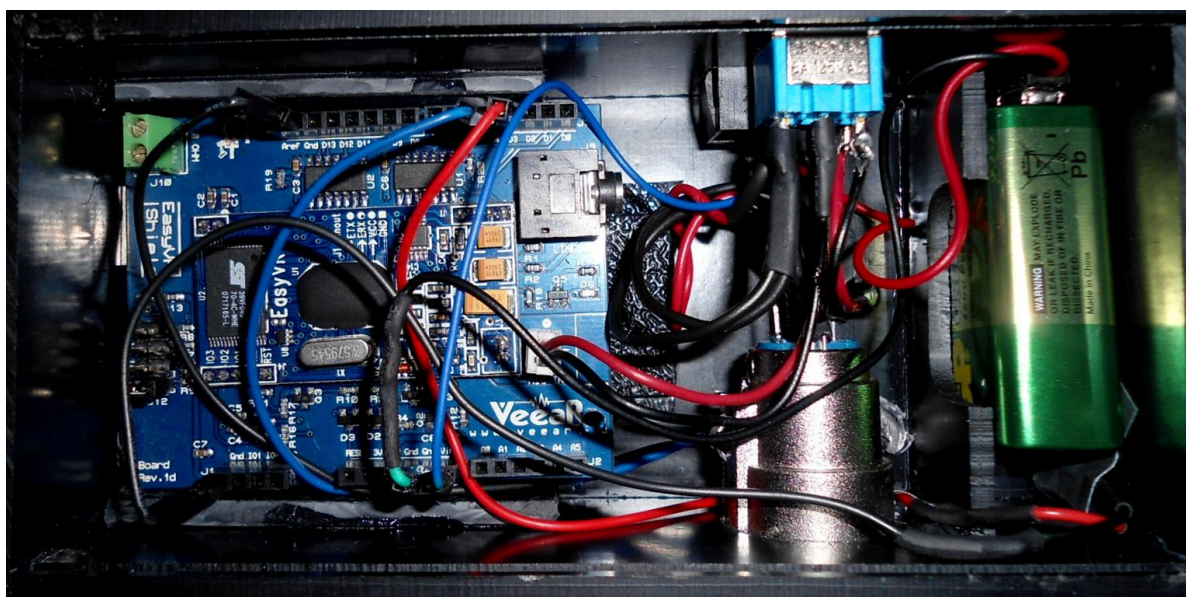


Figura 37. Parte interior del Dispositivo Controlador.



6. LIMITACIONES.

Una vez probado el dispositivo controlador se encontraron diferentes limitaciones para su respectivo funcionamiento, donde se menciona como afectan las distintas variables estudiadas respecto al proceso de reconocimiento de comandos de voz.

- **Comandos similares:**

A medida que fueron realizando las diferentes ases de entrenamiento para distintos comandos de voz, se observa que el software EasyVR Commander NO permite el entrenamiento de comandos similares en un mismo grupo (Ejemplo: Prender – Aprender).

- **Fase de entrenamiento de comandos de voz:**

A continuación se muestran las condiciones referentes a los niveles de presión sonoras que permiten llevar a cabo una fase de entrenamiento de comandos de voz satisfactoria.

Tabla 32.Limitaciones de funcionamiento.

LIMITACIONES FASE DE ENTRENAMIENTO		
MICRÓFONO	RELACIÓN SEÑAL RUIDO MÍNIMA [dB]	NIVEL DE RUIDO DE FONDO MÁXIMO [dB]
HORN	20	55
SHURE SM58	-5	80
AT8015	5	65

- **Fase de reconocimiento de comandos de voz:**

A continuación se exponen los niveles de presión sonora que permiten llevar a cabo una fase de reconocimiento de comandos de voz de manera satisfactoria, teniendo como punto de referencia los niveles de presión sonora presentes en el punto donde se ubique al micrófono utilizado.

Dado el caso en que el recinto en donde se vaya a usar el dispositivo controlador no presente variación alguna respecto al nivel de ruido de fondo, se tienen las siguientes limitaciones.

- **Micrófono HORN:** al momento de usar el dispositivo controlador teniendo como transductor de entrada al micrófono HORN, se puede llevar a cabo una fase de reconocimiento satisfactoria únicamente si el nivel de presión sonora de ruido de fondo no supera los 65 dB ni el nivel de presión sonora del comando de voz pronunciado.
- **Micrófono Shure SM58:** al momento de usar el dispositivo controlador teniendo como transductor de entrada al micrófono Shure SM58, se puede llevar a cabo una fase de reconocimiento satisfactoria siempre y cuando el nivel de presión sonora del ruido de fondo no supere al nivel de los comandos de voz por 15dB.
- **Micrófono AT 8015:** al momento de usar el dispositivo controlador teniendo como transductor de entrada al micrófono AT 8015, se puede llevar a cabo una fase de reconocimiento satisfactoria siempre y cuando el nivel de presión sonora del ruido de fondo no supere al nivel de los comandos de voz por 10dB.

Adicionalmente, en el caso en el que se use el dispositivo controlador en un espacio cuyo nivel de presión sonora presente variaciones significativas en el nivel de presión sonora, se tiene como limitante al transductor de entrada y la distancia entre el mismo y la boca del locutor, permitiendo llevar a cabo un proceso de reconocimiento de voz satisfactorio al usar exclusivamente un micrófono de baja sensibilidad de patrón polar cardiode, ubicado a 15 centímetros de la boca del locutor.

7. RECOMENDACIONES DE USO.

Al momento de realizar las fases de entrenamiento y de reconocimiento de comandos de voz, se recomiendan tener en cuenta las siguientes condiciones.

- Realizar la fase de entrenamiento de comandos de voz ubicando el dispositivo controlador en su posición de uso final.
- Ubicar el dispositivo controlador apuntando el IRled transmisor hacia los sistemas de entretenimiento.
- Pronunciar el comando de voz cuando el led azul del dispositivo controlador permanezca encendido.
- Considerando que la tecnología implementada para la comunicación entre los mandos a distancia y los sistemas de entretenimiento está desarrollada con el fin de controlar un solo dispositivo a la vez, se recomienda ubicar los sistemas de entretenimiento lo más cerca posible.
- Teniendo en cuenta que los sistemas de entretenimiento casero actuales poseen distintos modos de operación (CD, Tunner, USB, etc.) y varias entradas de audio y video seleccionables, carentes de mensajes infrarrojos estándar definidos, se deben seleccionar manualmente las entradas y los modos de operación pertinentes en los distintos sistemas de entretenimiento.

8. RECOMENDACIONES PARA TRABAJOS A FUTURO.

Inicialmente, considerando que en la web no existen mayores ejemplos respecto a la programación del módulo de reconocimiento de voz EasyVR, se recomienda tomar como ejemplo los códigos realizados para la realización de este proyecto, expuestos en el Anexo A, teniendo en cuenta que existen diferentes funciones y sentencias para comandos dependientes e independientes de locutor

Es de vital importancia revisar el manual de operación del módulo mencionado, en donde se puede encontrar el listado completo de funciones y sentencias referentes a la programación del mismo que permiten aprovechar al máximo el dispositivo, considerando que para este proyecto no se implementaron en su totalidad.

En cuanto al sistema de emisión de mensajes infrarrojos, es importante recordar que la librería IRremote.h utilizada en el control de los pulsos luminosos del IRLed transmisor soporta varios protocolos de comunicación vía infrarrojo, permitiendo llegar a controlar simultáneamente sistemas de entretenimiento de distintos fabricantes. Dado el caso en que se haga necesaria la implementación de un protocolo no soportado por la librería mencionada, es posible desarrollar un código de programación que permita controlar los pulsos luminosos del IRLed transmisor por medio de las funciones y pines del Arduino UNO destinados a trabajar con modulación por ancho de pulso.

9. CONCLUSIONES.

- Según las mediciones de efectividad realizadas se concluye que las fases de entrenamiento y reconocimiento de comandos de voz dependientes de locutor realizadas por el módulo EasyVR, no se ven afectadas de ninguna manera al variar el locutor, el idioma o el comando de voz, evidenciando los mismos resultados para las pruebas realizadas con los tres locutores involucrados en el estudio.
- Una vez realizado el análisis de los datos adquiridos para cada micrófono, en la tabla 34 se puede observar que los procesos involucrados en el reconocimiento de comandos voz se ven directamente afectados por la directividad y la sensibilidad del transductor utilizado, obteniendo los mejores resultados con un Shure SM58, micrófono de baja sensibilidad con patrón polar cardioide.
- Se evidenció que al programar el reconocimiento de comandos dependientes de locutor en el nivel 1 de dificultad, se obtienen resultados similares a un proceso de reconocimiento de comandos de voz independientes del locutor, permitiendo usar cualquier comando que el usuario desee en esta configuración.
- Se observa que los comandos de voz independientes de locutor predefinidos en el módulo EasyVR no mantienen los mismos niveles de efectividad, presenciando mejores resultados para los comandos “CERO”, “DOS”, “CUATRO” y “NUEVE”. Teniendo cuenta que se desconoce el algoritmo desarrollado por el fabricante, se reemplazó este proceso por el método que reconoce comandos dependientes de locutor en el nivel 1 de dificultad, el cual

permite llevar a cabo un proceso satisfactorio de reconocimiento de comandos de voz independientes de locutor para cualquier palabra entrenada.

- El reconocimiento de comandos independientes por parte del módulo EasyVR, permite unos niveles de ruido de fondo mayores respecto a los comandos dependientes, sin que esto signifique un porcentaje de reconocimiento mayor.
- Para obtener una efectividad mayor al momento de controlar distintos sistemas de entretenimiento por medio de comandos de voz, se debe usar un micrófono cardioide de baja sensibilidad, ubicado a 15 centímetros de la boca del usuario.
- Es posible evidenciar que al momento de controlar simultáneamente los tres sistemas de entretenimiento casero a través del dispositivo desarrollado, se optimiza el proceso habitual de control sobre los mismos, permitiéndole al usuario manipular distintas de funciones a partir de la pronunciación de un solo comando de voz.
- Comparando los resultados obtenidos en la prueba estándar y en la prueba real, se observa que el proceso de reconocimiento de comandos de voz es más efectivo al tener un nivel de ruido de fondo constante.
- En la tablas 31 se observa claramente que el rango de niveles de presión sonora referentes al nivel de ruido de fondo presente al momento de realizar un proceso de reconocimiento de voz, se mejora o aumenta significativamente al reemplazar el micrófono HORN por los micrófonos Shure SM 58 y AT8015, notando que se puede realizar un proceso de reconocimiento de comandos de voz satisfactorio teniendo un nivel de presión sonora de ruido de fondo máximo de 80 dB que puede estar hasta 5dB por encima del nivel de presión sonora generado al pronunciar un comando de voz.

