



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

DISEÑO E IMPLEMENTACIÓN DE UNA UNIDAD DE PROCESAMIENTO DE  
SONIDO PARA OPTIMIZAR LA COMUNICACIÓN DE VOZ EN MÁSCARAS DE  
PROTECCIÓN RESPIRATORIA

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELECTRICISTA

ERIK ANDRÉS ATENAS ORELLANA

PROFESOR GUÍA:  
EDUARDO VERA SOBRINO

MIEMBROS DE LA COMISIÓN:  
MARCOS ORCHARD CONCHA  
MAURICIO CONTRERAS HERNÁNDEZ

SANTIAGO DE CHILE  
2012

## Resumen

RESUMEN DE LA MEMORIA  
PARA OPTAR AL TITULO DE  
INGENIERO CIVIL ELECTRICISTA  
POR: ERIK ATENAS O.  
PROF. GUIA: Sr. EDUARDO VERA S.

“DISEÑO E IMPLEMENTACIÓN DE UNA UNIDAD DE PROCESAMIENTO DE SONIDO PARA OPTIMIZAR LA COMUNICACIÓN DE VOZ EN MÁSCARAS DE PROTECCIÓN RESPIRATORIA”

Actualmente la industria se esfuerza por disminuir la tasa de enfermedades laborales, con especial énfasis en el caso de la silicosis. Para ello el principal elemento de control es la máscara de protección respiratoria. Sin embargo ésta, introduce otro problema: bloquea la salida clara de la voz e impide la comunicación. Además, como el ambiente está dominado por altos niveles ruido, los trabajadores optan por no usarla a la hora de transmitir los mensajes, perdiendo toda protección. En este contexto nace la idea de resolver esta disyuntiva: proteger y simultáneamente permitir la comunicación.

El objetivo de este trabajo es desarrollar un nuevo dispositivo para máscaras respiratorias, que permita comunicarse por voz sin la necesidad de quitársela. Estando además acoplado a los sistemas de radios portátiles, permite la comunicación sin tener que hacer uso del botón “push-to-talk”, liberando al operario de tener que ocupar sus manos al comunicarse.

Se propone una solución en forma de un detector de comandos de voz que active/desactive el canal. Para ello, se analiza la necesidad de mejorar la calidad del audio ante efectos como la respiración del operario y ruidos externos.

El trabajo se divide en cuatro etapas: cancelación de ruido, para disminuir los niveles de ruido externo que enmascaran la voz; filtrado de señales, para disminuir el paso de sonidos molestos que no estén en el rango de la voz; reconocimiento de voz, para crear un detector de comandos y finalmente, la activación del canal, con el propósito de abrir el canal de la radio con la voz reemplazando al botón “push-to-talk”.

Los resultados indican que es posible reducir hasta en 20 dB los niveles de ruido externo manteniendo los de la voz, bajo ruido ambiental de hasta 100 dB. El filtrado elimina considerablemente las frecuencias de audio no deseadas, pudiéndose aplicar en diferentes rangos volviendo el sistema adaptable. El detector de comandos en ambientes con bajos niveles de ruido funciona correctamente. Sin embargo, para aquellos en los cuales las condiciones de ruido son muy adversas los resultados no son tan positivos. La activación del canal de comunicación logra el objetivo final y ante un correcto reconocimiento abre la comunicación entre las radios.

Dado que el sistema de activación entrega la suficiente independencia como para ser integrado a cualquier dispositivo, se propone controlar otros sistemas de interés a través comandos de voz dictados desde la máscara. De todos modos, como punto de partida, debe considerar la mejora del sistema de reconocimiento por uno mas robusto que no falle ante condiciones de ruido extremo.

*... tus sueños pueden hacerse realidad  
sí tienes el coraje de perseguirlos.*

## **Agradecimientos**

Mis más sinceros agradecimientos a mi profesor guía Eduardo Vera por la oportunidad y confianza brindada. Al profesor Marcos Orchard por el tiempo y los consejos y a Mauricio Contreras por el apoyo la dedicación y las oportunidades. También agradecer a todo SoluNova y a la gente de Blumos que me han hecho sentir como en casa.

Quiero dar las gracias a mi familia que a pesar de pasar muchos momentos difíciles durante este periodo siempre me han dado la fuerza para seguir adelante y confiaron en que conseguiría la meta. A mi mamá que ha realizado un gran esfuerzo durante todo este tiempo para que sus hijos logren ser profesionales y ha luchado sola por sacar su familia adelante, ya la mitad de la tarea está completa... A mi hermana que ha dedicado mucho tiempo a ayudarme... A mi papá que siempre ha creído en mí. En especial quiero agradecer a dos personas que me hubiese encantado que pudiesen compartir este momento, mis tatas, sin ellos nada de esto hubiese sido igual y sé que siempre me están acompañando y que desde donde estén disfrutan de este logro.

También quiero dar las gracias una persona muy especial, que llegó en el momento que más lo necesitaba y que ha sido el apoyo más grande estos últimos años, ha sido mi motivación y mi pilar. Tú sabes todo lo que significas para mí y quiero que en este párrafo quede reflejado que éste resultado también es tuyo.

A mis compañeros, mis amigos y a toda la gente que conocí durante mi paso por la universidad, estén seguros que todos han aportado un granito de arena para esta meta. Siempre me he sentido un afortunado de la gente que he podido conocer, del apoyo y cariño que siento de todos.

Espero que esto no sea el término, sino que el comienzo de una nueva etapa donde pueda seguir cumpliendo mis sueños y que ojalá todas las personas que han sido importantes en mi vida lo sigan siendo por siempre.

## Índice de contenido

Portada .....	1
Resumen .....	2
Agradecimientos.....	4
Índice de contenido .....	5
Índice de tablas.....	7
Índice de Ilustraciones .....	8
1. Introducción.....	12
2. Revisión Bibliográfica.....	14
2.1. Aparato fonador: .....	14
2.1.1. Voz .....	15
2.1.2. La voz humana como señal.....	16
2.1.3. Entendiendo el habla .....	16
2.1.4. Factores que afectan la inteligibilidad en sistemas de sonido.....	17
2.2. Ruido .....	18
2.2.1. Parámetros básicos de las ondas sonoras .....	18
2.2.2. Nivel de presión acústica .....	20
2.2.3. Suma de niveles de presión acústica.....	21
2.2.4. Niveles de presión acústica ponderados .....	24
2.2.5. Presencia del ruido en el trabajo .....	25
2.2.6. Los efectos del ruido.....	26
2.2.7. Cancelación activa de ruido .....	27
2.3. Filtros.....	29
2.3.1. Tipos de filtros .....	29
2.4. Reconocimiento automático del habla (RAH) .....	36
2.5. Tipos de reconocimiento de voz .....	37
2.5.1. Palabras aisladas.....	38
2.5.2. Palabras relacionadas .....	38
2.5.3. Discurso continuo.....	38
2.5.4. Habla espontánea.....	38
2.6. Verificación/identificación de voz .....	38

2.6.1. Representación de señales de voz en reconocimiento.....	38
2.6.2. Condiciones adversas .....	40
2.6.3. Incorporación de información acústica fonética .....	41
2.7. Usos y aplicaciones de sistemas de RAH .....	41
2.7.1. Dictado .....	41
2.7.2. Mando y Control.....	42
2.7.3. Telefonía .....	42
2.7.4. Médico / Discapacidad.....	42
2.7.5. Teléfonos móviles.....	42
2.8. Radios portátiles .....	42
3. Implementación.....	43
3.1. Sistema cancelador de ruido .....	47
3.1.1. Diseño cancelador de ruido.....	47
3.2. Diseño filtros .....	55
3.3. Reconocimiento de voz .....	56
3.4. Activación canal de comunicación .....	59
4. Discusión de resultados .....	62
4.1. Resultados cancelación de ruido .....	62
4.2. Filtros.....	75
4.3. Reconocimiento .....	79
4.4. Activación Canal .....	82
5. Conclusiones .....	85
6. Bibliografía.....	87
7. Anexos .....	88
7.1. Código Arduino .....	88
7.2. Esquemático placa de desarrollo LMV1089 .....	91

## Índice de tablas

Tabla 2.2-1: Presión acústica según actividad .....	21
Tabla 3.1-1: Familia de integrados LMV.....	51

## Índice de Ilustraciones

Ilustración 2.1-1 Aparato Fonador: Conjunto del aparato fonador y respiratorio (Imagen: E. Martínez Cedrón, 1984).....	14
Ilustración 2.1-2 Espectros: En azul promedio hombres, en rojo promedio mujeres, en verde espectro ideal.....	16
Ilustración 2.1-3: Sistema de amplificación.....	17
Ilustración 2.2-1: Equivalencias para suma de decibeles.....	22
Ilustración 2.2-2: Diferencia 4dB.....	23
Ilustración 2.2-3: Diferencia 8,5dB.....	24
Ilustración 2.2-4: Curvas de ponderación.....	25
Ilustración 2.2-5: Suma de ondas desfasadas en tiempo.....	27
Ilustración 2.2-6: Suma de ondas desfasadas en 180°.....	28
Ilustración 2.2-7: Forma de onda palabra hablada.....	28
Ilustración 2.3-1: Estructura de un filtro.....	29
Ilustración 2.3-2: Filtro pasa bajos ideal.....	30
Ilustración 2.3-3: Filtro pasa bajos real.....	30
Ilustración 2.3-4: Filtro pasa altos ideal.....	31
Ilustración 2.3-5: Filtro pasa banda ideal.....	31
Ilustración 2.3-6: Filtro rechaza banda ideal.....	32
Ilustración 2.3-7: Filtros en cascada para formar pasa banda.....	32
Ilustración 2.3-8: Filtros en paralelo para formar pasa bandas.....	33
Ilustración 2.3-9: Parámetros de un filtro.....	33
Ilustración 2.3-10: Respuesta en frecuencia filtro de Butterworth.....	34
Ilustración 2.3-11: Filtro de Chebyshev.....	34
Ilustración 2.3-12: Respuesta filtro de Bessel.....	35
Ilustración 2.3-13: Comparación de respuestas de filtros.....	35
Ilustración 2.8-1: Diagrama de flujo desarrollo propuesto.....	43
Ilustración 2.8-2: Formatos de comunicación.....	44
Ilustración 2.8-3: Máscara de protección respiratoria Supreme.....	45
Ilustración 2.8-4: Pieza frontal.....	45
Ilustración 2.8-5: Modelo del sistema.....	46
Ilustración 2.8-6: Sistema intercomunicador.....	46
Ilustración 3.1-1: Circuito diseñado versión 1.....	48



Ilustración 3.1-2: Diseño PCB cancelador de ruido v1.....	48
Ilustración 3.1-3: Esquema funcionamiento LMV 1089 .....	49
Ilustración 3.1-4: Diagrama de bloques LMV1089.....	50
Ilustración 3.1-5: Placa de desarrollo LMV1090 .....	52
Ilustración 3.1-6: Conexión placa de desarrollo PC .....	52
Ilustración 3.1-7: LMV1090 Software .....	53
Ilustración 3.1-8: Sony Sound Forge Pro 10.....	53
Ilustración 3.1-9: PCB LMV1089 v1.....	54
Ilustración 3.1-10: Modelos de tapas a utilizar .....	54
Ilustración 3.2-1: Esquemático Filtro LTC .....	55
Ilustración 3.2-2: Layout filtro LTC .....	56
Ilustración 3.3-1: Easy VR.....	56
Ilustración 3.3-2: EasyVR Shield para Arduino UNO .....	57
Ilustración 3.3-3: EasyVR Commander .....	58
Ilustración 3.4-1: Descripción pines de acceso radio Motorola pro5150 .....	60
Ilustración 3.4-2: Esquema conexión sistema activación del canal de comunicación.....	61
Ilustración 4.1-1: Arriba ruido blanco con voz; abajo ruido blanco con voz y cancelación de ruido .....	62
Ilustración 4.1-2: Comparación ruido blanco voz con y sin cancelación de ruido; en negro sin cancelación; en gris la muestra con cancelación.....	63
Ilustración 4.1-3: Arriba ruido rosa con voz, abajo ruido rosa con voz y cancelación de ruido .....	63
Ilustración 4.1-4: Comparación ruido rosa voz con y sin cancelación de ruido; en negro sin cancelación y en gris la muestra con cancelación.....	64
Ilustración 4.1-5: Arriba reducción de ruido para ruido blanco; abajo reducción de ruido para ruido rosa.....	64
Ilustración 4.1-6: Espectro de frecuencias muestra de ruido nº1 .....	65
Ilustración 4.1-7: Comparación MR1 18dB pre amplificación 6DB post amplificación sin máscara, en negro sin cancelación de ruido en gris con cancelación.....	66
Ilustración 4.1-8: Comparación MR1 36dB pre amplificación 6DB post amplificación sin máscara; en negro sin cancelación de ruido; en gris con cancelación.....	66