10. BIBLIOGRAFÍA.

- [1]. KEMBLE, Kimberlee, "An Introduction to Speech Recognition". {En línea}. Disponible en: (http://vidcat.org/user_data/28ab9a896e1e95ed3914cb462f36b1aabd272693.pdf).
- [2]. B.H. Juang & R. Lawrence, "Automatic Speech Recognition – A Brief History of the Technology Development". {En línea}. Disponible en: (http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/354_LALI-ASRHistory-final-10-8.pdf).
- [3]. OROPEZA J. Luis & SUAREZ Sergio, "Algoritmos y Métodos para el Reconocimiento de Voz en Español". {En línea}. Disponible en: (<http://www.scielo.org.mx/pdf/cys/v9n3/v9n3a7.pdf>).
- [4]. KIRRIEMUIR, John, "Speech Recognition Technologies". {En línea}. Disponible en: (http://www.jisc.ac.uk/uploaded_documents/tsw_03-03.pdf).
- [5]. MEJÍA, Francisco, "Reconocimiento de Voz Aplicado a la Domótica". {En línea}. Disponible en: (<http://bibdigital.epn.edu.ec/bitstream/15000/4914/1/RECONOCIMIENTO%20DE%20VOZ%20APLICADO.pdf>).
- [6]. MARTÍNEZ, Javier, "Control Domótico por Voz". {En línea}. Disponible en: (<http://riunet.upv.es/bitstream/handle/10251/17631/Memoria.pdf?sequence=1>).
- [7]. FALCONI Luis & JIMENEZ Carlos. "Estudio e Implementación de Domótica Activado por Comandos de Voz y Comunicación en Zigbee". {En línea}. Disponible en: (<http://dspace.esPOCH.edu.ec/bitstream/123456789/165/1/38T00156.pdf>).
- [8]. PORTELO F. & MARTÍNEZ M, "Radicación Infrarroja". {En línea}. Disponible en: (http://www.sld.cu/galerias/pdf/sitios/rehabilitacion-fis/radiacion_infrarroja.pdf).
- [9]. DE LOPE, Javier, "Emisión y Recepción de Infrarrojos". {En línea}. Disponible en: (<http://www.dia.fi.upm.es/~jdllope/docs/delopez01a.pdf>).

- [10] INNOTECH SYSTEMS INCORPORATED, "Press Release (Accenda Remote Control)". {En línea}. Disponible en: (<http://www.innotechsystems.com/accenda/guide.htm>).
- [11]. INNOTECH SYSTEMS INCORPORATED, "Accenda User's Guide". {En línea}. Disponible en: (<http://www.innotechsystems.com/accenda/guide.htm>).
- [12]. "Sony SIRC Infrared Protocol". {En línea}. Disponible en: (<http://www.cypress.com/?docID=46755>).
- [13]. VEEAR, "EasyVR User Manual Release 3". {En línea}. Disponible en: (<http://www.robotshop.com/media/files/pdf/easyvr-user-manual-vrb-001.pdf>).
- [14]. "VRBot Guía Rápida". {En línea}. Disponible en: (<http://www.msebilbao.com/notas/downloads/VRbot%20Guia%20rapida.pdf>).
- [15]. MIYARA, Federico, "La Voz Humana". {En línea}. Disponible en: (<http://www.fceia.unr.edu.ar/prodivoz/fonatorio.pdf>).
- [16]. BARTÍ DOMINGO, Robert, Título: Acústica Medioambiental Vol. I. Editorial Club Universitario, 2010, 134p.
- [17]. BELTRAN, Diana & OLAYA, Cesar & IVAN, Cepeda. Control Electrónico de voz para una silla de ruedas. Bogotá, 2003, 195p. Trabajo de grado (Ingeniería Electrónica). Universidad de San Buenaventura. Facultad de Ingeniería.
- [18]. TORRES, Daniel & ANGARITA, Edwin. Diseño e implementación de un sistema de iluminación domótico automatizado por comandos de voz. Bogotá, 2012, 81p. Trabajo de grado (Ingeniería de Sonido). Universidad de San Buenaventura. Facultad de Ingeniería.
- [19]. BECERRA, Nicolai & Rodríguez Simón. Diseño e implementación de un control lumínico por medio de la voz. Bogotá 2008, 106p. Trabajo de grado (Ingeniería de Sonido). Universidad de San Buenaventura. Facultad de Ingeniería.

[20]. INNOTECH SYSTEMS INCORPORATED, "Accenda Remote Control". {En línea}. {Agosto 2013}. Disponible en: (<http://www.innotechsystems.com/accenda/>).

[21]. INNOTECH SYSTEMS INCORPORATED, "Surfboard Remote Control". {En línea}. {Agosto 2013}. Disponible en: (<http://www.innotech.com/surfboard/>).

[22]. BABYLON 10, the world's best online dictionary. "Radiación Infrarroja". {En línea}. {Agosto 2013}. Disponible en: (<http://diccionario.babylon.com/infrarrojo?&tl=/>).

[23]. SBPROJECTS, Modulation. "SIRC Protocol". {En línea}. {Agosto 2013}. Disponible en: (<http://www.sbprojects.com/knowledge/ir/sirc.php>).

[24]. SBPROJECTS, Protocol. "SIRC Protocol". {En línea}. {Agosto 2013}. Disponible en: (<http://www.sbprojects.com/knowledge/ir/sirc.php>).

[25]. VEEAR, EasyVR. {En línea}. {Agosto 2013}. Disponible en: (<http://www.veear.eu/products/easyvr/>).

[26]. ZAGROS ROBOTICS, EasyVR. "EasyVR Speech Recognition Shield v2.0". {En línea}. {Agosto 2013}. Disponible en: (<http://www.zagrosrobotics.com/shop/item.aspx?itemid=865>).

[27].ROBOTSHOP, EasyVR User Manual. "Technical specifications, Physical dimensions and pin assignment". {En línea}. {Agosto 2013}. Disponible en: (<http://www.robotshop.com/media/files/pdf/easyvr-user-manual-vrb-001.pdf>).

[28]. OPNE STAX CNX, Modeling the Speech Signal. "Vocal Tract". {En línea}. {Febrero 2014}. Disponible en: (<http://www.robotshop.com/media/files/pdf/easyvr-user-manual-vrb-001.pdf>).

[29]. GRAMATICAS.NET, Fonemas – Definición y ejemplos. "Lista de Fonemas en Español". {En línea}. {Febrero 2014}.
Disponible en: (<http://www.gramaticas.net/2011/05/fonemas-definicion-y-ejemplos.html>)

[30]. WIKIMEDIA COMMONS, "Pink noise spectrum". {En línea}. {Febrero 2014}.
Disponible en: (http://commons.wikimedia.org/wiki/File:Pink_noise_spectrum.png).

ANEXO A. CÓDIGOS DE PROGRAMACIÓN.

Pág.

1.	Código 1. Comprobación de mensajes infrarrojos.....	121
2.	Código 2. Reconocimiento de comandos independientes de locutor.....	122
3.	Código 3. Reconocimiento de comandos dependientes de locutor.....	125
4.	Código 4. Código final.....	129

1. Código 1. Comprobación mensajes infrarrojos.

```
#include <IRremote.h>
IRsend irsend;
void setup()
{
    Serial.begin(9600);
}
void loop()
{
    for (int i = 0; i < 4; i++)
    {
        irsend.sendSony(0b101010010000, 12); // Prender TV. (Tabla 8).
        delay(40);
    }
    delay(5000);
}
```

2. Código 2. Reconocimiento de comandos independientes de locutor.

```
#if defined(ARDUINO) && ARDUINO >= 100
    #include "Arduino.h"
    #include "SoftwareSerial.h"
    SoftwareSerial port(12,13);
#else // Arduino 0022 - use modified NewSoftSerial
    #include "WProgram.h"
    #include "NewSoftSerial.h"
    NewSoftSerial port(12,13);
#endif
#include "EasyVR.h"
#include <IRremote.h>
IRsend irsend;
EasyVR easyvr(port);
EasyVRBridge bridge;
int8_t group, idx;
int verde = 4;
int azul = 5;
int rojo = 6;
void setup()
{
    pinMode(verde, OUTPUT);
    pinMode(azul, OUTPUT);
    pinMode(rojo, OUTPUT);
    if (bridge.check())
    {
        cli();
        bridge.loop(0, 1, 12, 13);
    }
    Serial.begin(9600);
    port.begin(9600);
}
```

```

if (!easyvr.detect())
{
    Serial.println("EasyVR not detected!");
    for (;;)
    }
digitalWrite(verde,LOW);
digitalWrite(azul,LOW);
digitalWrite(rojo,LOW);
easyvr.setPinOutput(EasyVR::IO1, LOW);
Serial.println("EasyVR detected!");
easyvr.setTimeout(6);
easyvr.setLanguage(4);
easyvr.setLevel(5);
}
void loop()
{
    digitalWrite(verde,HIGH);
    easyvr.setPinOutput(EasyVR::IO1, HIGH);
    Serial.print("Diga un comando del Wordset 3");
    easyvr.recognizeWord(3);
    do
    {
    }
    while (!easyvr.hasFinished());
    digitalWrite(verde,LOW);
    easyvr.setPinOutput(EasyVR::IO1, LOW); // LED off
    idx = easyvr.getWord();
    if (idx >= 0)
    {
        action();
    }
    else
    {
        if (easyvr.isTimeout())
            Serial.println("Timed out, try again...");
    }
}

```

```

    int16_t err = easyvr.getError();
    if (err >= 0)
    {
        digitalWrite(rojo,HIGH);
        Serial.print("Error ");
        Serial.println(err, HEX);
        delay(150);
        digitalWrite(rojo,LOW);
    }
}

void action()
{
    switch (idx)
    {
        case 0:
            digitalWrite(azul,HIGH);
            easyvr.playSound(0, EasyVR::VOL_FULL);
            Serial.println("CERO");
            break;
        .
        .
        .
        case 9:
            digitalWrite(azul,HIGH);
            easyvr.playSound(0, EasyVR::VOL_FULL);
            Serial.println("NUEVE");
            break;
        default:
            digitalWrite(rojo,HIGH);
            Serial.println("NADA");
    }
    digitalWrite(azul,LOW);
}

```

3. Código 3. Reconocimiento de comandos dependientes de locutor.

```
#if defined(ARDUINO) && ARDUINO >= 100
    #include "Arduino.h"
    #include "SoftwareSerial.h"
    SoftwareSerial port(12,13);
#else // Arduino 0022 - use modified NewSoftSerial
    #include "WProgram.h"
    #include "NewSoftSerial.h"
    NewSoftSerial port(12,13);
#endif
#include "EasyVR.h"
#include <IRremote.h>
IRsend irsend;
EasyVR easyvr(port);
EasyVRBridge bridge;
enum Groups
{
    GROUP_1 = 1, // PELICULAS
};
enum Group1
{
    PRENDER = 0,
    .
    .
    .
    APAGAR = 15,
};
int8_t group, idx;
int verde = 4;
int azul = 5;
int rojo = 6;
```

```

void setup()
{
    pinMode(verde, OUTPUT);
    pinMode(azul, OUTPUT);
    pinMode(rojo, OUTPUT);
    if (bridge.check())
    {
        cli();
        bridge.loop(0, 1, 12, 13);
    }
    Serial.begin(9600);
    port.begin(9600);
    if (!easyvr.detect())
    {
        Serial.println("EasyVR not detected!");
        for (;;)
        {
            digitalWrite(verde,LOW);
            digitalWrite(azul,LOW);
            digitalWrite(rojo,LOW);
            easyvr.setPinOutput(EasyVR::IO1, LOW);
            Serial.println("EasyVR detected!");
            easyvr.setTimeout(6);
            easyvr.setKnob(4);
            group=1;
        }
    }
    void loop()
    {
        digitalWrite(verde,HIGH);
        easyvr.setPinOutput(EasyVR::IO1, HIGH);
        Serial.print("Diga un comando en el grupo 1 ");
        easyvr.recognizeCommand(group);
        do
        {
        }
    }
}

```

```

while (!easyvr.hasFinished());
digitalWrite(verde,LOW);
easyvr.setPinOutput(EasyVR::IO1, LOW); // LED off
idx = easyvr.getCommand();
if (idx >= 0)
{
    uint8_t train = 0;
    char name[32];
    Serial.print("Command: ");
    Serial.print(idx);
    if (easyvr.dumpCommand(group, idx, name, train))
    {
        Serial.print(" = ");
        Serial.println(name);
    }
    else
        Serial.println();
    action();
}
else
{
    if (easyvr.isTimeout())
        Serial.println("Timed out, try again...");
    int16_t err = easyvr.getError();
    if (err >= 0)
    {
        digitalWrite(rojo,HIGH);
        Serial.print("Error ");
        Serial.println(err, HEX);
        delay(150);
        digitalWrite(rojo,LOW);
    }
}
}
}

```

```

void action()
{
    digitalWrite(azul,HIGH);
    switch (group)
    {
        case GROUP_1:
            switch (idx)
            {
                case PRENDER:
                    easyvr.playSound(0, EasyVR::VOL_FULL);
                    break;
                .
                .
                .
                case APAGAR:
                    easyvr.playSound(0, EasyVR::VOL_FULL);
                    break;
            }
        }
    digitalWrite(azul,LOW);
}

```

4. Código Final.

```
#if defined(ARDUINO) && ARDUINO >= 100
    #include "Arduino.h"
    #include "SoftwareSerial.h"
    SoftwareSerial port(12,13);
#else // Arduino 0022 - use modified NewSoftSerial
    #include "WProgram.h"
    #include "NewSoftSerial.h"
    NewSoftSerial port(12,13);
#endif
#include "EasyVR.h"
#include <IRremote.h>
IRsend irsend;
EasyVR easyvr(port);
EasyVRBridge bridge;
enum Groups
{
    GROUP_1 = 1, // PELICULAS
};
enum Group1
{
    PRENDER          = 0,
    PLAY              = 1,
    MENÚ              = 2,
    VOLUMEN           = 3,
    BAJAR             = 4,
    PAUSA             = 5,
    APAGAR            = 6,
};
int8_t group, idx;
int azul = 4;
```

```

void setup()
{
    pinMode(azul, OUTPUT);
    if (bridge.check())
    {
        cli();
        bridge.loop(0, 1, 12, 13);
    }
    Serial.begin(9600);
    port.begin(9600);
    if (!easyvr.detect())
    {
        Serial.println("EasyVR not detected!");
        for (;;)
        {
            digitalWrite(azul,LOW);
            easyvr.setPinOutput(EasyVR::IO1, LOW);
            Serial.println("EasyVR detected!");
            easyvr.setTimeout(6);
            easyvr.setLevel(5);    // 1 para comandos independientes configurables.
                                // 5 para comando dependientes.

            group=1;
        }
    }
    void loop()
    {
        digitalWrite(azul,HIGH);
        easyvr.setPinOutput(EasyVR::IO1, HIGH);
        Serial.print("Diga un comando en el grupo 1 ");
        easyvr.recognizeCommand(group);
        do
        {
        }
        while (!easyvr.hasFinished());
        digitalWrite(azul,LOW);
    }
}

```

```

easyvr.setPinOutput(EasyVR::IO1, LOW); // LED off
idx = easyvr.getCommand();
if (idx >= 0)
{
    uint8_t train = 0;
    char name[32];
    Serial.print("Command: ");
    Serial.print(idx);
    if (easyvr.dumpCommand(group, idx, name, train))
    {
        Serial.print(" = ");
        Serial.println(name);
    }
    else
        Serial.println();
    action();
}
else
{
    if (easyvr.isTimeout())
        Serial.println("Timed out, try again...");
    int16_t err = easyvr.getError();
    if (err >= 0)
    {
        digitalWrite(azul,HIGH);
        delay(50);
        digitalWrite(azul,LOW);
        delay(50);
        digitalWrite(azul,HIGH);
        delay(50);
        digitalWrite(azul,LOW);
        delay(50);
    }
}

```

```

        Serial.print("Error ");
        Serial.println(err, HEX);
        delay(50);
    }
}

void action()
{
    switch (group)
    {
        case GROUP_1:
            switch (idx)
            {
                case PRENDER:
                    if (gate == 0)
                    {
                        gate=1;
                        for (int i = 0; i < 4; i++)
                        {
                            irsend.sendSony(0b10101000101110010010, 20);
                            delay(40);
                        }
                        for (int i = 0; i < 4; i++)
                        {
                            irsend.sendSony(0b101010010000, 12);
                            delay(40);
                        }
                        for (int i = 0; i < 4; i++)
                        {
                            irsend.sendSony(0b101010000001, 12);
                            delay(40);
                        }
                    }
                }
            }
        }
    }
}

```

```

    for (int i = 0; i < 4; i++)
    {
        irsend.sendSony(0b001010010000, 12);
        delay(40);
    }
    delay(4000);
    for (int i = 0; i < 4; i++)
    {
        irsend.sendSony(0b01101000101110010010, 20);
        delay(40);
    }
}
break;

case PLAY:
    if (gate==1)
    {
        for (int i = 0; i < 4; i++)
        {
            irsend.sendSony(0b01001100101110010010, 20);
            delay(40);
        }
    }
    break;

case MENU:
    if (gate==1)
    {
        for (int i = 0; i < 4; i++)
        {
            irsend.sendSony(0b01001100101110010010, 20);
            delay(40);
        }
    }
    break;

```

```

case VOLUMEN:
    if (gate==1)
    {

        for (int i = 0; i < 8; i++)
        {
            irsend.sendSony(0b010010000001, 12);
            delay(40);
        }
    }
    break;

case BAJAR:
    if (gate==1)
    {
        for (int i = 0; i < 8; i++)
        {
            irsend.sendSony(0b110010000001, 12);
            delay(40);
        }
    }
    break;

case PAUSA:
    if (gate==1)
    {

        for (int i = 0; i < 4; i++)
        {
            irsend.sendSony(0b10011100101110010010, 20);
            delay(40);
        }
    }
    break;

```

```

case APAGAR:
    if (gate==1)
    {
        for (int i = 0; i < 4; i++)
        {
            irsend.sendSony(0b101010010000, 12);
            delay(40);
        }
        for (int i = 0; i < 4; i++)
        {
            irsend.sendSony(0b101010000001, 12);
            delay(40);
        }
        for (int i = 0; i < 4; i++)
        {
            irsend.sendSony(0b01101000101110010010, 20);
            delay(40);
        }
        delay(10000);
        for (int i = 0; i < 4; i++)
        {
            irsend.sendSony(0b10101000101110010010, 20);
            delay(40);
        }
        gate=0;
    }
    break;
}
}

```

ANEXO B. Resultados prueba estándar.

	Pág.
1. Medición de ruido de fondo.....	137
2. Medición de normalización de niveles de ruido de fondo.....	142
3. Medición de normalización de niveles de comandos de voz.....	147
4. Prueba de niveles de presión sonora en fase de entrenamiento.....	149
5. Variación de niveles de presión sonora con relación señal ruido constante.....	152
6. Variación de niveles dificultad con niveles de presión sonora distintos a la fase entrenamiento.....	154
7. Niveles de presión sonora de prueba para comandos independientes.....	157
7.1. Micrófono HORN.....	157
7.2. Micrófono Shure SM58.....	161
7.3. Micrófono AT 8015.....	163

1. Medición de ruido de fondo.

A continuación se exponen los datos adquiridos a través del software SVAN PC, mostrando los niveles de presión sonora por 1/3 de octava y los niveles equivalentes con y sin ponderación.

POSICIÓN 1 TOMA 1					
F [Hz]	Lev[dB]	F [Hz]	Lev[dB]	F [Hz]	Lev[dB]
25	42.9	315	22	4000	22.8
31.5	39.1	400	19.3	5000	24.7
40	31.5	500	20.2	6300	24.1
50	34.6	630	23	8000	26.2
63	27.9	800	22.1	10000	26.6
80	24.4	1000	23.6	12500	27.5
100	21.6	1250	26	16000	27.8
125	21.7	1600	22	20000	33.7
160	23.3	2000	20.1	TOT_A	34.5
200	19.8	2500	21	TOT_C	42.5
250	20.2	3150	21.8	TOT_Lin	46.4

POSICIÓN 1 TOMA 2					
F [Hz]	Lev[dB]	F [Hz]	Lev[dB]	F [Hz]	Lev[dB]
25	41.3	315	22.1	4000	22.6
31.5	35.8	400	19.3	5000	24.5
40	31.3	500	20.2	6300	23.8
50	32.5	630	23.1	8000	25.9
63	29.2	800	22.1	10000	26.4
80	27.5	1000	23.6	12500	27.3
100	28.9	1250	26	16000	27.6
125	25.1	1600	21.9	20000	33.7
160	24.5	2000	19.9	TOT_A	34.3
200	22.9	2500	21	TOT_C	41.4
250	20.4	3150	21.4	TOT_Lin	44.9

POSICIÓN 2 TOMA 1					
F [Hz]	Lev[dB]	F [Hz]	Lev[dB]	F [Hz]	Lev[dB]
25	38.9	315	22.1	4000	22.7
31.5	35.3	400	19.3	5000	24.7
40	32.7	500	20.2	6300	24.1
50	35.3	630	23.1	8000	26.1
63	31.5	800	22.1	10000	26.6
80	29.5	1000	23.6	12500	27.4
100	26.3	1250	26	16000	27.8
125	23.5	1600	22	20000	33.7
160	22.9	2000	20.1	TOT_A	34.5
200	19.6	2500	21	TOT_C	41.1
250	19.6	3150	21.8	TOT_Lin	44.3

POSICIÓN 2 TOMA 2					
F [Hz]	Lev[dB]	F [Hz]	Lev[dB]	F [Hz]	Lev[dB]
25	38.4	315	22	4000	22.4
31.5	32.2	400	19.1	5000	24.4
40	30.9	500	20.2	6300	23.8
50	31.3	630	23	8000	25.9
63	27.2	800	22.1	10000	26.3
80	28	1000	23.6	12500	27.2
100	25	1250	26	16000	27.6
125	22.5	1600	21.7	20000	33.6
160	23.3	2000	19.7	TOT_A	34.2
200	20	2500	20.6	TOT_C	39.7
250	19.6	3150	21.4	TOT_Lin	43.4

POSICIÓN 3 TOMA 1					
F [Hz]	Lev[dB]	F [Hz]	Lev[dB]	F [Hz]	Lev[dB]
25	28.6	315	22.1	4000	22.5
31.5	24.8	400	19.3	5000	24.5
40	27.2	500	20.2	6300	24
50	34.1	630	23.1	8000	26
63	26.1	800	22	10000	26.5
80	26.3	1000	23.5	12500	27.4
100	29.4	1250	26	16000	27.8
125	26.3	1600	22	20000	33.7
160	25.4	2000	19.9	TOT_A	34.5
200	21.7	2500	20.6	TOT_C	38.5
250	20.7	3150	21.7	TOT_Lin	41.1

POSICIÓN 3 TOMA 2					
F [Hz]	Lev[dB]	F [Hz]	Lev[dB]	F [Hz]	Lev[dB]
25	28.8	315	22.4	4000	22.2
31.5	22.6	400	19.3	5000	24.3
40	25	500	20.5	6300	23.7
50	33.7	630	23.3	8000	25.8
63	26.9	800	22.1	10000	26.2
80	25.3	1000	23.6	12500	27.1
100	27.2	1250	25.9	16000	27.5
125	26.5	1600	21.7	20000	33.6
160	23.1	2000	19.7	TOT_A	34.2
200	21.2	2500	20.6	TOT_C	38.1
250	19.6	3150	21.3	TOT_Lin	41.3