Ilustración 4.1-9: Comparación MR1 36dB pre-amplificación 12dB post amplificación sin máscara; en negro sin cancelación de ruido; en gris con cancelación.....	67
Ilustración 4.1-10: Vista frontal y lateral ubicación micrófonos muestra tapa nº1.....	67
Ilustración 4.1-11: Prueba MR1 con tapa MT1 22 pre amplificación y 6 post amplificación; arriba izquierda sin cancelación; arriba derecha con cancelación; abajo comparación, en gris con cancelación en negro sin cancelación.....	68
Ilustración 4.1-12: Pruebas distintas posiciones micrófonos y modelos tapas .....	69
Ilustración 4.1-13: Placa LMV1089 fabricación propia.....	70
Ilustración 4.1-14: Salida placa diseño propio .....	71
Ilustración 4.1-15: Muestra ruido mina .....	71
Ilustración 4.1-16: Comparación MR1 con MRM; en azul claro MRM, en azul oscuro MR1.....	72
Ilustración 4.1-17: Comparación MRM 6dB pre amplificación 6dB post amplificación sin máscara; en negro con cancelación de ruido; en gris sin cancelación.....	73
Ilustración 4.1-18: Comparación MRM con cancelación; de abajo hacia arriba 6, 12, 18 y 22 dB .....	73
Ilustración 4.1-19: Vistas esquemáticas frontal y lateral nueva pieza .....	74
Ilustración 4.1-20: Comparación MRM con pieza modificada, 6 pre amplificación 6 post amplificación; en gris sin cancelación, en negro con cancelación.....	75
Ilustración 4.2-1: MR1 con filtro pasabajos 1Khz pre amplificación 22dB post amplificación 6 dB; arriba izquierda sin cancelación; arriba derecha con cancelación; abajo comparación .....	76
Ilustración 4.2-2: MR1 con filtro pasabajos 1Khz pre amplificación 22dB post amplificación 6 dB; verde sin cancelación; azul sin cancelación y con filtro pasa bajos.....	77
Ilustración 4.2-3: MR1 con filtro pasabajos 1Khz pre-amplificación 22dB post-amplificación 6 dB; azul claro sin cancelación; naranja sin cancelación y con filtro pasa bajos.....	77
Ilustración 4.2-4: Filtro LTC .....	78
Ilustración 4.2-5: Resultados filtro pasa bajos; arriba izquierda 39,9 Hz 156 mV; arriba derecha 196 Hz 152 mV; abajo derecha 690 Hz 117 mV; abajo derecha 1050 Hz 45 mV .....	78

Ilustración 4.3-1: Entrenamiento comando .....	79
Ilustración 4.3-2: Resultados reconocimiento .....	81
Ilustración 4.3-3: Captura osciloscopio activación y desactivación salida DigitalOut11 .....	81
Ilustración 4.4-1: Circuito abierto, canal desactivado .....	82
Ilustración 4.4-2: Circuito cerrado, canal abierto .....	82
Ilustración 4.4-3: Resultados simulación activación canal .....	83
Ilustración 4.4-4: Montaje circuito activación .....	84
Ilustración 7.2-1: Esquemático cancelador.....	91

## 1. Introducción

Los trabajadores continuamente se mueven en entornos riesgosos; de hecho muchos accidentes laborales ocurren diariamente, algunos incluso con desenlace fatal. Es por ello que existen constantes esfuerzos en desarrollar elementos que protejan de estos riesgos.

De acuerdo a la legislación vigente, se entiende por elemento de protección personal (EPP) “todo equipo, aparato o dispositivo especialmente proyectado y fabricado para preservar el cuerpo humano, en todo o en parte, de riesgos específicos de accidentes del trabajo o enfermedades profesionales”<sup>1</sup>.

Es aquí donde el desarrollo de nuevas tecnologías juega un importante rol. Las industrias invierten grandes cantidades de dinero con el fin de que los EPP sean de mejor calidad y a la vez provean la mayor protección para sus trabajadores. Cabe destacar que la empresa no puede cobrarle a sus trabajadores por estos elementos<sup>2</sup>, de manera que su costo debe asumirse completamente, entonces además se debe considerar este factor como una restricción.

Este trabajo se centra en uno de los elementos más usados, el protector respiratorio. Existen diversos sistemas que proveen esta protección, ya sea algo tan básico como un pañuelo puesto sobre la boca y la nariz hasta sistemas bastante complejos como mascarillas antigases o submarinas. Sin embargo a pesar de traer un beneficio, agregan un problema, y es que en la mayoría de los casos dificultan la comunicación, pues impiden la salida limpia de la voz, ya sea distorsionándola o disminuyendo el volumen. Es común ver en industrias como la minería a los trabajadores que estando en ambientes donde es estrictamente necesario el uso de la mascarilla, evitan su uso con el fin de comunicarse oralmente.

El objetivo general de este trabajo es desarrollar un nuevo dispositivo para máscaras respiratorias, que permita comunicarse por voz sin la necesidad de quitársela. Estando además acoplado a los sistemas de radios portátiles, permitiendo la comunicación sin tener que hacer uso del botón “push-to-talk”, liberando al operario de tener que ocupar sus manos al comunicarse. Para lo que se propone como objetivos específicos lograr una solución en forma de un detector de comandos de voz que active/desactive el canal y que considere un dispositivo que mejore la calidad del audio de la comunicación ante efectos como la respiración del operario y ruidos

---

<sup>1</sup> Art. 1º del Decreto Supremo Nº173, de 1982, del Ministerio de Salud de Chile.

<sup>2</sup> Ley 16.744 sobre Accidentes del Trabajo y Enfermedades Profesionales, Artículo nº 68

externos, que sea de bajo costo, reducido tamaño y que considere la tanto la posibilidad de patentarse como las opciones de comercialización.

La memoria se divide en 7 capítulos, incluido el correspondiente a esta introducción. El capítulo 2 contiene información sobre los temas a tratar y que permite comprender de mejor forma los términos o conceptos utilizados, entre ellos: el funcionamiento del aparato fonador, la definición de ruido, sus principales características y cómo influye en la vida laboral, algunos tipos de sistemas de reconocimiento de voz y formas de verificación e identificación de la voz, además de usos y aplicaciones típicas de estos sistemas. En el capítulo 3 se detalla la implementación del dispositivo, para lo que se divide el trabajo en 4 etapas: diseño de un sistema cancelador de ruido, filtrado de señales, reconocimiento de voz y activación del canal de la radio. En cada uno se intenta resolver tanto las características técnicas como los requerimientos mencionados en los objetivos específicos. En el capítulo 4 se muestran los resultados obtenidos de las pruebas realizadas y se analiza las influencias que tienen en el resultados general y como podría mejorarse los sistemas desarrollados. El capítulos 5 contiene las conclusiones del trabajo y plantea desafíos futuros al sistema. El capítulo 6 y 7 contienen la bibliografía y los anexos respectivamente.

## 2. Revisión Bibliográfica

Para desarrollar un nuevo sistema de detección de comandos de voz es necesario conocer de antemano los factores que influyen en la comunicación oral, como son el aparato fonador y la articulación de las palabras o la respiración del operario, y para el caso específico del sistema en estudio, las influencias ambientales como el ruido externo, la humedad, los golpes y las vibraciones.

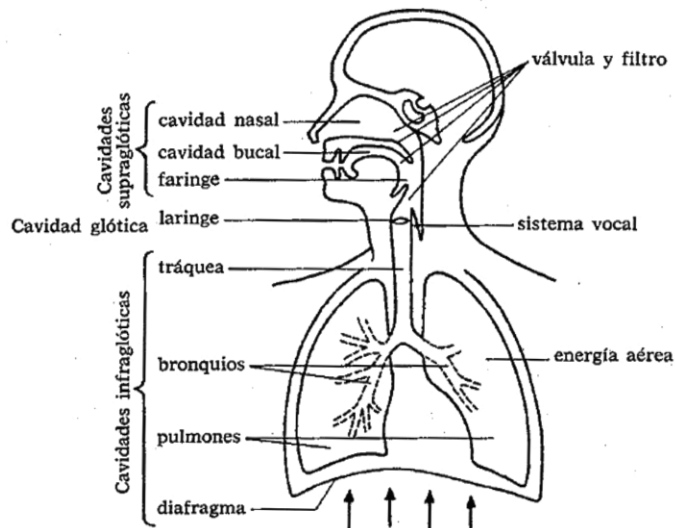
### 2.1. Aparato fonador [1] [2]:

A grandes rasgos el aparato fonador es un conjunto de órganos implicados en la producción del habla, los cuales no están específicamente diseñados para esta tarea pues todos ellos cumplen otra función (respiración, masticación, etc.). La evolución humana ha sido decisiva en la especialización de estos órganos.

El aparato fonador se divide en 3 partes:

- Cavidades infra glóticas
- Cavidad Glótica
- Cavidades supra glóticas

Cada una de ellas es responsable de alguno de los procesos implicados en la producción del sonido en el habla.



**Ilustración 2.1-1 Aparato Fonador: Conjunto del aparato fonador y respiratorio (Imagen: E. Martínez Cedrón, 1984)**

Los órganos que componen el aparato fonador se observan en la Ilustración 2.1-1 Aparato Fonador, donde las funciones de cada una de las partes se pueden resumir de la siguiente forma:

- Cavidades infra glóticas: Proporcionan el aire necesario para producir los sonidos del habla, son los órganos responsables de la intensidad de la voz y la duración de la misma.
- Cavity glótica: Es el elemento más importante del aparato fonador, pues contiene los repliegues vocales, los que si se encuentran juntos la corriente de aire que proviene de los pulmones los pone en vibración, lo que se conoce como producción de voz. La forma en que estos repliegues vibran dan origen a los diferentes tipos de fonación o cualidades de la voz.
- Cavity supra glótica: Se encargan de la articulación, para producir diferentes sonidos. La disposición de estos órganos modifica el tamaño y forma de las cavidades lo que produce formas distintas de resonancia, esto es decisivo a la hora de obtener un sonido u otro. También actúan como un filtro, amplificando o atenuando determinados componentes de frecuencia de la onda sonora producida.

### **2.1.1. Voz**

La voz se produce voluntariamente por el aparato fonatorio (ver sección 2.1), el cual por su conformación permite dar características distintas a los sonidos del habla durante su emisión. Sin embargo existen algunas de ellas que son convenientes de conocer a la hora de realizar un sistema de comunicación.

- Oralidad y nasalidad: Los fonemas en los que el aire pasa por la cavidad nasal se denominan nasales, y aquellos en los que sale por la boca orales. Esta diferencia viene dada por el resonador principal que los produce (cavidad nasal u oral respectivamente). En castellano los sonidos nasales son las consonantes m, n y ñ. Esto puede ser un punto de importancia a la hora de implementar el sistema en las mascarillas dado que muchas veces el ajuste de estas impide que los sonidos nasales puedan ser ejecutados con normalidad.
- Tonalidad: Aquellos fonemas que son producidos por la vibración de las cuerdas vocales se denominan tonales. Esta característica produce en los sonidos generados un espectro cuasi periódico<sup>3</sup>. Todas las vocales son tonales, pero existen varias consonantes que también tienen esta característica, por ejemplo b, d y m.

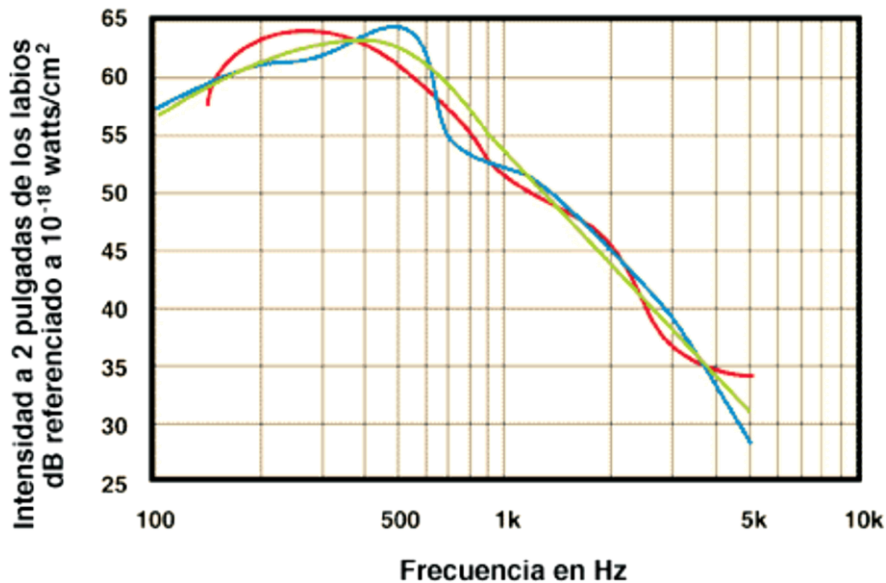
---

<sup>3</sup> Son periódicos cuando el hablante hace un esfuerzo por emitir un sonido sin inflexiones.

### 2.1.2. La voz humana como señal [3]

La voz humana es una onda continua con una frecuencia fundamental en el rango de los 100 a 400 Hz, aunque en promedio el hombre esta más cerca de los 100Hz y la mujer de los 400Hz. Los armónicos de estas señales se denominan formantes son determinados por las características de resonancia de la región que los produce. Los sonidos consonantes por ejemplo están en el rango de 2 a 9 kHz.

Las vocales cuya duración promedio esta entre 30 y 300 ms son las responsables de la potencia sonora. En cambio la inteligibilidad es producida por las consonantes con duración promedio de 10 a 100 ms. La potencia de la señal varía como un todo, pero la potencia de cada rango de frecuencia varia en función de los cambios de los formantes<sup>4</sup>.



**Ilustración 2.1-2 Espectros: En azul promedio hombres, en rojo promedio mujeres, en verde espectro ideal.**

### 2.1.3. Entendiendo el habla

Los intervalos de sonido pueden no corresponder directamente a divisiones de palabras o sílabas, dado que la mayoría de estos eventos son discontinuos, se combinan y se superponen en el tiempo además la articulación de un fonema difiere entre cada orador. Hasta ahora no se tiene

<sup>4</sup> Un formante consiste en una alta concentración de energía que se da en una determinada banda de frecuencias.



una explicación clara para como el oído decodifica el habla continua. Existen diversos factores que afectan la percepción y comprensión del habla, el ruido, la duración y el contenido espectral son algunos de ellos.

En ese sentido la pérdida de inteligibilidad está asociada a la pérdida de información producto de la participación de todos los factores influyentes. Los ruidos pueden enmascarar la voz, la dirección desde la que se genera el habla y el ruido también pueden alterar la percepción del mensaje. Otra de las cosas que influye a la hora de entender lo que se habla es lo predecible del mensaje

#### 2.1.4. Factores que afectan la inteligibilidad en sistemas de sonido

El objetivo de un sistema destinado a reproducir la voz humana es hacerlo con la mayor claridad posible y así facilitar la comprensión. Si consideramos que la señal de la voz humana es complicada por si misma, lograr una alta calidad al amplificarla y reproducirla es aún más complejo.

Normalmente un sistema de refuerzo sonoro esta compuesto por un micrófono, un preamplificador, un amplificador de potencia y un emisor (comúnmente un parlante); los que se ven afectados por los factores antes mencionados además de una serie de ruidos propios de los sistemas electrónicos, como son los ruidos térmicos o la saturación de los componentes (ver Ilustración 2.1-3: Sistema de amplificación).

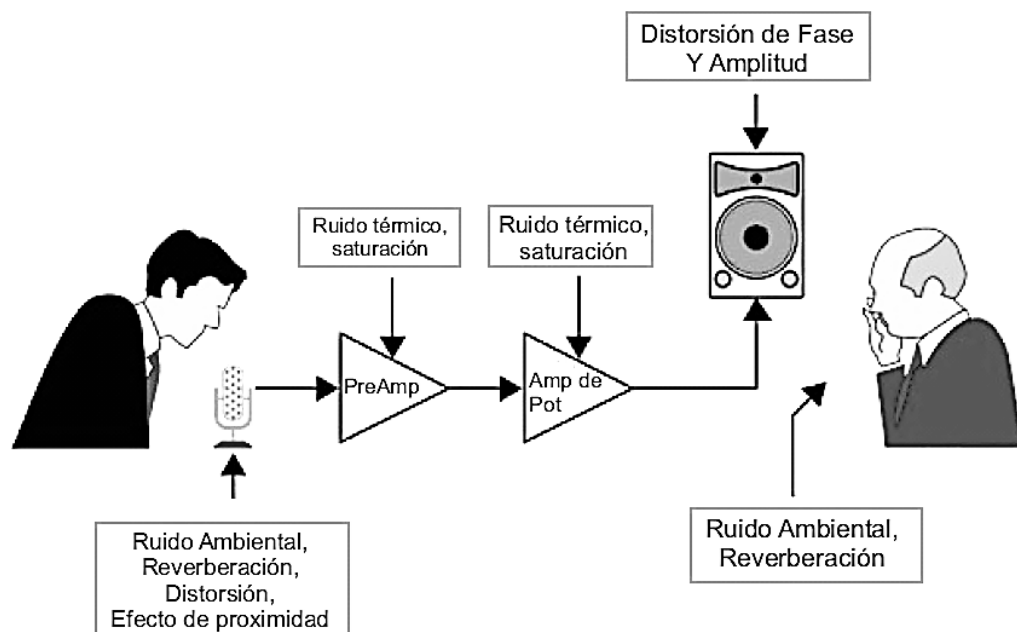


Ilustración 2.1-3: Sistema de amplificación

## **2.2. Ruido [4]**

Se puede definir ruido como cualquier sonido no deseado que interfiera en la recepción de otro sonido. El ruido acústico está formado por una mezcla de ondas sonoras de distintas frecuencias y amplitudes, esto produce que el resultado sea un sonido que conjuga frecuencias fundamentales con los armónicos que los acompañan.

El ruido acústico puede ser clasificado, al igual que otros ruidos según su intensidad y la frecuencia.

Tipos de ruido según la intensidad:

- Ruido Fluctuante: Es aquel cuya intensidad varía a lo largo del tiempo, ya sea de manera periódica o aleatoria.
- Ruido Impulsivo: Es aquel cuya intensidad aumenta bruscamente durante un impulso, la duración de este es breve en comparación al tiempo entre los impulsos.

Tipos de ruido según frecuencia:

- Ruido blanco: Se denomina así por la asociación con la luz blanca, tiene una distribución uniforme en el espectro audible (20 a 20kHz). Su respuesta en frecuencia es plana, es decir su intensidad es constante en todas las frecuencias.
- Ruido rosado: A diferencia del ruido blanco no tiene respuesta en frecuencia plana, su intensidad cae 3 decibeles por octava.
- Ruido marrón: Es un ruido poco común sin embargo existe, esta compuesto principalmente por ondas graves y medias.

### **2.2.1. Parámetros básicos de las ondas sonoras [5]**

Para describir una onda sonora existen una serie de parámetros básicos, entre los que se encuentran la frecuencia, el periodo, la velocidad del sonido, longitud de onda y la presión acústica.

- Frecuencia (f): Es el número de variaciones de presión por segundo, determina el tono grave o agudo de un sonido, su unidad de medida es el Hertz (Hz). El oído humano es capaz de reconocer sonidos en el rango de los 20 a los 20000 Hz.
- Periodo (T): Es el inverso de la frecuencia e indica el tiempo en que se produce un ciclo completo de onda. Su unidad es el segundo.
- Velocidad del sonido: El sonido no se propaga en vacío, necesita un medio para propagarse. Su velocidad depende de las características propias del medio, para el aire en condiciones normales de presión y

temperatura es de 344 m/s, en el agua 1500 m/s y en el hierro 3700 m/s.

- Presión acústica (P): Cuando una onda sonora se propaga en cualquier medio genera una variación de la presión medida sobre la presión atmosférica, su valor puede ser representado por:

$$p(t) = P_0 \cdot \sin(\omega t) \quad (1)$$

Donde  $p(t)$  es la presión instantánea y  $\omega$  la frecuencia angular.

Para el caso de los temas en estudio, esta forma no es muy útil, dado que en muchas aplicaciones de audio necesitamos un único valor para expresar los niveles de presión y el valor medio de la presión puede ser más interesante.

$$P_{medio} = \frac{1}{T} \cdot \int_0^T p(t) dt \quad (2)$$

Si se sustituye el valor de  $p(t)$ , en (2) se tiene:

$$P_{medio} = \frac{1}{2\pi} \cdot \int_0^{2\pi} P_0 \cdot \sin(\omega t) dt = 0 \quad (3)$$

Este resultado no nos entrega información, pero si se considera el valor eficaz:

$$P_{rms} = \sqrt{\frac{1}{T} \cdot \int p^2(t) dt} \quad (4)$$

Se tiene:

$$P_{rms} = \sqrt{\frac{1}{2\pi} \int_0^{2\pi} (P_0 \sin(\omega t))^2 dt} = \frac{P_0}{\sqrt{2}} \quad (5)$$

Entonces desde ahora se trabajara con la presión acústica como presión eficaz.

Cabe destacar que la energía sonora es proporcional al cuadrado del valor eficaz de la presión, por lo que al momento de sumar energías debe considerarse que están relacionadas con la suma de los cuadrados de las presiones efectivas.

### 2.2.2. Nivel de presión acústica

En el SI<sup>5</sup> la unidad de medida de presión acústica es el Pascal

$$P = \frac{\text{Fuerza}}{\text{Superficie}} = \frac{\text{Newton}}{\text{m}^2} = \text{Pascal} \quad (6)$$

El nivel mas bajo de presión acústica que somos capaces de detectar son los  $20 \cdot 10^{-6}$  [Pascal] y el valor límite son los 200 [Pascal] que es cuando la presión causa dolor y se hace insoportable. Como el rango es muy extenso se normaliza a una nueva unidad llamada decibelio.

$$L_p(\text{dB}) = 10 \log \left[ \frac{P_{rms}}{P_0} \right]^2 \quad (7)$$

Donde  $L_p$  es el nivel de presión acústica en dB,  $P_{rms}$  el valor eficaz de la presión acústica en Pascales y  $P_0$  la presión de referencia  $20 \cdot 10^{-6}$  [Pascales]

Con la definición anterior podemos determinar en dB los valores umbral y límites para la audición:

$$\text{Umbral: } L_{umbral} = 10 \log \left[ \frac{20 \cdot 10^{-6}}{20 \cdot 10^{-6}} \right]^2 = 0 \text{ dB} \quad (8)$$

$$\text{Límite: } L_{límite} = 10 \log \left[ \frac{200}{20 \cdot 10^{-6}} \right]^2 = 140 \text{ dB} \quad (9)$$

Con esta escala se tiene una forma mas reducida de representar los ruidos por niveles. La Tabla 2.2-1 muestra los niveles de ruido en dB para distintas actividades de la vida cotidiana.