POSICIÓN 4 TOMA 1					
F [Hz]	Lev[dB]	F [Hz]	Lev[dB]	F [Hz]	Lev[dB]
25	37.4	315	22.4	4000	22.5
31.5	31.2	400	19.6	5000	24.6
40	31.6	500	20.4	6300	23.9
50	31.4	630	23.2	8000	26
63	29.1	800	22.1	10000	26.5
80	27.2	1000	23.6	12500	27.3
100	26.8	1250	26	16000	27.8
125	24.7	1600	22	20000	33.6
160	23.2	2000	19.9	TOT_A	34.4
200	20.9	2500	20.8	TOT_C	39.5
250	20	3150	21.7	TOT_Lin	42.9

POSICIÓN 4 TOMA 2					
F [Hz]	Lev[dB]	F [Hz]	Lev[dB]	F [Hz]	Lev[dB]
25	36.5	315	22	4000	22.1
31.5	31.1	400	19.1	5000	24.2
40	30.3	500	20	6300	23.6
50	28.7	630	23	8000	25.7
63	24.6	800	22	10000	26.2
80	26	1000	23.6	12500	27.1
100	23.1	1250	25.9	16000	27.5
125	22.9	1600	21.7	20000	33.6
160	24	2000	19.5	TOT_A	34.1
200	19.1	2500	20.6	TOT_C	38.3
250	18.9	3150	21.1	TOT_Lin	41.9

POSICIÓN 5 TOMA 1					
F [Hz]	Lev[dB]	F [Hz]	Lev[dB]	F [Hz]	Lev[dB]
25	44.7	315	22.1	4000	22.6
31.5	39.4	400	19.8	5000	24.6
40	35	500	20.5	6300	23.9
50	34.8	630	23.2	8000	26
63	29.3	800	22.2	10000	26.4
80	25.2	1000	23.6	12500	27.3
100	26.1	1250	26	16000	27.7
125	24.6	1600	21.9	20000	33.6
160	23.3	2000	19.9	TOT_A	34.4
200	20.7	2500	21	TOT_C	43.8
250	19.8	3150	21.5	TOT_Lin	47.7

POSICIÓN 5 TOMA 2					
F [Hz]	Lev[dB]	F [Hz]	Lev[dB]	F [Hz]	Lev[dB]
25	40.7	315	22	4000	22.2
31.5	35.5	400	19.3	5000	24.3
40	31.8	500	20.2	6300	23.6
50	30.9	630	23.1	8000	25.7
63	26.2	800	22	10000	26.2
80	24.3	1000	23.6	12500	27
100	24.7	1250	25.9	16000	27.3
125	23.3	1600	21.7	20000	33.6
160	22.5	2000	19.5	TOT_A	34
200	20.4	2500	20.6	TOT_C	40.4
250	19.1	3150	21.1	TOT_Lin	43.9

2. Normalización de niveles de ruido de fondo.

Las siguientes tablas llevan por título el nivel de ganancia en el cual se ubicó del fader del secuenciador utilizado, en donde se reprodujo la señal de ruido rosa mencionada.

FADER CANAL 0 DBFS					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	38.8	315	66	4000	68.5
31.5	53.7	400	68.1	5000	67.7
40	68.2	500	66.7	6300	68.2
50	78.3	630	68.9	8000	65.9
63	71.8	800	69.2	10000	64.4
80	72.6	1000	68	12500	63.6
100	71.1	1250	68.6	16000	62.1
125	67	1600	67.8	20000	58.8
160	68	2000	66.9	TOT_A	79.7
200	68.8	2500	68.8	TOT_C	83
250	68.5	3150	68.7	TOT_Lin	83.8

FADER CANAL -1 DBFS					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	38.6	315	65	4000	67.5
31.5	52.8	400	67.1	5000	66.7
40	67.2	500	65.7	6300	67.2
50	77.3	630	67.9	8000	64.9
63	70.9	800	68.1	10000	63.4
80	71.6	1000	67	12500	62.6
100	70.1	1250	67.6	16000	61.1
125	66	1600	66.8	20000	57.8
160	67	2000	65.9	TOT_A	78.7
200	67.8	2500	67.7	TOT_C	82
250	67.5	3150	67.7	TOT_Lin	82.8

FADER CANAL -2 DBFS					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	37.2	315	64	4000	66.4
31.5	51.7	400	66.1	5000	65.7
40	66.3	500	64.6	6300	66.2
50	76.4	630	66.9	8000	63.9
63	69.9	800	67.2	10000	62.4
80	70.5	1000	66	12500	61.6
100	69.1	1250	66.6	16000	60.1
125	65	1600	65.9	20000	56.8
160	66	2000	64.9	TOT_A	77.7
200	66.8	2500	66.8	TOT_C	81.1
250	66.5	3150	66.7	TOT_Lin	81.8

FADER CANAL -3DBFS					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	37	315	62.9	4000	65.4
31.5	50.8	400	65.1	5000	64.7
40	65.3	500	63.7	6300	65.2
50	75.4	630	65.9	8000	62.9
63	68.9	800	66.2	10000	61.4
80	69.6	1000	65	12500	60.6
100	68.1	1250	65.6	16000	59.1
125	64	1600	64.9	20000	55.9
160	65	2000	63.9	TOT_A	76.7
200	65.8	2500	65.7	TOT_C	80.1
250	65.5	3150	65.7	TOT_Lin	80.8

FADER CANAL -4 DBFS					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	37.1	315	62	4000	64.4
31.5	49.8	400	64.1	5000	63.7
40	64.3	500	62.7	6300	64.2
50	74.1	630	64.9	8000	61.9
63	67.9	800	65.2	10000	60.3
80	68.5	1000	64	12500	59.6
100	67.1	1250	64.6	16000	58.1
125	63	1600	63.9	20000	54.8
160	63.9	2000	62.9	TOT_A	75.7
200	64.8	2500	64.7	TOT_C	79
250	64.5	3150	64.7	TOT_Lin	79.7

FADER CANAL -5 DBFS					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	34.8	315	61	4000	63.4
31.5	48.8	400	63.1	5000	62.7
40	63.4	500	61.6	6300	63.2
50	73.4	630	63.9	8000	60.9
63	66.9	800	64.2	10000	59.4
80	67.6	1000	63	12500	58.6
100	66.1	1250	63.6	16000	57.1
125	62	1600	62.9	20000	53.9
160	63	2000	61.9	TOT_A	74.7
200	63.8	2500	63.7	TOT_C	78.1
250	63.5	3150	63.7	TOT_Lin	78.8

FADER CANAL -6DBFS					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	35	315	60	4000	62.4
31.5	47.8	400	62.1	5000	61.7
40	62.4	500	60.6	6300	62.2
50	72.4	630	62.9	8000	59.9
63	65.9	800	63.2	10000	58.4
80	66.6	1000	62	12500	57.6
100	65.1	1250	62.7	16000	56.1
125	61	1600	61.9	20000	52.9
160	62	2000	60.9	TOT_A	73.7
200	62.8	2500	62.7	TOT_C	77.1
250	62.5	3150	62.8	TOT_Lin	77.8

FADER CANAL -7DBFS					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	34.6	315	59	4000	61.4
31.5	47	400	61.1	5000	60.7
40	61.4	500	59.6	6300	61.1
50	71.3	630	61.9	8000	58.9
63	64.9	800	62.2	10000	57.3
80	65.6	1000	61	12500	56.6
100	64.1	1250	61.6	16000	55.1
125	60.1	1600	60.9	20000	51.9
160	61	2000	60	TOT_A	72.7
200	61.8	2500	61.7	TOT_C	76.1
250	61.5	3150	61.7	TOT_Lin	76.8

FADER CANAL -8 DBFS					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	38.7	315	58	4000	60.4
31.5	46	400	60.1	5000	59.6
40	60.4	500	58.6	6300	60.2
50	70.4	630	60.9	8000	57.9
63	63.9	800	61.2	10000	56.4
80	64.6	1000	60	12500	55.6
100	63.1	1250	60.7	16000	54.1
125	59.1	1600	59.9	20000	50.9
160	60	2000	59	TOT_A	71.7
200	60.8	2500	60.7	TOT_C	75.1
250	60.5	3150	60.8	TOT_Lin	75.8

FADER CANAL -9DBFS					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	35.7	315	57	4000	59.4
31.5	45.1	400	59.1	5000	58.6
40	59.4	500	57.6	6300	59.2
50	69.4	630	59.9	8000	56.9
63	62.9	800	60.2	10000	55.3
80	63.6	1000	59	12500	54.6
100	62.2	1250	59.7	16000	53.1
125	58.1	1600	58.9	20000	49.9
160	59.1	2000	58	TOT_A	70.7
200	59.9	2500	59.7	TOT_C	74.1
250	59.5	3150	59.8	TOT_Lin	74.8

3. Normalización de niveles de comandos de voz.

A continuación se exponen los niveles de presión sonora adquiridos al llevar a cabo el proceso de normalización de niveles para comandos de voz, realizado independientemente para cada locutor.

LOCUTOR 1					
F [Hz]	Leq [dB]	F [Hz]	Leq [dB]	F [Hz]	Leq [dB]
25.0	33.0	315.0	76.6	4000.0	67.0
31.5	28.7	400.0	74.3	5000.0	55.5
40.0	29.8	500.0	79.4	6300.0	56.2
50.0	42.9	630.0	77.7	8000.0	60.5
63.0	41.1	800.0	63.1	10000.0	51.0
80.0	58.4	1000.0	58.8	12500.0	46.4
100.0	69.4	1250.0	65.0	16000.0	37.5
125.0	64.4	1600.0	75.8	20000.0	34.6
160.0	67.5	2000.0	76.8	TOT_A	83.3
200.0	64.5	2500.0	75.0	TOT_C	84.4
250.0	69.0	3150.0	71.5	TOT_Lin	84.5

LOCUTOR 2					
F [Hz]	Leq [dB]	F [Hz]	Leq [dB]	F [Hz]	Leq [dB]
25.0	34.5	315.0	65.1	4000.0	66.1
31.5	31.9	400.0	66.8	5000.0	56.0
40.0	32.2	500.0	76.9	6300.0	54.7
50.0	46.3	630.0	72.8	8000.0	66.1
63.0	42.2	800.0	73.6	10000.0	63.0
80.0	38.0	1000.0	59.3	12500.0	56.1
100.0	47.1	1250.0	61.8	16000.0	48.4
125.0	35.2	1600.0	64.3	20000.0	39.6
160.0	36.2	2000.0	77.8	TOT_A	83.9
200.0	69.2	2500.0	78.1	TOT_C	84.2
250.0	74.6	3150.0	71.3	TOT_Lin	84.4

LOCUTOR 3					
F [Hz]	Leq [dB]	F [Hz]	Leq [dB]	F [Hz]	Leq [dB]
25.0	31.4	315.0	67.0	4000.0	65.1
31.5	28.4	400.0	68.6	5000.0	63.1
40.0	33.9	500.0	75.4	6300.0	65.8
50.0	44.5	630.0	65.0	8000.0	64.4
63.0	43.9	800.0	71.5	10000.0	66.5
80.0	39.8	1000.0	60.2	12500.0	58.1
100.0	41.7	1250.0	61.9	16000.0	55.8
125.0	36.4	1600.0	55.8	20000.0	43.8
160.0	61.9	2000.0	69.8	TOT_A	80.0
200.0	75.1	2500.0	70.3	TOT_C	84.5
250.0	79.1	3150.0	72.1	TOT_Lin	84.6

4. Prueba de niveles en la fase de entrenamiento.

En las siguientes tablas, las celdas mostradas en rojo con una “X” representan un proceso de entrenamiento de comandos fallido, mientras que las celdas mostradas en verdes con una “Y” representan un proceso satisfactorio.