---

<sup>5</sup> Sistema Internacional de medidas

**Tabla 2.2-1: Presión acústica según actividad**

<b>Nivel en <math>\mu</math>Pascales</b>	<b>Nivel en dB</b>	<b>Actividad</b>
<b>200.000.000</b>	140	Aeropuerto
<b>20.000.000</b>	120	Sala de compresores
<b>2.000.000</b>	100	Martillos neumáticos
<b>200.000</b>	80	Calle con tráfico
<b>20.000</b>	60	Oficinas
<b>2.000</b>	40	Biblioteca
<b>200</b>	20	Zona rural aislada
<b>20</b>	0	Umbral de la audición

Si bien es cierto esta escala cuenta con ventajas, también posee varias desventajas. Una de ellas es que se puede menospreciar niveles que si bien son altos por la normalización parecen numéricamente bajos. Por ejemplo una diferencia de 3 dB implica el doble de la energía sonora.

Numéricamente, si tenemos  $L_1$  nivel de presión inicial,  $L_2$  nivel de presión final y  $L_1 - L_2 = 3 \text{ dB}$

$$L_1 = 10 \log \left[ \frac{P_1^2}{P_0^2} \right] \text{ De donde } P_1^2 = P_0^2 \cdot 10^{\frac{L_1}{10}}$$

$$L_2 = 10 \log \left[ \frac{P_2^2}{P_0^2} \right] \text{ De donde } P_2^2 = P_0^2 \cdot 10^{\frac{L_2}{10}}$$

Luego

$$\frac{P_1^2}{P_2^2} = 10^{\frac{L_1 - L_2}{10}} = 10^{\frac{3}{10}} \approx 2 \quad (10)$$

### **2.2.3. Suma de niveles de presión acústica**

Los niveles de presión sonora no se suman de manera lineal, esto porque se adoptó una transformación logarítmica para transformar desde pascales a dB. Existen 2 métodos, uno es simplemente mediante remplazo de los valores en las formulas presentadas y otro un procedimiento gráfico.

Por ejemplo si se quiere sumar 86, 90 y 82 dB y sabiendo que  $P_0 = 20 \cdot 10^{-6}$  [Pascal] se tendrá:

$$L_1 = 10 \log \left[ \frac{P_1^2}{P_0^2} \right] \rightarrow P_1^2 = P_0^2 \cdot 10^{\frac{L_1}{10}} \rightarrow P_1^2 = (20 \cdot 10^{-6})^2 \cdot 10^{\frac{86}{10}} = 0,1592428$$

$$L_2 = 10 \log \left[ \frac{P_2^2}{P_0^2} \right] \rightarrow P_2^2 = P_0^2 \cdot 10^{\frac{L_2}{10}} \rightarrow P_2^2 = (20 \cdot 10^{-6})^2 \cdot 10^{\frac{90}{10}} = 0,4$$

$$L_3 = 10 \log \left[ \frac{P_3^2}{P_0^2} \right] \rightarrow P_3^2 = P_0^2 \cdot 10^{\frac{L_3}{10}} \rightarrow P_3^2 = (20 \cdot 10^{-6})^2 \cdot 10^{\frac{82}{10}} = 0,0633957$$

Con lo que

$$P_1^2 + P_2^2 + P_3^2 = 0,6226385$$

Entonces

$$L_{suma} = 10 \log \left[ \frac{0,6226385}{(20 \cdot 10^{-6})^2} \right] = 91,92 \text{ dB} \quad (11)$$

La segunda forma de sumar decibeles es mediante el método gráfico para realizarlo primero se deben conocer las diferencias lineales entre los niveles de ruido y luego obtener su valor en decibeles según la siguiente gráfica:

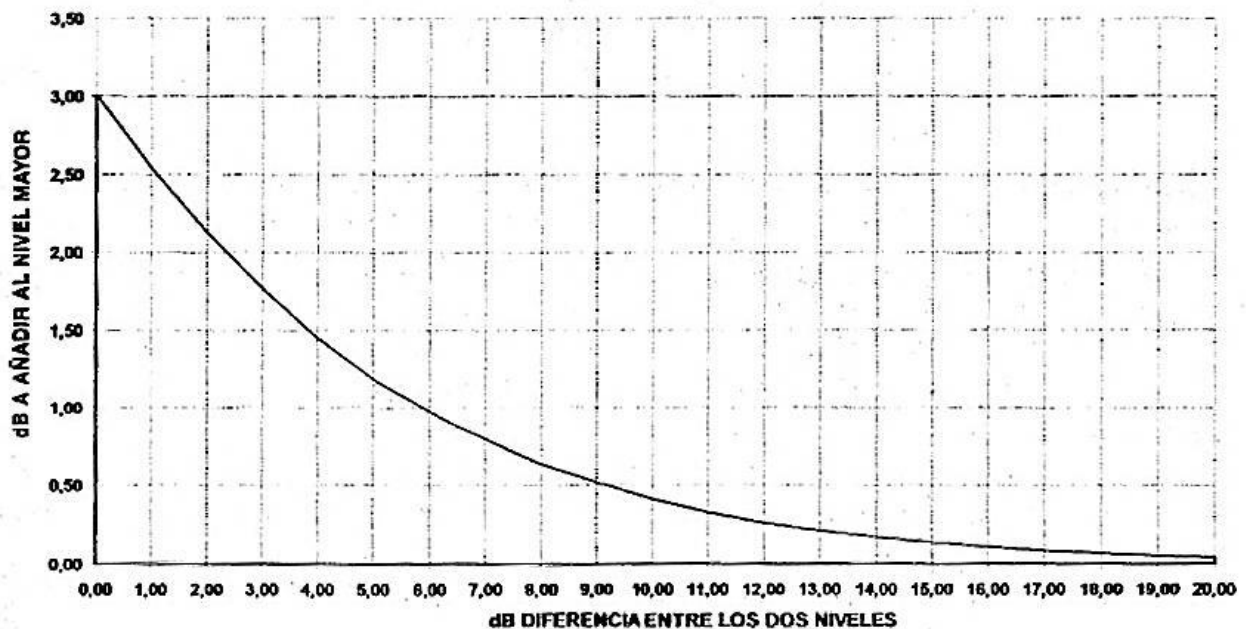


Ilustración 2.2-1: Equivalencias para suma de decibeles

Una vez obtenido el valor en decibeles para la diferencia, se lo sumamos al nivel más alto de los sumandos y obtenemos el valor de suma entre ambos niveles de ruido en dB.

Para el ejemplo anterior el proceso es el siguiente:

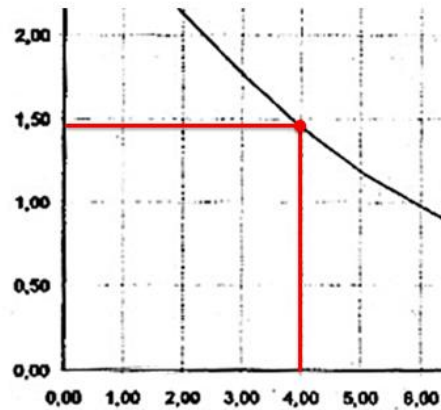
$$L_1 = 86 \text{ dB}$$

$$L_2 = 90 \text{ dB}$$

$$L_3 = 82 \text{ dB}$$

$$L_2 - L_1 = 90 - 86 = 4$$

Si vemos la tabla



**Ilustración 2.2-2: Diferencia 4dB**

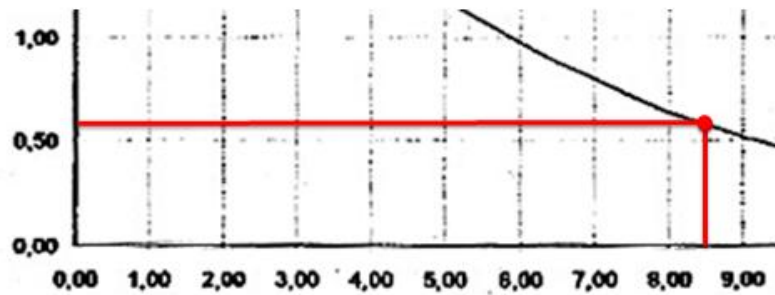
Encontramos que para una diferencia de 4 se suman aproximadamente 1,4 dB al mayor sumando:

Entonces:

$$L_n = L_2 + L_1 = 90 + 1,4 = 91,4 \text{ dB} \quad ( 12 )$$

Faltaría sumarlo al tercer valor para lo cual realizamos el mismo procedimiento

$$L_n - L_3 = 91,4 + 83 = 8,5$$



**Ilustración 2.2-3: Diferencia 8,5dB**

Del gráfico obtenemos que se debe sumar aproximadamente 0,6 al sumando mas alto, con esto

$$L_n + L_3 = 91,4 + 0,6 = 92 \text{ dB} \quad ( 13 )$$

Si comparamos con el resultado obtenido con el calculo matematico estos difieren en 0,08 DB, lo cual es bastante aceptable considerando la simpleza del segundo método.

#### **2.2.4. Niveles de presión acústica ponderados [6]**

La percepción del sonido por el oído humano es un fenómeno complejo de explicar que depende de la frecuencia y el nivel de presión sonora sin que exista alguna relación lineal entre ellos.

Si se desea analizar los riesgos de la exposición del ruido en el trabajo se debe lograr alguna forma que represente como percibe el ruido el trabajador.

Para ello se creó 4 escalas de ponderación denominadas A,B,C y D que son aceptadas internacionalmente y permiten a los instrumentos de medición corregir sus lecturas adaptándolas a la respuesta del oído.

Cada una de estas escalas están diseñadas para representar la respuesta ante distintos niveles de presión sonora, la A niveles bajo los 55dB, la B niveles intermedios 55 a 85 dB y la C niveles altos sobre los 85 dB. La escala D esta pensada para niveles muy altos de presión sonora.

Actualmente se ha llegado a la conclusión que la escala A es la que representa de mejor forma la respuesta real del oído ante la peligrosidad para el órgano auditivo, por esto se ha adoptado mundialmente como la norma.



Para diferenciar en que escala se encuentra la medida, se agrega la letra de la ponderación a la unidad dB, definiendo como dBA, la unidad de presión sonora según la ponderación A, dBB para la escala B, dBC en el caso de la C y dBD en la D.

La forma de las curvas que representan las escalas las podemos observar en la Ilustración 2.2-4:

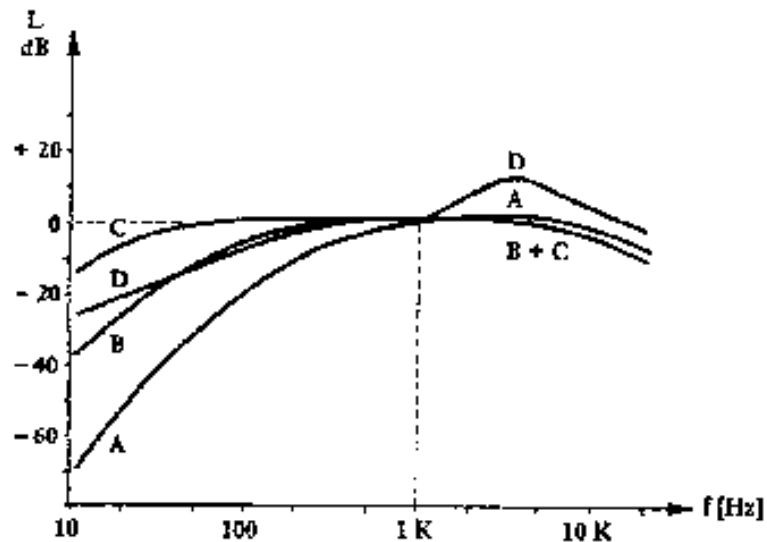


Ilustración 2.2-4: Curvas de ponderación

### 2.2.5. Presencia del ruido en el trabajo [4]

El ruido es uno de los peligros más comunes en el trabajo. Estadísticas indican que millones de trabajadores en el mundo se ven expuestos a niveles muy altos de ruido, por ejemplo en Estados Unidos 9 millones de trabajadores están sometidos a un promedio de 85dBA y 5,2 millones trabajan en entornos más ruidosos aún, esto es alrededor del 35% de personas del sector fabricación.

Identificar estos niveles es fácil y en muchos casos controlarlos también, para ello se utilizan distintas técnicas como remodelar equipos, instalar sistemas de aislación o modificar procesos. Estas soluciones en algunos casos pueden tener elevados costos, sobre todo cuando se quiere bajar de los 85dBA por lo que algunas industrias no hacen esfuerzos por implementarlas.