MICRÓFONO HORN				
LOC.	R. S/R. [dB]	R.FONDO [dB]	NIVEL VOZ [dB]	FASE ENTR.
1	-20	90	70	X
	-15	85	70	X
	-10	80	70	X
	-5	75	70	X
	0	70	70	X
	5	65	70	X
	10	60	70	X
	15	55	70	X
	20	50	70	Y
	25	45	70	Y
2	-20	90	70	X
	-15	85	70	X
	-10	80	70	X
	-5	75	70	X
	0	70	70	X
	5	65	70	X
	10	60	70	X
	15	55	70	X
	20	50	70	Y
	25	45	70	Y
3	-20	90	70	X
	-15	85	70	X
	-10	80	70	X
	-5	75	70	X
	0	70	70	X
	5	65	70	X
	10	60	70	X
	15	55	70	X
	20	50	70	Y
	25	45	70	Y

MICRÓFONO SM58				
LOC.	R. S/R. [dB]	R.FONDO [dB]	NIVEL VOZ [dB]	FASE ENTR.
1	-20	90	70	X
	-15	85	70	X
	-10	80	70	X
	-5	75	70	Y
	0	70	70	Y
	5	65	70	Y
	10	60	70	Y
	15	55	70	Y
	20	50	70	Y
	25	45	70	Y
2	-20	90	70	X
	-15	85	70	X
	-10	80	70	X
	-5	75	70	Y
	0	70	70	Y
	5	65	70	Y
	10	60	70	Y
	15	55	70	Y
	20	50	70	Y
	25	45	70	Y
3	-20	90	70	X
	-15	85	70	X
	-10	80	70	X
	-5	75	70	Y
	0	70	70	Y
	5	65	70	Y
	10	60	70	Y
	15	55	70	Y
	20	50	70	Y
	25	45	70	Y

MICRÓFONO AT 8015				
LOC.	R. S/R. [dB]	R.FONDO [dB]	NIVEL VOZ [dB]	FASE ENTR.
1	-20	90	70	X
	-15	85	70	X
	-10	80	70	X
	-5	75	70	X
	0	70	70	X
	5	65	70	Y
	10	60	70	Y
	15	55	70	Y
	20	50	70	Y
	25	45	70	Y
2	-20	90	70	X
	-15	85	70	X
	-10	80	70	X
	-5	75	70	X
	0	70	70	X
	5	65	70	Y
	10	60	70	Y
	15	55	70	Y
	20	50	70	Y
	25	45	70	Y
3	-20	90	70	X
	-15	85	70	X
	-10	80	70	X
	-5	75	70	X
	0	70	70	X
	5	65	70	Y
	10	60	70	Y
	15	55	70	Y
	20	50	70	Y
	25	45	70	Y

5. Variación de niveles de presión sonora con relación señal ruido constante.

En las siguientes tablas, las celdas mostradas en rojo con una “X” representan un proceso de entrenamiento de comandos fallido, mientras que las celdas mostradas en verdes con una “Y” representan un proceso satisfactorio.

MICRÓFONO HORN				
LOCUTOR	R. S/R. [dB]	R. FONDO [dB]	NIVEL VOZ [dB]	FASE ENTR.
1	20	45	65	Y
		50	70	Y
		55	75	Y
		60	80	X
		65	85	X
		70	90	X
2	20	45	65	Y
		50	70	Y
		55	75	Y
		60	80	X
		65	85	X
		70	90	X
3	20	45	65	Y
		50	70	Y
		55	75	Y
		60	80	X
		65	85	X
		70	90	X

MICRÓFONO SM58				
LOCUTOR	R. S/R. [dB]	R. FONDO [dB]	NIVEL VOZ [dB]	FASE ENTR.
1	-5	65	60	X
		70	65	X
		75	70	Y
		80	75	Y
		85	80	X
		90	85	X
2	-5	65	60	X
		70	65	X
		75	70	Y
		80	75	Y
		85	80	X
		90	85	X
3	-5	65	60	X
		70	65	X
		75	70	Y
		80	75	Y
		85	80	X
		90	85	X

MICRÓFONO AT 8015				
LOCUTOR	R. S/R. [dB]	R. FONDO [dB]	NIVEL VOZ [dB]	FASE ENTR.
1	5	55	60	Y
		60	65	Y
		65	70	Y
		70	75	X
		75	80	X
		80	85	X
2	5	55	60	Y
		60	65	Y
		65	70	Y
		70	75	X
		75	80	X
		80	85	X
3	5	55	60	Y
		60	65	Y
		65	70	Y
		70	75	X
		75	80	X
		80	85	X

6. Variación de niveles dificultad con niveles de presión sonora distintos a la fase entrenamiento.

Las celdas cuyo título es “NIVELES DE ENTR.” Hacen referencia a la relación señal ruido y a los niveles de presión sonora con los cuales se llevó a cabo la fase de entrenamiento de comandos de voz. Las celdas mostradas en verde representan los niveles para los cuales la fase de reconocimiento de comandos de voz dependientes de locutor fue satisfactoria.

MICRÓFONO HORN																	
NIVELES DE ENTR. [dB]	NIVELES DE PRUEBA		NIVELES DE DIFICULTAD														
	RUIDO DE FONDO [dB]	NIVEL VOZ [dB]	1			2			3			4			5		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
-20 (45/65)	45	65	10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	70		9	9	9	9	9	9	9	9	9	9	9	9	9	9	
	75		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	80		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-20 (50/70)	45	70	10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	70		7	7	7	7	7	7	7	7	7	7	7	7	7	7	
	75		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	80		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
-20 (55/75)	45	75	10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	70		4	4	4	4	4	4	4	4	4	4	4	4	4	4	
	75		0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	80		0	0	0	0	0	0	0	0	0	0	0	0	0	0	

MICRÓFONO SM58																	
NIVELES DE ENTR. [dB]	NIVELES DE PRUEBA		NIVELES DE DIFICULTAD														
	RUIDO DE FONDO [dB]	NIVEL VOZ [dB]	1			2			3			4			5		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
5(75/70)	45	70	10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	85		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5(80/75)	45	75	10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	85		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	90		10	10	10	10	10	10	10	10	10	10	10	10	10	10	
	95		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MICRÓFONO AT 8015																	
NIVELES ENTR. [dB]	NIVELES DE PRUEBA		NIVELES DE DIFICULTAD														
	RUIDO DE FONDO [dB]	NIVEL VOZ [dB]	1			2			3			4			5		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
-5(55/60)	45	60	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
	80		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	85		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-5(60/65)	45	65	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
	85		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-5(65/70)	45	70	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	85		4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

7. Niveles de prueba para comandos independientes.

Las celdas mostradas en verde representan los niveles para los cuales la fase de reconocimiento de comandos de voz independientes de locutor fue satisfactoria.

7.1. Micrófono HORN.

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
	R. FONDO	NIVEL VOZ	0			1			2			3			4		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
HORN	45	60	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	70		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	75		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	80		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	85		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
	R. FONDO	NIVEL VOZ	0			1			2			3			4		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
HORN	45	65	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
	75		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	80		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	85		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
	R. FONDO	NIVEL VOZ	0			1			2			3			4		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
HORN	45	70	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	80		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	85		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
	R. FONDO	NIVEL VOZ	0			1			2			3			4		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
HORN	45	75	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	85		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	95		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
	R. FONDO	NIVEL VOZ	0			1			2			3			4		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
HORN	45	80	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	85		4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	95		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
	R. FONDO	NIVEL VOZ	0			1			2			3			4		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
HORN	45	85	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	85		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	90		6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
	95		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
	R. FONDO	NIVEL VOZ	0			1			2			3			4		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
HORN	45	90	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	85		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	90		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	95		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

7.2. Micrófono Shure SM58.

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
	R. FONDO	NIVEL VOZ	0			1			2			3			4		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
SM 58	45	75	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	85		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	90		5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	95		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
	R. FONDO	NIVEL VOZ	0			1			2			3			4		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
SM 58	45	80	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	85		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	90		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	95		6	6	6	6	6	6	6	6	6	6	6	6	6	6	6

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
	R. FONDO	NIVEL VOZ	0			1			2			3			4		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
SM 58	45	85	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	85		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	90		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	95		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
	R. FONDO	NIVEL VOZ	0			1			2			3			4		
			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
SM 58	45	90	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	85		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	90		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	95		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10

7.3. Micrófono AT8015.

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
			0			1			2			3			4		
	R. FONDO	NIVEL VOZ	LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
AT 8015	45	60	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	75		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	80		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	85		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
			0			1			2			3			4		
	R. FONDO	NIVEL VOZ	LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
AT 8015	45	65	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	80		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	85		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
			0			1			2			3			4		
	R. FONDO	NIVEL VOZ	LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
AT 8015	45	70	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
	85		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
			0			1			2			3			4		
	R. FONDO	NIVEL VOZ	LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
AT 8015	45	75	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	85		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	90		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	95		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
			0			1			2			3			4		
	R. FONDO	NIVEL VOZ	LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
AT 8015	45	80	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	85		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	90		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	95		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
			0			1			2			3			4		
	R. FONDO	NIVEL VOZ	LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
AT 8015	45	85	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	85		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	90		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	95		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MIC.	NIVELES DE PRUEBA [dB]		NIVELES DIFICULTAD														
			0			1			2			3			4		
	R. FONDO	NIVEL VOZ	LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR			LOCUTOR		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
AT 8015	45	90	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	50		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	55		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	60		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	65		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	70		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	75		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	80		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	85		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	90		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	95		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10

ANEXO C. Resultados prueba real.

	Pág.
1. Medición de ruido de fondo.....	168
2. Medición de niveles de presión sonora generados por el equipo de sonido.....	171
3. Medición de efectividad para la prueba real.....	176

1. Medición de ruido de fondo.

A continuación se exponen los datos adquiridos a través del software SVAN PC, mostrando los niveles de presión sonora por 1/3 de octava y los niveles equivalentes con y sin ponderación.

POSICIÓN 1 TOMA 1					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	55.7	315	39.8	4000	25.5
31.5	47.9	400	39.1	5000	24.9
40	46	500	37.3	6300	22.7
50	45.8	630	34.2	8000	23.3
63	38.4	800	34.8	10000	20.3
80	40.5	1000	37.6	12500	20.3
100	41.5	1250	31.5	16000	14.3
125	43.2	1600	29.7	20000	33.4
160	38.2	2000	28.5	TOT_A	44.3
200	41.2	2500	26.6	TOT_C	55.5
250	44.1	3150	25.8	TOT_Lin	58.6

POSICIÓN 1 TOMA 2					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	56.5	315	35.4	4000	22.5
31.5	48.3	400	32.8	5000	23
40	46.5	500	33.2	6300	18.1
50	45.9	630	32	8000	22.3
63	37.4	800	30.9	10000	15.3
80	40	1000	31.1	12500	20
100	40.2	1250	30.8	16000	14.3
125	38.3	1600	28.6	20000	33.9
160	36.9	2000	26.1	TOT_A	40.2
200	37.6	2500	25	TOT_C	54.9
250	38.2	3150	23.1	TOT_Lin	58.6

POSICIÓN 1 TOMA 3					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	56.8	315	40.1	4000	23
31.5	49.6	400	39.6	5000	23.5
40	50	500	38.3	6300	20.5
50	47.7	630	34.6	8000	22.6
63	39.9	800	35	10000	16
80	42.2	1000	38.7	12500	20.3
100	41	1250	32.3	16000	14.3
125	42	1600	29.6	20000	34.1
160	39	2000	27.5	TOT_A	44.6
200	41.7	2500	24.3	TOT_C	56.6
250	43.4	3150	22.4	TOT_Lin	59.8

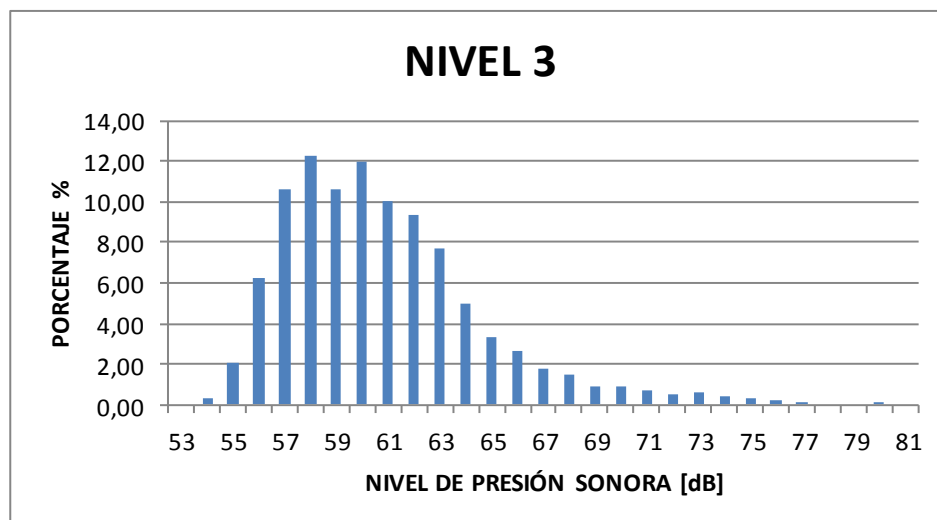
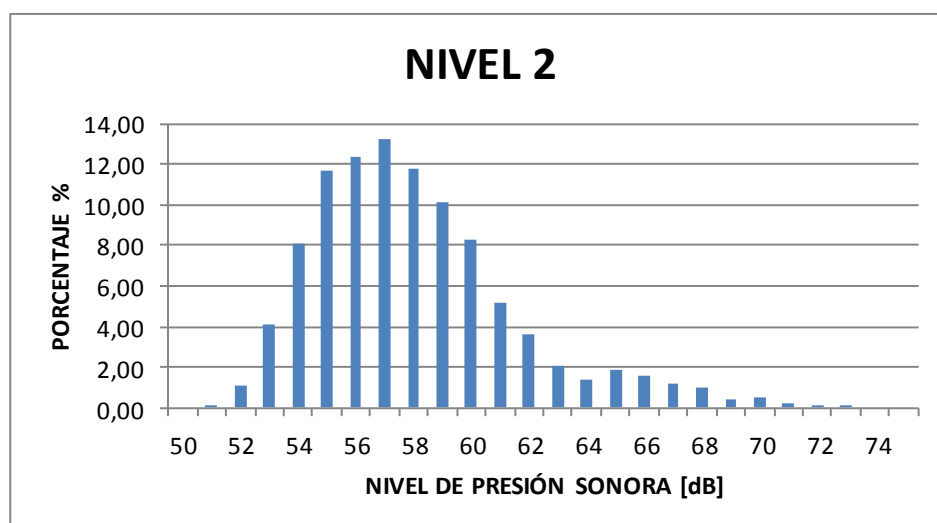
POSICIÓN 2 TOMA 1					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	53.8	315	46.8	4000	36
31.5	45.6	400	42.6	5000	33.5
40	45.3	500	54.6	6300	29.1
50	37.4	630	50.8	8000	28.7
63	37.1	800	40.8	10000	24.1
80	42.8	1000	37.6	12500	23.4
100	38.9	1250	35.2	16000	14.3
125	40.6	1600	40.2	20000	33.9
160	41.9	2000	40.2	TOT_A	54.6
200	40.1	2500	38.1	TOT_C	58.5
250	40.7	3150	36.5	TOT_Lin	59.8

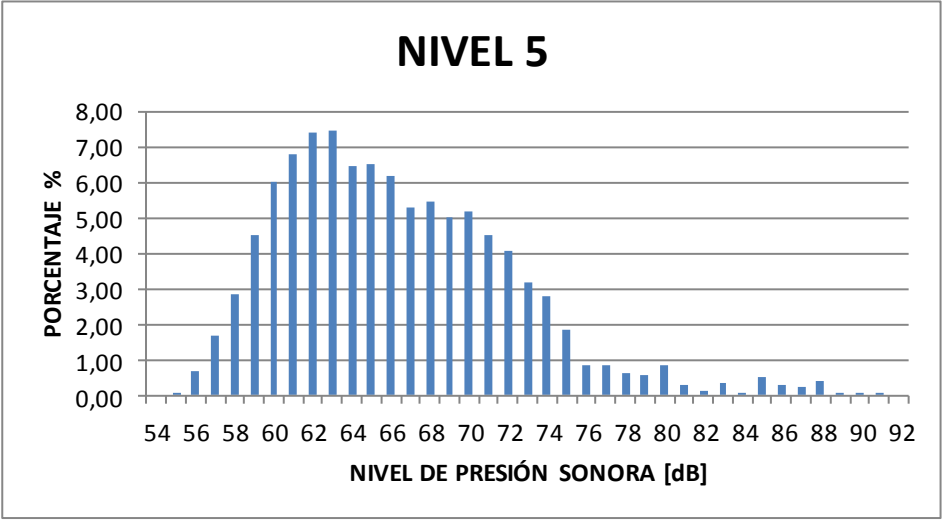
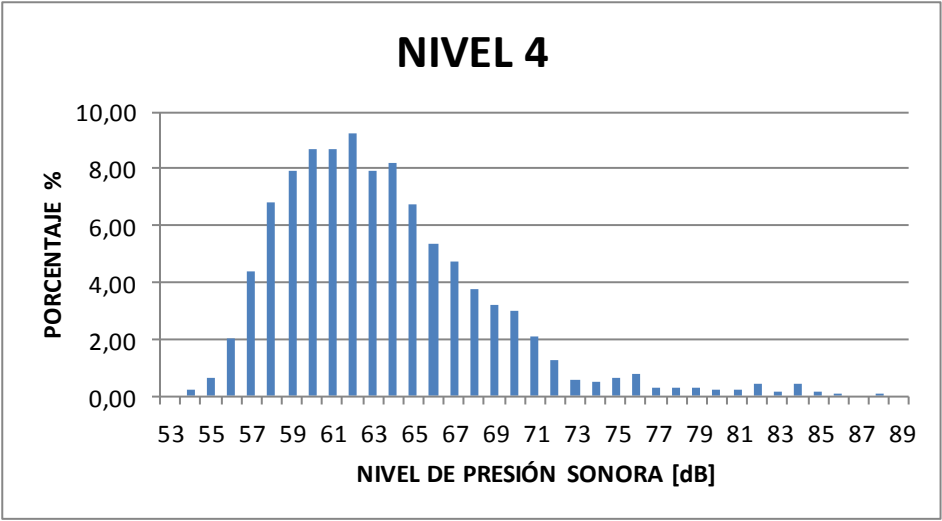
POSICIÓN 2 TOMA 2					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	54.9	315	40.2	4000	23.8
31.5	47.9	400	39.3	5000	23.7
40	43.1	500	38	6300	18.5
50	35.9	630	34.4	8000	22.3
63	36.1	800	39	10000	15.3
80	43.1	1000	41.7	12500	20.3
100	39.1	1250	40.7	16000	14.3
125	40.2	1600	36.6	20000	34.1
160	42	2000	27.2	TOT_A	47.3
200	39.8	2500	24.7	TOT_C	55
250	42.6	3150	22.4	TOT_Lin	58

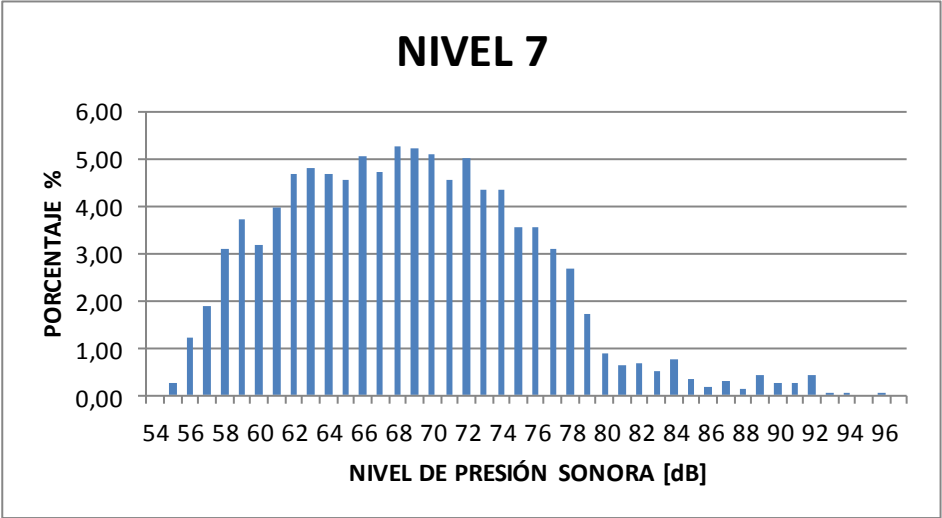
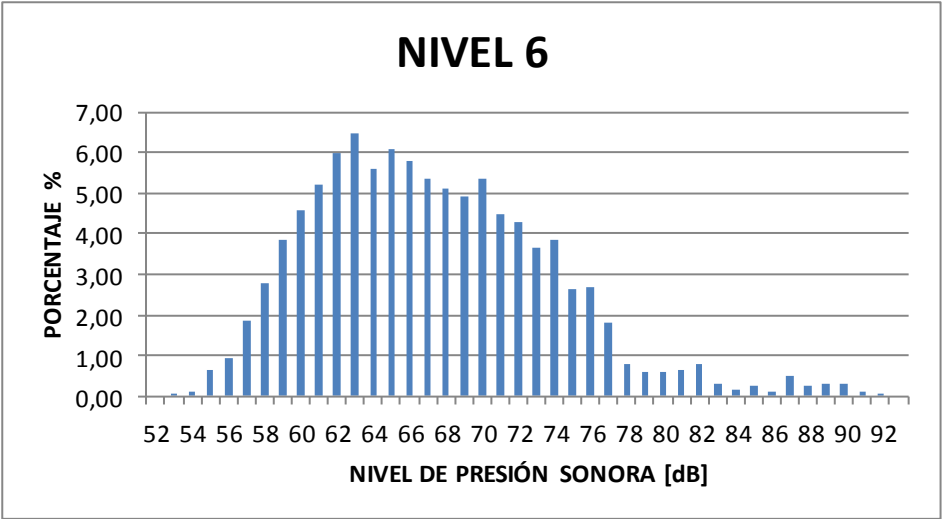
POSICIÓN 2 TOMA 3					
F[Hz]	Lev[dB]	F[Hz]	Lev[dB]	F[Hz]	Lev[dB]
25	54.1	315	34.3	4000	24.6
31.5	45.1	400	31.6	5000	24.4
40	43.3	500	32.4	6300	20.3
50	36.7	630	30.9	8000	23.3
63	35.3	800	29.4	10000	18.1
80	40.8	1000	30	12500	21.2
100	34.9	1250	30.2	16000	14.3
125	36.7	1600	27.4	20000	34.2
160	38.7	2000	26.3	TOT_A	39.3
200	36.4	2500	24.8	TOT_C	52.5
250	34.8	3150	23.4	TOT_Lin	56.3