## **2.2.6. Los efectos del ruido**

La pérdida de la capacidad auditiva es uno de los efectos perjudiciales del ruido mas conocidos, pero no es el único. Acufenos (sensación de zumbido en los oídos), interferencia en la comunicación hablada y en la percepción de las señales de alarma, alteraciones del ánimo y del rendimiento. Las protecciones auditivas por ende no solo sirven para proteger contra el ruido sino que también favorecen y mejoran el ambiente laboral.

### **2.2.6.1. Deterioro auditivo**

El ruido del trabajo no es la única causa de pérdida auditiva, existen otras fuentes extra laborales que la producen y cuyos efectos son difíciles de diferenciar. Generalmente estas fuentes son denominadas socioacusias, música a alto volumen, disparo de armas de fuego, ruido de motores de automóviles son algunas de ellas.

Es así como la exposición a ruidos extra laborales resulta importante pues se suma a la que se produce dentro del trabajo, y ocupar protectores puede influir de sobremanera en reducir la perdida auditiva a futuro.

Uno de los primeros efectos que se produce son los acufenos, que en palabras simples puede ser descrito como la sensación de zumbido en los oídos, el cual puede ser muy suave a veces como severo en otras. Normalmente un daño alto se detecta al sentir acufenos en ambientes silenciosos. Esto indica que se tiene un daño a las células sensoriales del oído interno y es predictor de una perdida auditiva por ruido.

### **2.2.6.2. La interferencia con la comunicación y la seguridad**

Un punto a destacar es la interferencia que provoca el ruido en la comunicación, puede enmascarar la comunicación hablada y las señales de alarma. Muchos procesos industriales pueden ejecutarse sin problemas con un mínimo de comunicación entre los trabajadores pero otros, como los realizados por pilotos, conductores de carros blindados y vehículos pesados, trabajadores de faenas mineras, entre otros, dependen en gran medida de la comunicación hablada. Estudios han demostrado que con niveles de ruido por sobre los 80 dBA es necesario hablar muy alto y si supera los 85 dBA se debe gritar. Para casos en que los niveles se sitúan por encima de los 95dBA, además de los efectos anteriores, los hablantes deben encontrarse a distancias muy cortas.

Los trabajadores reconocen además que los protectores tanto auditivos como respiratorios no solo atenúan señales externas también entorpecen la percepción de las palabras y de las alarmas, por ello es interesante agregar sistemas que no solo permitan mejorar las características protectoras sino

que además mediante dispositivos electrónicos cancelen o reduzcan el ruido y amplifiquen la voz.

### **2.2.7. Cancelación activa de ruido [7]**

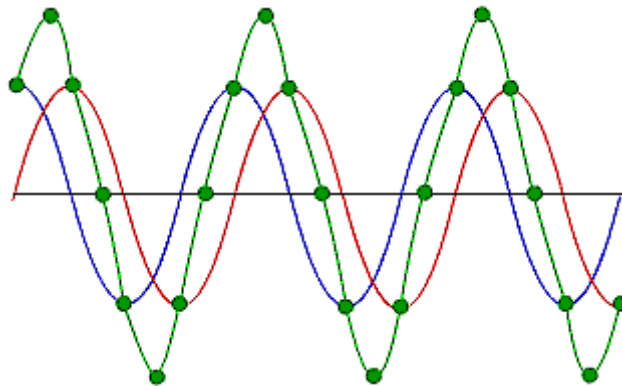
La cancelación activa del ruido es un método encargado de reducir o cancelar completamente los ruidos no deseados. Se le llama activa porque participan en ella elementos electrónicos activos que la realizan en tiempo real.

La cancelación de ruido se basa en un principio bastante simple, la interferencia de ondas, más específicamente la destructiva. Cuando se tiene un sonido de frecuencia única, basta con sumar la misma con su fase invertida. El uso de esta tecnología en dispositivos de audio es nueva y poco utilizada, se puede encontrar principalmente en audífonos.

#### **2.2.7.1. Cancelación de onda simple**

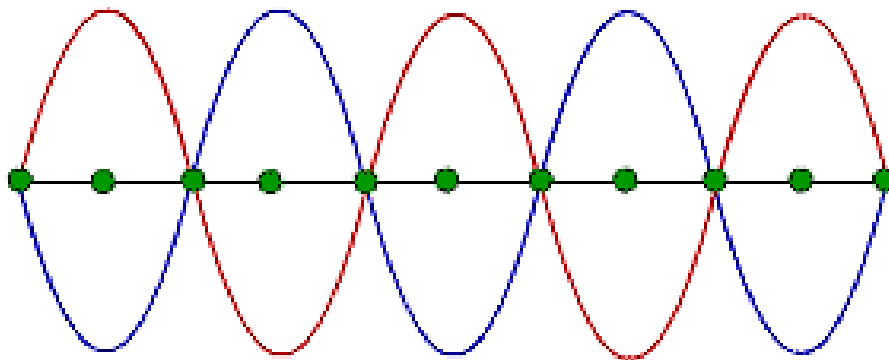
A pesar de que el sonido es una onda de compresión, cada tono o frecuencia puede ser representado por una onda sinusoidal de longitud dada. Esta representación nos permite el análisis de las interferencias en el sonido.

Si tenemos dos ondas con la misma frecuencia y longitud pero desfasadas en el tiempo, la suma de ambas nos entrega una onda que tiene puntos en donde la onda tiene una amplitud igual al doble de la que inicialmente poseían ambas por serado. Gráficamente sería:



**Ilustración 2.2-5: Suma de ondas desfasadas en tiempo**

Ahora si se realiza el mismo ejercicio pero con una de las ondas desfasada en  $180^\circ$  se producirá el efecto de interferencia destructiva total y el resultado será la cancelación completa del sonido emitido.

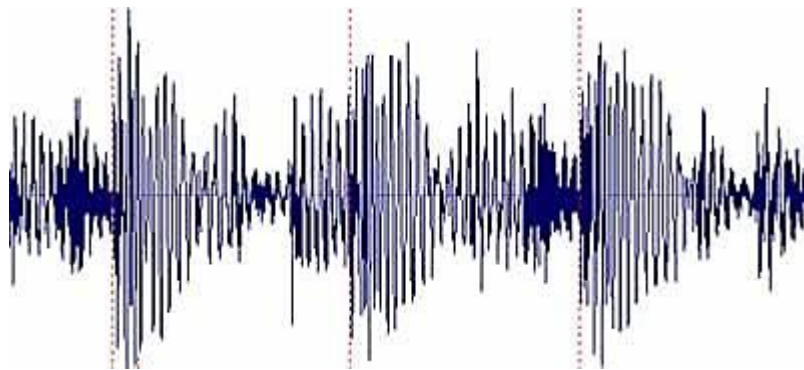


**Ilustración 2.2-6: Suma de ondas desfasadas en 180°**

Si llevamos este resultado a la realidad del sonido, tenemos que al emitirlo los máximos y los mínimos de las ondas representan compresión o expansión del medio por el cual se transmite.

### **2.2.7.2. Cancelación de ondas complejas**

Electrónicamente resolver el problema planteado en la sección 2.2.7.1 sería relativamente simple. Bastaría con tomar la señal invertirla en fase a través de un dispositivo como un amplificador operacional o mediante sistemas digitales y sumarlo a la señal original, pero la mayoría de los sonidos están lejos de parecerse a una onda sinusoidal simple.



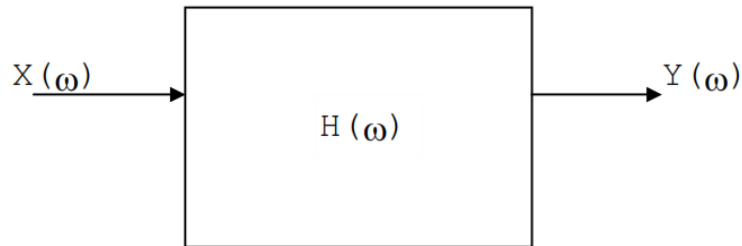
**Ilustración 2.2-7: Forma de onda palabra hablada**

Como se ha mencionado las palabras habladas están compuestas por un espectro de frecuencias de diferentes amplitudes. Lo que significaría que para cancelarlas habría que filtrar y detectar cada una de las frecuencias y luego crear la misma invertida en fase y sumarla, pero claramente es un proceso prácticamente imposible de llevar a cabo. Lo que normalmente se hace es seleccionar una banda estrecha de frecuencias y con el promedio realizar el proceso lo que entrega un resultado bastante bueno aunque no perfecto.

## 2.3. Filtros [8]

Los filtros son circuitos que se caracterizan por permitir el paso de señales eléctricas a un rango de frecuencias determinadas e impide el paso del resto. Estos circuitos quedan descritos a través de una función de transferencia  $H(\omega)$ , cumpliendo que:

$$y(\omega) = H(\omega) \cdot x(\omega) \quad ( 14 )$$



**Ilustración 2.3-1: Estructura de un filtro**

La función de transferencia, para el caso ideal, tomará el valor 1 para todas aquellas frecuencias que se deseen pasar y 0 para las que no.

### 2.3.1. Tipos de filtros

Los filtros pueden clasificarse en base a tres conceptos:

- El tipo de tecnología con que se fabrica el filtro
- La respuesta en frecuencia
- La función de transferencia

Según la tecnología se pueden subdividir en:

- Pasivos
- Activos
- De capacidades conmutadas
- Digitales

Los filtros pasivos usan elementos de tipo pasivo, es decir R, C y L.

Los filtros activos utilizan R, C y amplificadores que pueden ser discretos o integrados.

Los filtros de capacidades conmutadas utilizan condensadores en conmutación en lugar de resistencias. Los valores resistivos que se obtienen dependen de la capacidad y de la frecuencia de conmutación. Por tanto

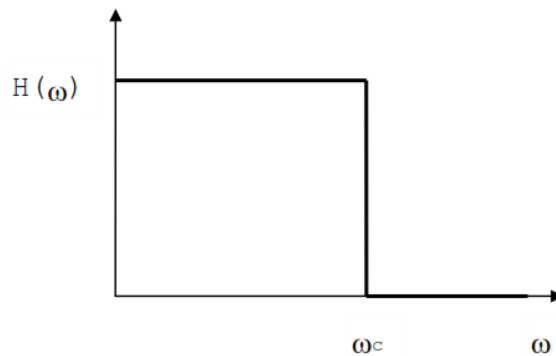
variando la frecuencia podemos cambiar el valor de las resistencias y la respuesta en frecuencia del filtro.

Los filtros digitales realizan la función del filtro a través de algoritmos numéricos, para efectos de este trabajo y por temas de costos, facilidad y factibilidad de diseño y consumo de potencia, se opta por realizar el prototipo con dispositivos analógicos.