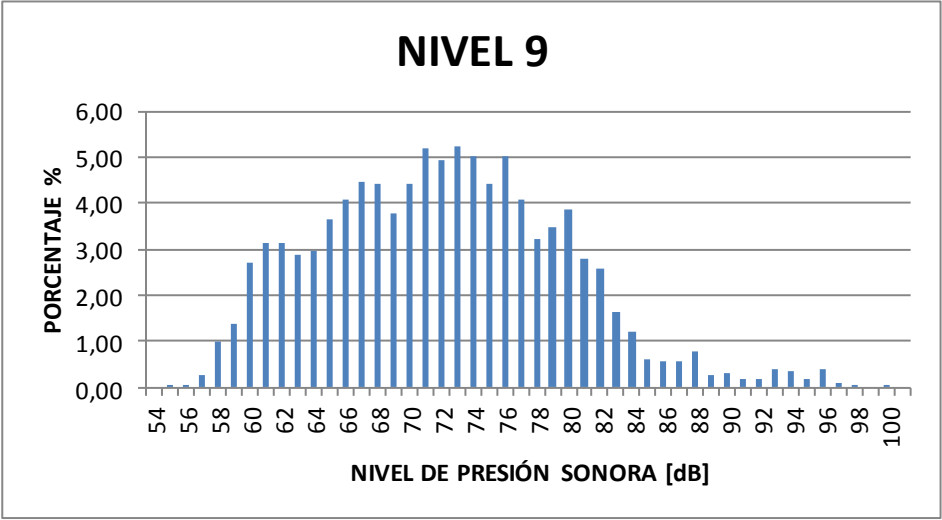
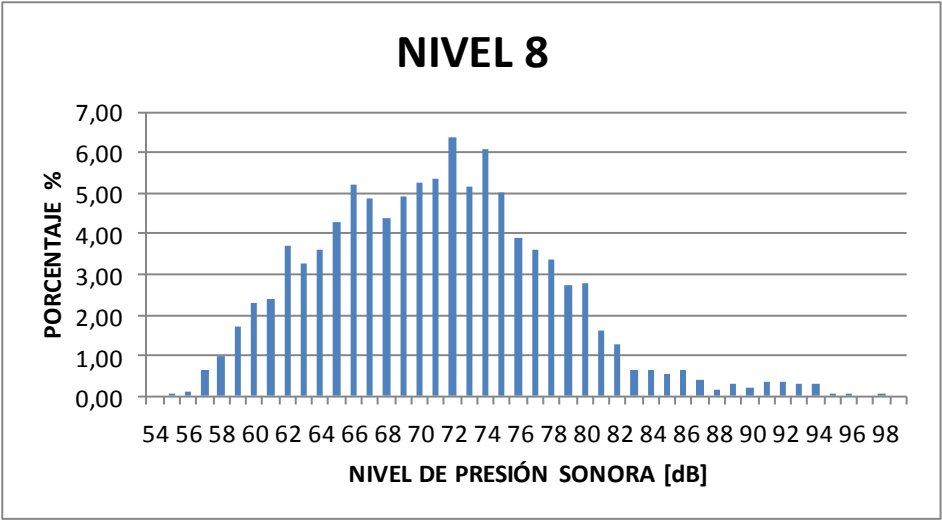
2. Medición de niveles de presión sonora generados por el equipo de sonido.

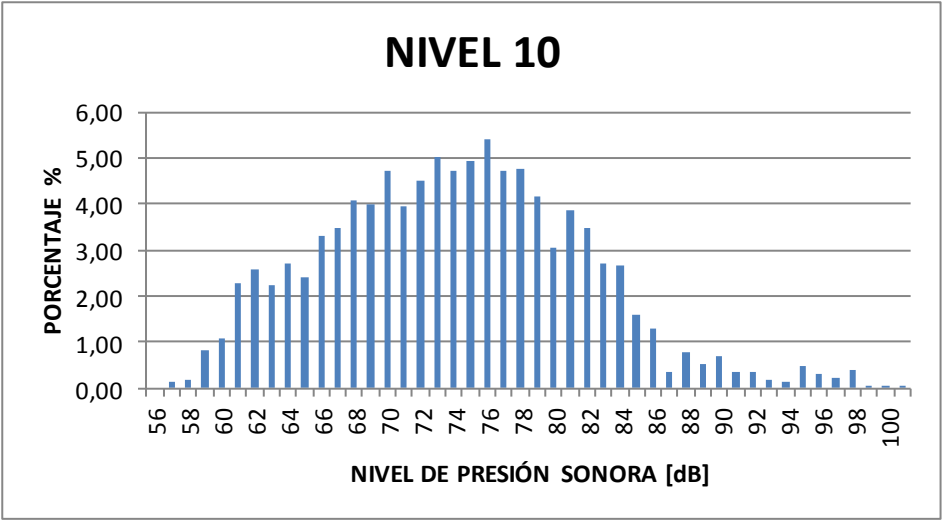
En las figuras mostradas a continuación, se exponen los histogramas de las mediciones realizadas para los niveles de ganancia del equipo de sonido al reproducir el audio de una película, mostrando los niveles de presión sonora con mayor porcentaje de muestras adquiridas.











3. Medición de efectividad prueba real.

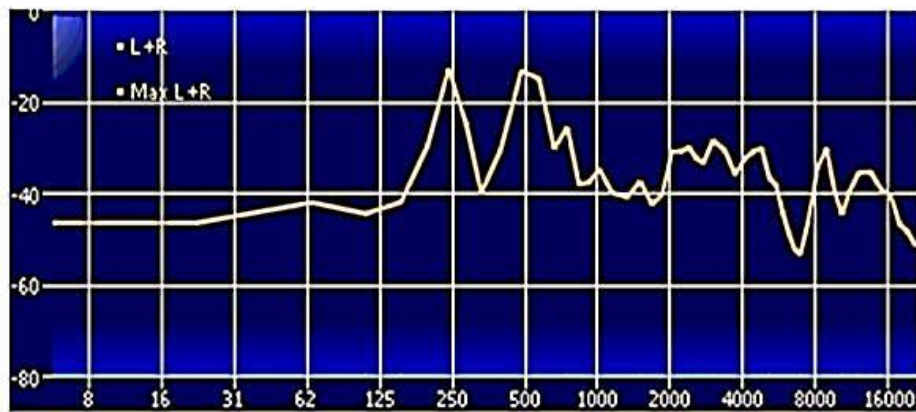
Las celdas mostradas en verde representan los niveles para los cuales la fase de reconocimiento de comandos de voz fue satisfactoria.

VARIACIÓN NIVELES DE DIFICULTAD Y GANANCIA												
MICRÓFONO	COMANDO	NIVEL DIFICULTAD	NIVEL DE GANANCIA									
			2	3	4	5	6	7	8	9	10	
HORN	PRENDER	5	10	10	4	0	0	0	0	0	0	
	APAGAR	5	10	10	6	0	0	0	0	0	0	
	PLAY	5	10	10	2	0	0	0	0	0	0	
	PAUSA	5	10	10	0	0	0	0	0	0	0	
	PRENDER	1	10	10	8	0	0	0	0	0	0	
	APAGAR	1	10	10	9	0	0	0	0	0	0	
	PLAY	1	10	10	4	0	0	0	0	0	0	
	PAUSA	1	10	10	8	0	0	0	0	0	0	
SM58	PRENDER	5	8	0	0	0	0	0	0	0	0	
	APAGAR	5	6	0	0	0	0	0	0	0	0	
	PLAY	5	9	0	0	0	0	0	0	0	0	
	PAUSA	5	6	0	0	0	0	0	0	0	0	
	PRENDER	1	10	0	0	0	0	0	0	0	0	
	APAGAR	1	10	0	0	0	0	0	0	0	0	
	PLAY	1	10	0	0	0	0	0	0	0	0	
	PAUSA	1	9	0	0	0	0	0	0	0	0	
AT 8015	PRENDER	5	10	10	8	0	0	0	0	0	0	
	APAGAR	5	10	10	8	0	0	0	0	0	0	
	PLAY	5	10	8	6	0	0	0	0	0	0	
	PAUSA	5	10	10	5	0	0	0	0	0	0	
	PRENDER	1	10	10	10	0	0	0	0	0	0	
	APAGAR	1	10	10	10	0	0	0	0	0	0	
	PLAY	1	10	10	6	0	0	0	0	0	0	
	PAUSA	1	10	10	9	0	0	0	0	0	0	

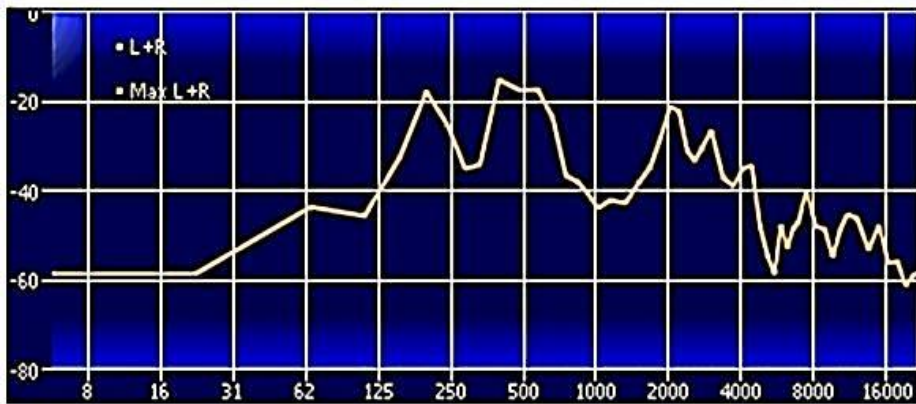
ANEXO D. ANALISIS DE CARACTERÍSTICAS GENERALES DE LA VOZ DE LOS LOCUTORES AL PRONUNCIAR EL COMANDO PRENDER.

	Pág.
1. Espectro con valores pico.....	178
2. Espectro con valores promedio.....	179
3. Espectro con visualización de formantes.....	180
4. Espectrogramas.....	181

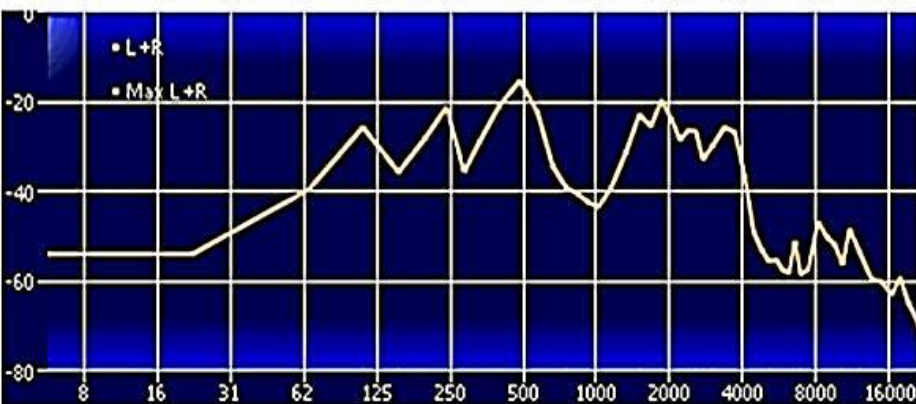
1. Espectro con valores pico.



LOCUTOR 1
NIÑO 12
AÑOS

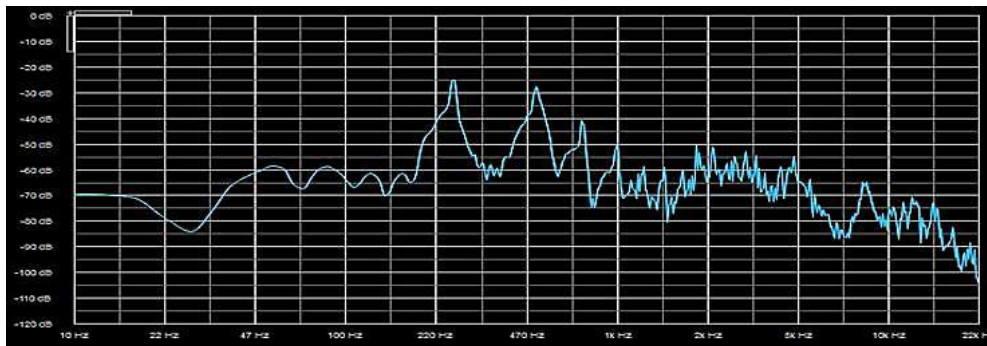


LOCUTOR 2
MUJER 50
AÑOS

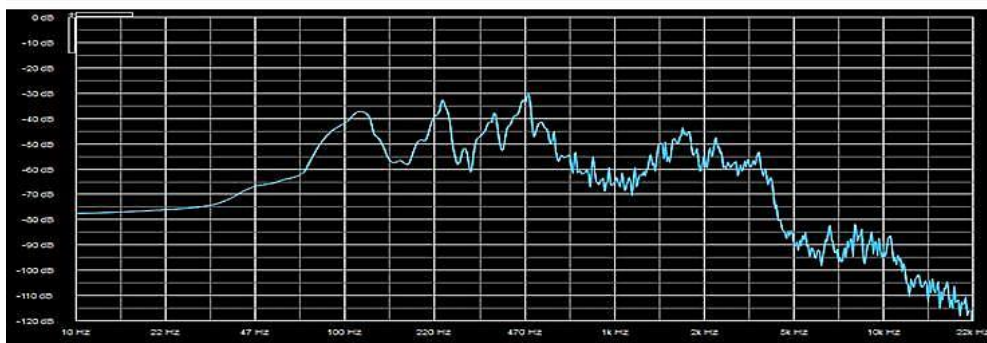


LOCUTOR 3
HOMBRE 23
AÑOS

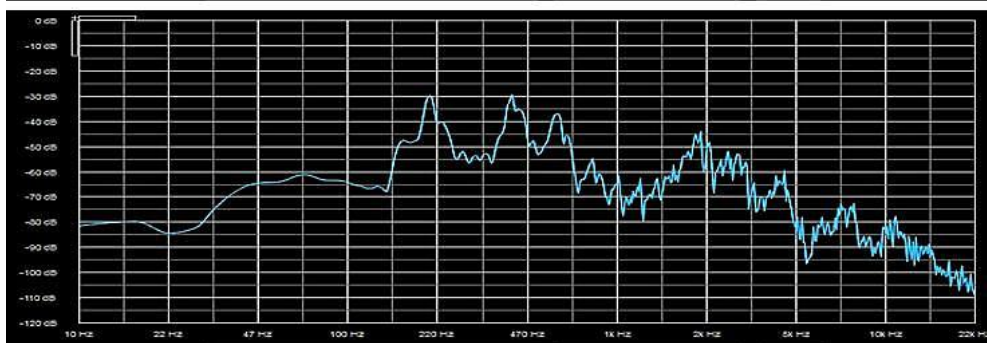
2. Espectro con valores promedio.



LOCUTOR 1
NIÑO 12
AÑOS

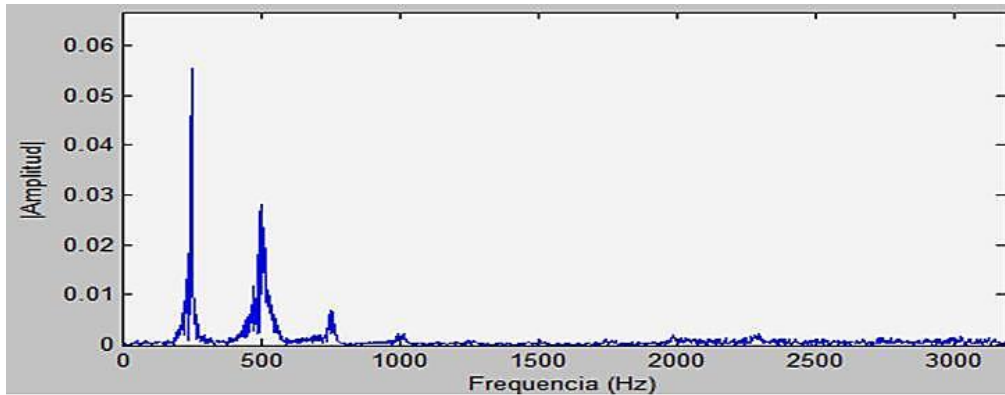


LOCUTOR 2
MUJER 50
AÑOS

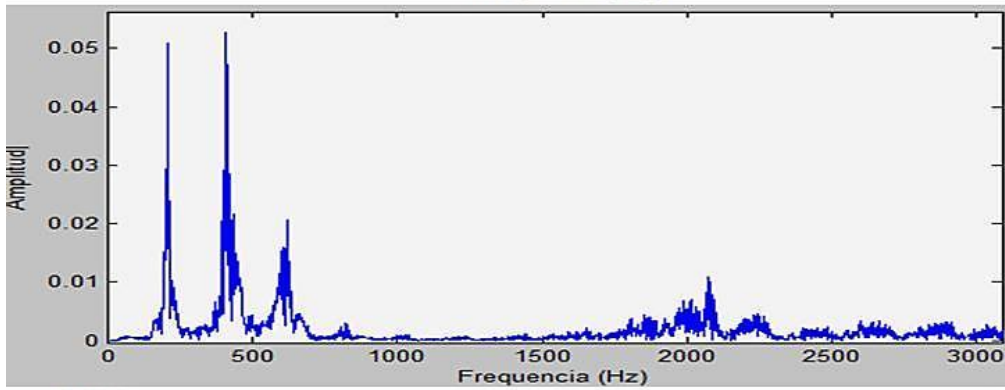


LOCUTOR 3
HOMBRE 23
AÑOS

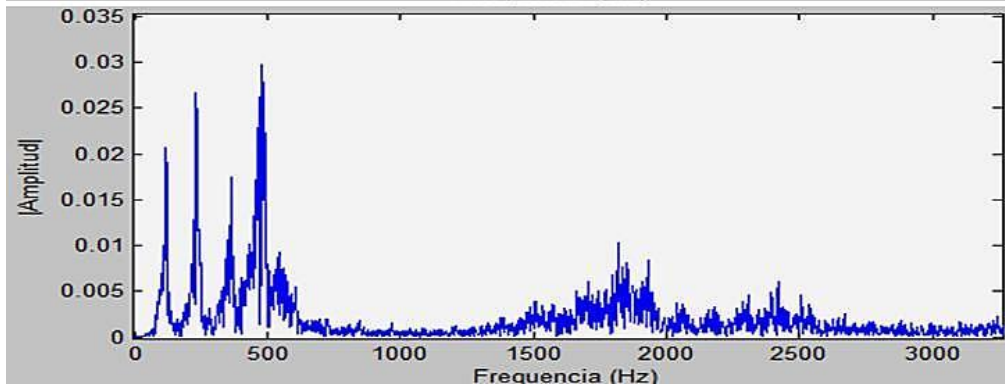
3. Espectro con visualización de formantes representativas.



LOCUTOR 1
NIÑO 12
AÑOS

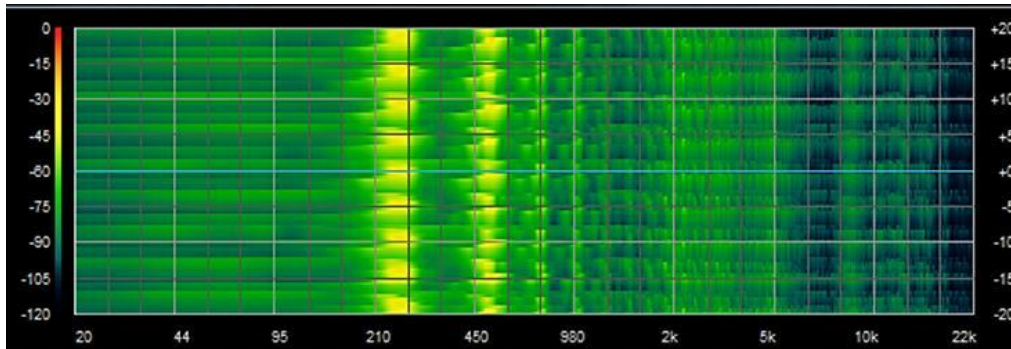


LOCUTOR 2
MUJER 50
AÑOS

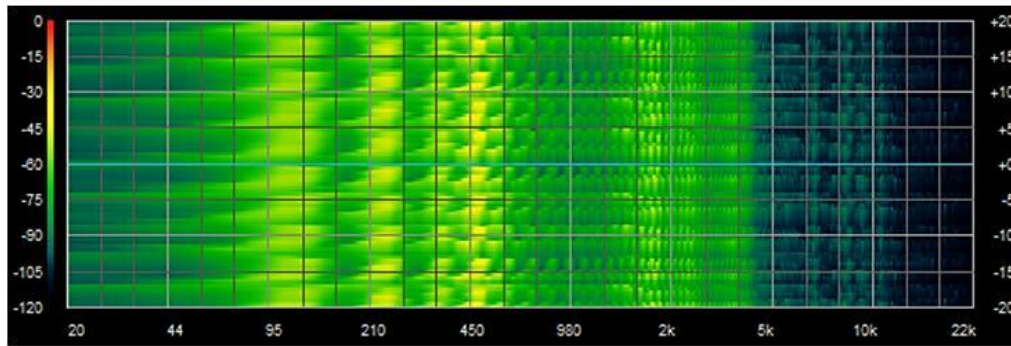


LOCUTOR 3
HOMBRE 23
AÑOS

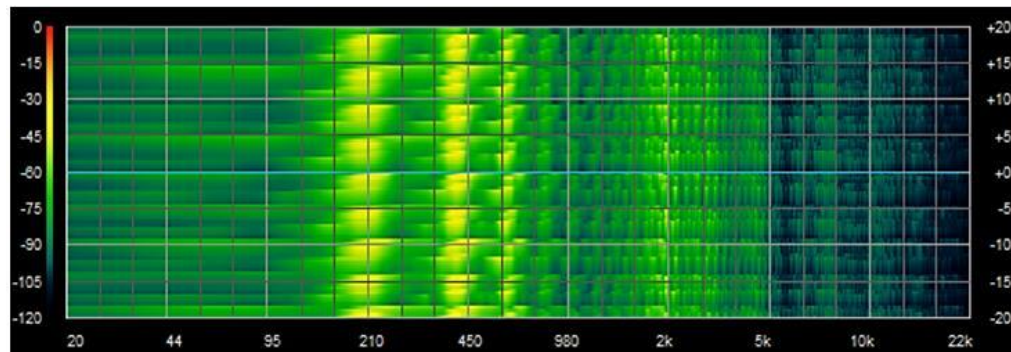
4. Espectrogramas.



LOCUTOR 1
NIÑO 12
AÑOS



LOCUTOR 2
MUJER 50
AÑOS



LOCUTOR 3
HOMBRE 23
AÑOS