Según la respuesta en frecuencia, se pueden subdividir:

- Pasa bajos
- Pasa altos
- Pasa banda
- Rechaza banda

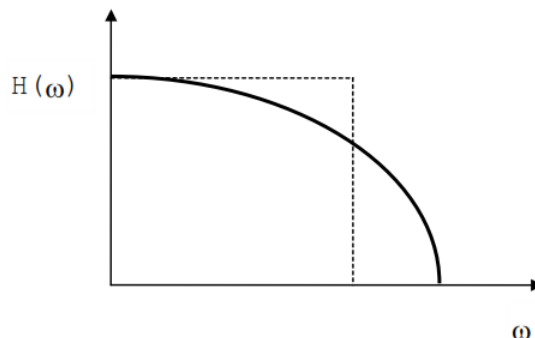
Un filtro pasa bajos tiene una respuesta ideal de la forma (Ilustración 2.3-2):



**Ilustración 2.3-2: Filtro pasa bajos ideal**

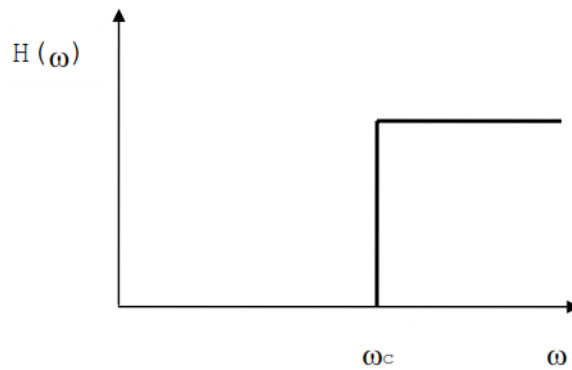
Donde  $|H(\omega)| = 1$  si  $\omega < \omega_c$  y  $|H(\omega)| = 0$  si  $\omega > \omega_c$

Una aproximación más real tendría una respuesta de la forma (Ilustración 2.3-3):



**Ilustración 2.3-3: Filtro pasa bajos real**

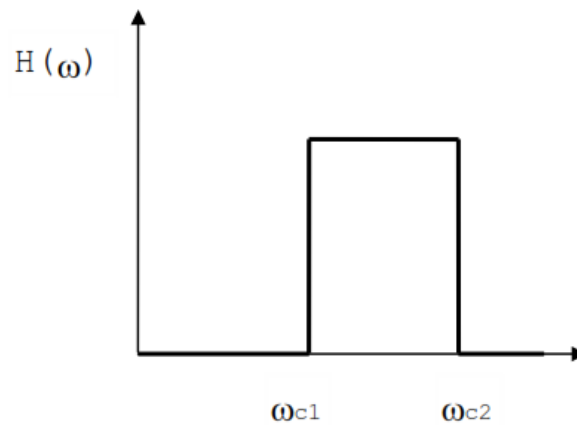
Para el caso de un pasa altos su comportamiento será (Ilustración 2.3-4):



**Ilustración 2.3-4: Filtro pasa altos ideal**

Donde  $|H(\omega)| = 1$  si  $\omega > \omega_c$  y  $|H(\omega)| = 0$  si  $\omega < \omega_c$

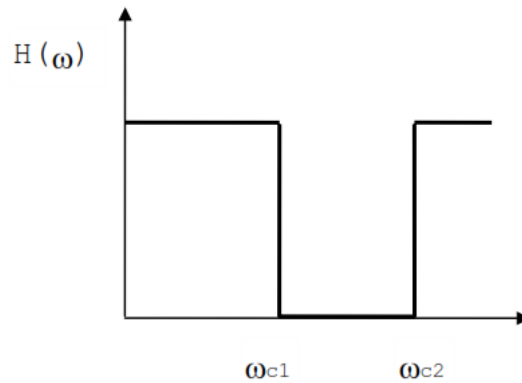
Para un filtro pasabanda se tiene (Ilustración 2.3-5):



**Ilustración 2.3-5: Filtro pasa banda ideal**

$|H(\omega)| = 1$  si  $\omega_{c1} < \omega < \omega_{c2}$  y  $|H(\omega)| = 0$  si  $\omega < \omega_{c1}$  y  $\omega > \omega_{c2}$

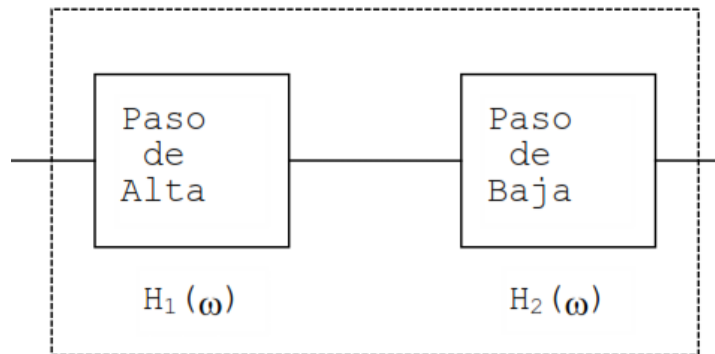
Finalmente en el caso de un filtro rechaza banda, la respuesta en frecuencia es de la siguiente forma (Ilustración 2.3-6):



**Ilustración 2.3-6: Filtro rechaza banda ideal**

Donde  $|H(\omega)| = 0$  si  $\omega_{c1} < \omega < \omega_{c2}$  y  $|H(\omega)| = 1$  si  $\omega < \omega_{c1}$  y  $\omega > \omega_{c2}$

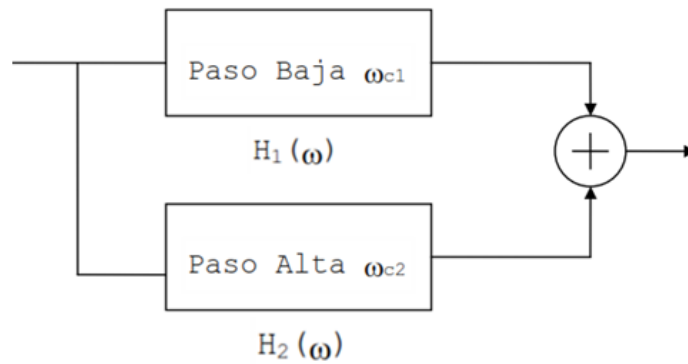
Si analizamos bien la categoría podría verse reducida a tan sólo dos filtros ya que tanto el pasabanda como el rechaza banda pueden ser obtenidos a partir de la mezcla de filtros pasabajos y pasa altos. Para el caso de un pasa banda bastaría con colocar un filtro pasa altos y uno pasabajos en cascada:



**Ilustración 2.3-7: Filtros en cascada para formar pasa banda**

Para el caso de un filtro rechaza banda se deben colocar en paralelo ambos filtros para obtener la respuesta deseada.





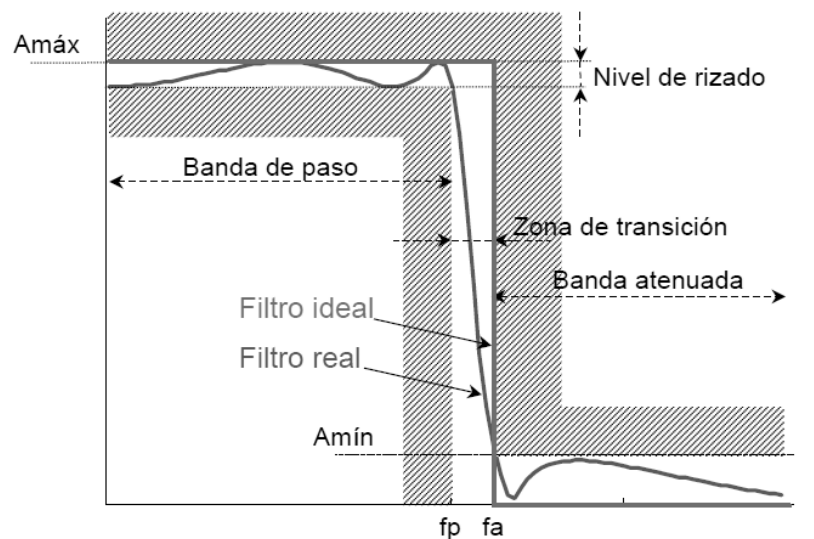
**Ilustración 2.3-8: Filtros en paralelo para formar pasa bandas**

Según la función de transferencia, pueden subdividirse en [9]:

- Butterworth
- Bessel
- Chebyshev
- Cauer

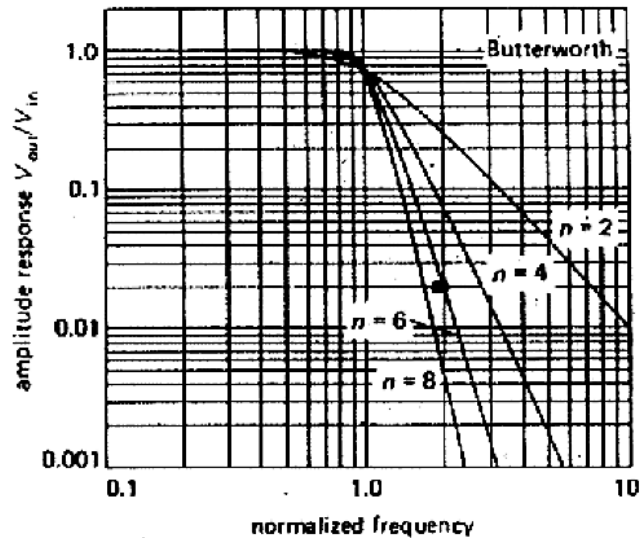
En esta clasificación no es posible desarrollar filtros ideales, pero si se han desarrollado aproximaciones matemáticas, así cada tipo de filtro es una aproximación que optimiza un aspecto particular.

Para una mejor comprensión de la lectura se entrega el siguiente gráfico que muestra diferentes parámetros que definen a un filtro:



**Ilustración 2.3-9: Parámetros de un filtro**

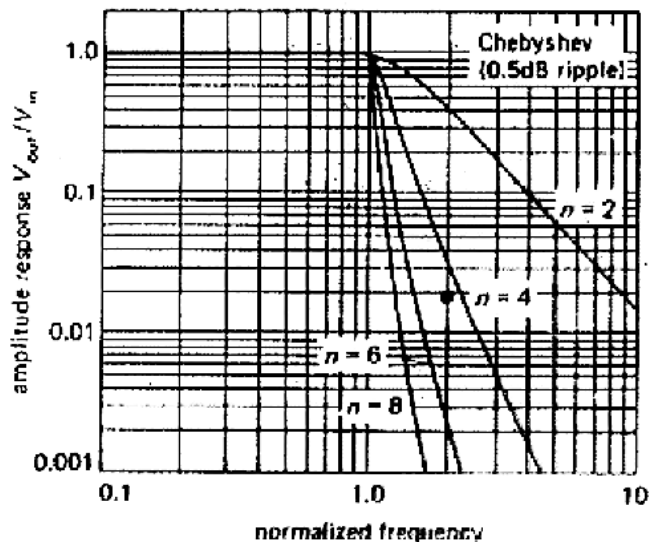
En el caso del filtro de Butterworth se obtiene una ganancia muy plana en la banda de paso con un corte poco abrupto, pero se tiene una respuesta transitoria buena. La respuesta en frecuencia queda graficada de la siguiente forma:



**Ilustración 2.3-10: Respuesta en frecuencia filtro de Butterworth**

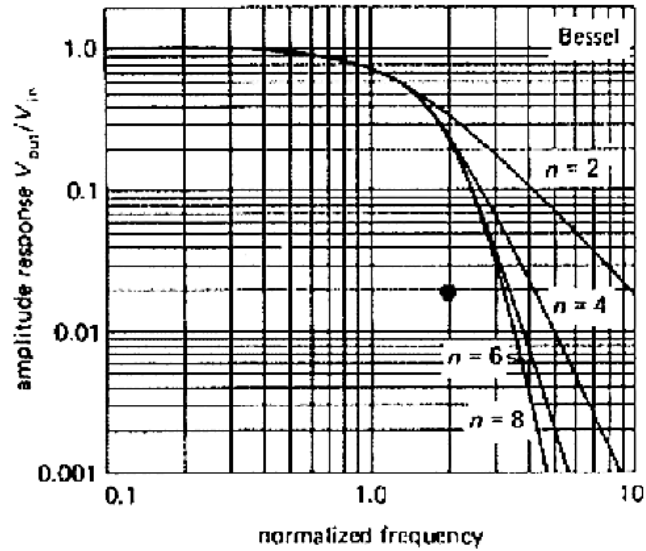
El filtro de Chebyshev produce un corte mas rápido pero agrega rizado en la banda de paso, además con una pendiente de transición más alta y con una respuesta transitoria peor de Butterworth.

Su respuesta se muestra en la Ilustración 2.3-11:



**Ilustración 2.3-11: Filtro de Chebyshev**

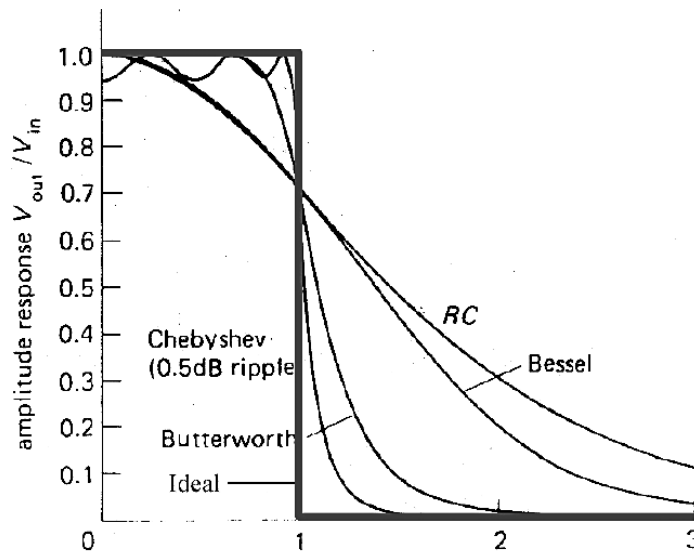
Para el caso de Bessel se tiene una mínima deformación de la señal pero una pendiente de transición peor que el caso de Butterworth.



**Ilustración 2.3-12: Respuesta filtro de Bessel**

Finalmente el filtro de Cauer o elíptico, tiene un corte muy abrupto con rizado en la banda de paso y en la atenuada, una respuesta transitoria peor que Chebyshev. Este tipo de filtro es adecuado para eliminar una frecuencia concreta.

Una comparación entre las respuestas para cada uno de los filtros se puede observar en la siguiente gráfica:



**Ilustración 2.3-13: Comparación de respuestas de filtros**

Normalmente los filtros son utilizados para el acondicionamiento de señales de entrada, digitalización de señales y acondicionamiento de las señales producidas

#### **2.4. Reconocimiento automático del habla (RAH) [10]**

El reconocimiento automático del habla es una parte de la inteligencia artificial en la que muchas disciplinas, como la teoría de señales o la fonética entre otras, juegan un importante papel. El objetivo de la RAH es permitir la comunicación entre humanos y computadoras.

La dificultad de estos sistemas esta en hacer cooperar el conjunto de la información en presencia de ambigüedades, incertidumbres y errores con el fin de obtener una interpretación lo mas fidedigna del mensaje recibido

Este conjunto de informaciones puede provenir de diversas fuentes, acústicas, fonéticas, fonológicas y léxicas, que a través de procedimientos interpretativos dan lugar a un módulo conocido como decodificador acústico fonético.

Una etapa importante a la hora de diseñar un sistema de RAH es la elección del tipo de aprendizaje que se utilice para construir las fuentes de conocimiento. Básicamente, existen dos tipos: el Aprendizaje Deductivo (AD) y el Aprendizaje Inductivo (AI) [11]

Las técnicas de AD se basan en la transferencia de los conocimientos que un experto humano posee a un sistema informático. Un ejemplo paradigmático de las metodologías que utilizan tales técnicas lo constituyen los sistemas basados en el conocimiento y, en particular, los sistemas expertos (Haton, 1985). Por otra parte, las técnicas de AI se basan en que el sistema pueda, automáticamente, conseguir los conocimientos necesarios a partir de ejemplos reales sobre la tarea que se desea modelar. En este segundo tipo, los ejemplos los constituyen aquellas partes de los sistemas basados en los modelos ocultos de Markov o en las redes neuronales artificiales que son configuradas automáticamente a partir de muestras de aprendizaje. [11]

En la actualidad la mayoría de las metodologías para RAH, utilizan AI, aunque generalmente no están basadas solo en ella, se asume un compromiso deductivo-inductivo.

La entrada al decodificador acústico-fonético es la señal vocal convenientemente representada. Para ello, es necesario que ésta sufra un pre proceso de parametrización. Donde se debe asumir algún modelo que la represente, los modelos actualmente utilizados son básicamente estáticos, donde se asume que la señal vocal es cuasi-estacionaria en intervalos cortos

de tiempo. Un estudio del comportamiento dinámico de la voz podría suministrar mayor información a un sistema de RAH. Otro tema poco estudiado es el del habla en ambientes adversos. Caracterizar el Efecto Lombard<sup>6</sup> sería de gran interés en el diseño de sistemas que deben funcionar en un ambiente con altos niveles de ruido.

Existe una gran cantidad de problemas en el diseño de un decodificador acústico-fonético, entre ellos la elección de las unidades sub léxicas<sup>7</sup> adecuadas al momento de crear un modelo

Para el diseño de modelos se utilizan en mayor medida unidades del tipo fonéticas. La idea final es relacionar las unidades sub léxicas con unidades lingüísticas para poder caracterizar el vocabulario con la aplicación concreta, para que no sean necesarias muestras de entrenamiento con todas las palabras que lo componen. La fonología por su parte puede agregar información sobre las reglas que se aplican a las posibles secuencias de unidades sub léxicas, lo que ayuda a la hora de depurar errores del sistema. Otro problema a enfrentar es el que tiene relación con modelar palabras cortas que pueden cambiar el significado de una pronunciación

Los sistemas actuales de RAH necesitan para su correcto funcionamiento grandes corpus lingüísticos de voz<sup>8</sup> para el aprendizaje y evaluación de los correspondientes sistemas. El aporte de expertos en fonética y sintaxis. es importantísimo para el diseño del corpus, especialmente, en las transcripciones, segmentación y estudio de la representatividad del material que se debe adquirir.

## **2.5. Tipos de reconocimiento de voz**

Los sistemas de reconocimiento de voz pueden ser clasificados de diferentes formas en función del tipo de enunciado que son capaces de reconocer. No es una clasificación cerrada, es decir, un enunciado puede ser parte de más de una clasificación a la vez.

---

<sup>6</sup> Efecto que produce aumentar la intensidad de voz inconscientemente al encontrarse en un ambiente ruidoso

<sup>7</sup> Las unidades sub léxicas son segmentos acústicos sin significado propio pero tales que, por concatenación de las unidades correspondientes, permiten la construcción del modelo acústico de cualquier palabra o frase.

<sup>8</sup> Un corpus lingüístico es un conjunto, habitualmente muy amplio, de ejemplos reales de uso de una lengua. Estos ejemplos pueden ser textos (lo más común) o muestras orales (generalmente transcritas).

### **2.5.1. Palabras aisladas**

Los sistemas de reconocimiento de palabras aisladas por lo general requieren un silencio entre cada frase, esto no significa que sólo acepta palabras individuales, pero si requieren una expresión a la vez. Normalmente se realiza el procesamiento durante estas pausas.

### **2.5.2. Palabras relacionadas**

Conecta los sistemas de palabras (o más correctamente conecta los enunciados), este sistema es similar al de palabras aisladas, pero permite pequeñas pausas entre los fonemas.

### **2.5.3. Discurso continuo**

Un sistema de reconocimiento de voz con capacidades continuas es de los más difíciles de crear porque se debe utilizar métodos especiales para determinar los límites de enunciación. Estos sistemas permiten a los usuarios hablar casi de manera natural, mientras que el equipo determina el contenido.

### **2.5.4. Habla espontánea**

Un sistema de RAH con la capacidad de habla espontánea debe ser capaz de manejar una variedad de características naturales, tales como palabras que se ejecuta en conjunto, "ums" y "ahs", e incluso tartamudeos leves.

## **2.6. Verificación/identificación de voz**

Algunos sistemas RAH tienen la capacidad de identificar a los usuarios específicos. Aunque para el caso de este trabajo no es necesario entrar en detalles.

### **2.6.1. Representación de señales de voz en reconocimiento [10]**

Para comenzar a diseñar el sistema de reconocimiento de la señal de voz, es esencial la extracción de valores y parámetros que tengan información relevante de la señal, idealmente el número de parámetros debe ser el mínimo posible.

La información relevante es toda aquella que permite al sistema reconocer el mensaje. Minimizar el número de parámetros va en el sentido de mejorar el tiempo de respuesta, desde que la señal entra hasta que el sistema ejecuta la acción deseada. Para mejorar el reconocimiento se utilizan técnicas de compresión de voz.

Entre los pasos necesarios para conseguir lo anterior está el limitar la voz en banda y digitalizarla. Generalmente la limitación se realiza a una frecuencia entre 3,7 kHz y 5 kHz, y se muestrea a una frecuencia entre 8 kHz y 10 kHz. El siguiente paso es segmentar la señal, esto es dividirla entre trozos de duración fija, normalmente 30ms pudiendo estar solapados entre sí en alrededor de 10 a 20 ms. Luego a estos segmentos se les realiza una compresión de datos, que es un análisis de frecuencia de la señal, que entregan como resultado un conjunto de parámetros representativos.

Este análisis de frecuencia puede realizarse según distintos métodos siendo los más utilizados los siguientes:

- Filtrado de la señal en bandas de frecuencia y cálculo de la energía en cada banda: El resultado es que cada segmento queda representado por N valores de energía, cada uno correspondiente a una banda.
- Cálculo de la envolvente del espectro: Está basado en un sencillo modelo articulatorio donde el tracto vocal se supone formado por una sucesión de tubos acústicos sin pérdidas. Aplicando técnicas de predicción lineal se extrae, para cada segmento, un conjunto de parámetros (el número oscila entre 8 y 12) que representan el modelo articulatorio y a su vez llevan información de la envolvente del espectro como ya se ha mencionado. Aparecen claramente representados los cuatro formantes más significativos. En este último método también se añade la energía asociada a cada segmento.

El oído no responde de una forma lineal a una excitación dada. El doble de potencia no significa que el oído lo perciba con el doble de intensidad, tiene una respuesta más bien logarítmica (decibelios) y por esta razón en los métodos de obtención de parámetros se tiene en cuenta una cierta compresión logarítmica en las características de frecuencia extraídas.

Posteriormente a la parametrización por cualquiera de los métodos mencionados se realiza una estimación de las características dinámicas del espectro, evolución temporal del espectro y evolución temporal de la energía. Finalmente con toda esta información se procede a realizar el reconocimiento.

El conocimiento de modelos auditivos y articulatorios puede mejorar las técnicas de representación de la señal.

Un simple análisis del modelo auditivo se ha traducido en realizar una compresión logarítmica del espectro. Si se trabaja con una representación de frecuencias más plana logrando enfatizar aquellos formantes que, de otra manera, quedarían totalmente enmascarados por el de mayor valor. Un análisis algo más profundo del comportamiento auditivo, podría aumentar la

calidad de la señal parametrizada teniendo en cuenta las siguientes consideraciones:

Existe un umbral fijo de audición por debajo del cual el oído es incapaz de percibir sonidos. Este umbral es dependiente de la frecuencia.

Existe un umbral variable de audición producido por un efecto de enmascaramiento psicoacústico. Un tono a una frecuencia determinada no permite percibir tonos por debajo de ese umbral dentro de lo que se denomina «banda crítica».

La parametrización basada en el modelo articulatorio es susceptible de ser mejorado. Por ejemplo el modelo no tiene en cuenta la posible resonancia producida por la cavidad nasal.

### **2.6.2. Condiciones adversas**

El sistema de reconocimiento diseñado para trabajar en un laboratorio puede verse altamente degradado cuando trabaja en un ambiente real. Intervienen muchos factores extras que pueden no haber sido considerados: Ruido, múltiples conversaciones simultáneas, factores psicológicos que alteran al locutor como stress, fatiga, factores ambientales entre otros.

Para el sistema en desarrollo conviene que el sistema sea robusto; es decir, que los parámetros obtenidos de los modelos anteriores no se vean modificados fuertemente por estos factores. Para ello existen varias técnicas.

El estudio del ruido es un tema bastante amplio, dado que existen muchas fuentes de ruido con diferentes características pero, por ejemplo, se puede utilizar un cancelador de ruido, que es un sistema que aprende como es el ruido, sus características en frecuencia, con el fin de cancelarlas de la señal original. Luego de esta cancelación se procede a parametrizar la señal. El estudio del ruido propiamente tal es un tema bastante amplio, dado que existen muchas fuentes de ruidos con diferentes características.

Para el caso de las conversaciones múltiples el oído es capaz de discriminar y seguir una de las conversaciones, en los sistemas desarrollados actualmente este es un tema no resuelto, sin embargo existen avances en modelos de discriminación tonal, esto es a grandes rasgos, tomar muestras de diversos lugares y comparar los tonos que se registran en cada una de ellas.



### **2.6.3. Incorporación de información acústica fonética**

Ya revisados los sistemas de obtención de parámetros, puede ser de interés introducir algunos temas, que desde un punto de vista acústico-fonético pueden mejorar la caracterización de la señal

Una representación tiempo frecuencia como el espectrograma puede dar a un experto en fonética información suficiente para realizar un reconocimiento con una tasa de error muy baja. Parte de esta información es condensada en los parámetros expuestos anteriormente, aunque existe aún mucha información no tomada en cuenta, como pueden ser parámetros típicos del lenguaje hablado, como por ejemplo:

- Entonación: Puede facilitar la búsqueda de puntos donde puede asegurarse que empieza o termina una palabra o una frase.
- Rasgos distintivos: Determinan unívocamente cada alófono<sup>9</sup>. Si bien es difícil determinar con exactitud todos los rasgos distintivos en un segmento de señal para asociarlo a un alófono dado, un análisis de los mismos puede permitir una clasificación de sonidos en agrupaciones estables y con significado fonético. Una posible aplicación es la segmentación de señales.
- Reglas fonológicas: El conjunto de sonidos de una lengua es importante, pero no menos importante es el conocimiento de las combinaciones posibles entre estos sonidos. Las reglas fonológicas permiten avanzar en el estudio de la coarticulación y establecen restricciones importantes en las transiciones entre fonemas. Las reglas fonológicas pueden ser utilizadas para realizar una etapa de reconocimiento entre agrupaciones de fonemas previa al reconocimiento definitivo.

### **2.7. Usos y aplicaciones de sistemas de RAH**

Las aplicaciones más comunes en las que se desarrollan sistemas de reconocimiento automático del habla son los siguientes:

#### **2.7.1. Dictado**

Constituye el uso más común para los sistemas de RAH en la actualidad. Esto incluye transcripciones médicas, el dictado legal y de negocios, así como procesamiento de textos en general. En algunos casos se utilizan vocabularios especiales para aumentar la precisión del sistema.

---

<sup>9</sup> Cada uno de los sonidos que en un idioma dado se reconoce como un determinado fonema, sin que las variaciones entre ellos tengan valor diferenciativo.

### **2.7.2. Mando y Control**

Existen sistemas de RAH que están diseñados para llevar a cabo funciones y acciones en sistemas de Mando y Control. Las expresiones como "atento" y "cambio" son ejemplos de ello. En este caso la acción realizada es la de abrir o cerrar el canal de comunicación de la radio, reemplazando al sistema push to talk

### **2.7.3. Telefonía**

Algunos sistemas de correo de voz permiten que las personas que llaman digan comandos en lugar de pulsar botones para enviar los mensajes específicos

### **2.7.4. Médico / Discapacidad**

Muchas personas tienen dificultad para escribir debido a las limitaciones físicas, tales como lesiones por esfuerzo repetitivo, distrofia muscular, entre otros. Por ejemplo, las personas con problemas de audición podrían utilizar un sistema conectado a su teléfono para convertir el discurso de la persona que llama en texto.

### **2.7.5. Teléfonos móviles**

Algunos teléfonos móviles más modernos incluyen reconocimiento de voz que permiten expresiones tales como "Llamar casa".

## **2.8. Radios portátiles**

Más conocidos por el inglés walkie-talkie, es un dispositivo transceptor<sup>10</sup> de radio portátil. En un comienzo desarrollados para uso militar, se han masificado y actualmente es posible encontrarlos en un sin número de aplicaciones, ya sean para la industria en general, la policía, bomberos y hasta el uso civil en áreas aisladas. Entre sus características principales encontramos que posea un canal semi dúplex, es decir solo uno de los radios puede transmitir a la vez, lo que no impide que pueda simultáneamente ser escuchada por muchas otras unidades, y un interruptor que permite iniciar la comunicación o abrir el canal, sistema conocido como push to talk.

Existen variadas versiones de estos dispositivos, siendo algunos modernos muy similares a las primeras versiones de los teléfonos celulares. Según la forma de transmisión de datos, hay modelos tanto análogos como digitales.

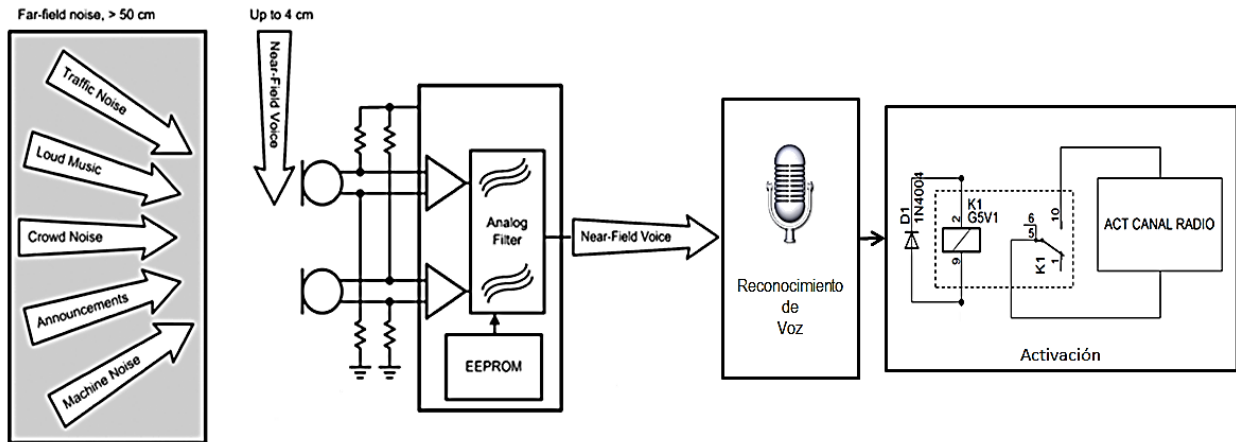
---

<sup>10</sup> Dispositivo capaz de realizar funciones tanto de transmisión como de recepción

### 3. Implementación

La implementación del sistema se puede dividir en cuatro etapas, cuyo origen está en los principales temas a resolver para la conclusión con éxito del sistema. Estas son:

- Sistema cancelador de ruido para micrófonos
- Filtros de señal
- Reconocimiento de voz
- Activación canal de comunicación



**Ilustración 2.8-1: Diagrama de flujo desarrollo propuesto**

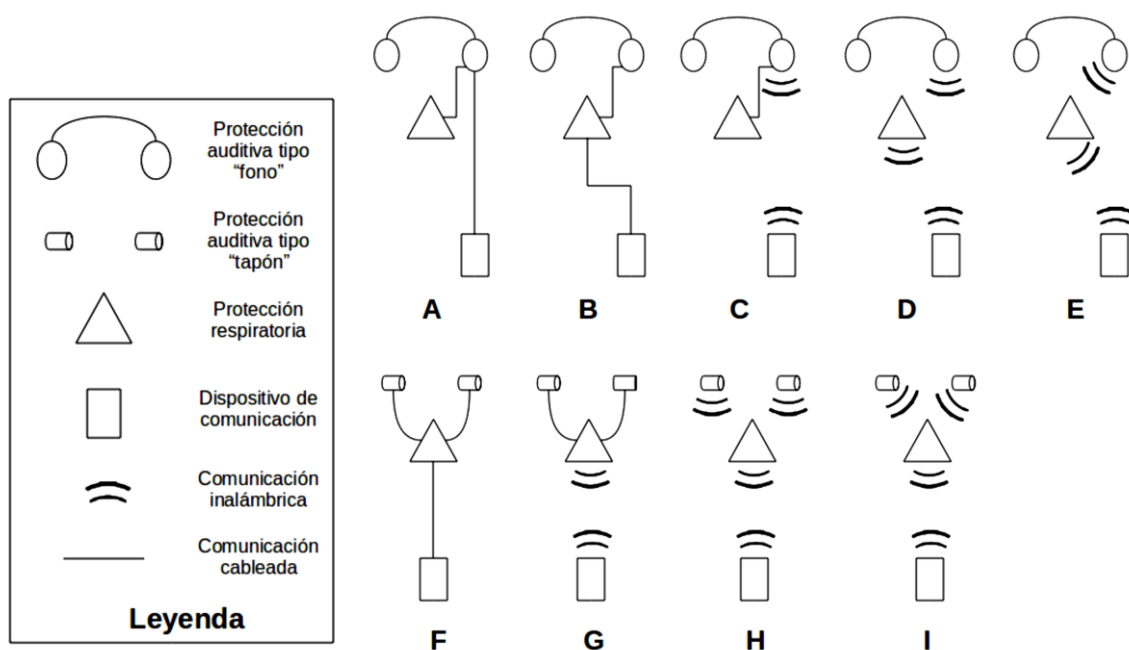
Los tiempos de trabajo de cada módulo dependen de la importancia que tengan en el resultado final, siendo el más importante el módulo de cancelación de ruido, dado que los otros tendrán mejores resultados en cuanto éste funcione mejor.

Otro punto a destacar es la forma física de la mascarilla y el formato de los sistemas de comunicación, dado que también definen muchas de las decisiones que se toman en torno a los diseños de los módulos.

Se define al sistema como una fusión entre los elementos de protección personal y los elementos electrónicos que se utilizan para la comunicación, siendo de vital importancia intervenir lo menos posible las características de seguridad que poseen.

En todo momento el trabajo se realiza en conjunto con la empresa SoluNova Ltda. la cual propone una serie de formatos de dispositivos finales a implementar, incluyendo la mascarilla y un sistema de comunicación usando los protectores auditivos. La representación de cada uno y su descripción se observan a continuación.

- A. Comunicación cableada entre dispositivo de comunicación y "fono", y cableada entre fono y máscara.
- B. Comunicación cableada entre dispositivo de comunicación y máscara, y cableada entre máscara y "fono".
- C. Comunicación inalámbrica entre dispositivo de comunicación y "fono", y cableada entre fono y máscara. Alternativamente, comunicación inalámbrica entre dispositivo de comunicación y máscara, y cableada entre fono y máscara.
- D. Comunicación inalámbrica entre dispositivo de comunicación y máscara, e inalámbrica entre dispositivo de comunicación y "fono".
- E. Comunicación inalámbrica entre dispositivo de comunicación y "fono", e inalámbrica entre máscara y "fono". Alternativamente, comunicación inalámbrica entre dispositivo de comunicación y máscara, e inalámbrica entre máscara y "fono".
- F. Comunicación cableada entre dispositivo de comunicación y tapones, y cableada entre tapones y máscara.
- G. Comunicación inalámbrica entre dispositivo de comunicación y máscara, y cableada entre máscara y tapones.
- H. Comunicación inalámbrica entre dispositivo de comunicación y máscara, e inalámbrica entre dispositivo de comunicación y tapones.
- I. Comunicación inalámbrica entre dispositivo de comunicación y máscara, e inalámbrica entre máscara y tapones. Alternativamente, comunicación inalámbrica entre dispositivo de comunicación y tapones, e inalámbrica entre máscara y tapones.



**Ilustración 2.8-2: Formatos de comunicación**

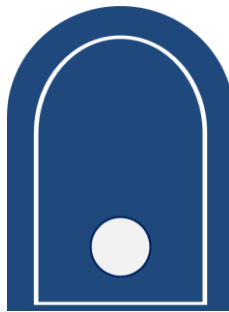
De todas ellas la opción A es la más fácil de realizar a corto plazo.

Las máscaras de protección respiratorias son provistas por Masprot, fabricantes chilenos de elementos de protección personal. El modelo con el que se trabaja es "Supreme".



**Ilustración 2.8-3: Máscara de protección respiratoria Supreme**

Posee una pieza frontal (Ilustración 2.8-4) pequeña en cual posteriormente se insertaran los componentes electrónicos, agregando al diseño condiciones límites para el tamaño de los circuitos diseñados, pues no se plantea agregar equipos extras, para no sobrecargar al operario, por lo que estos sistemas electrónicos no pueden ir en una caja separada.



**Ilustración 2.8-4: Pieza frontal**

Un diagrama modelado del sistema se muestra en la Ilustración 2.8-5



**Ilustración 2.8-5: Modelo del sistema**

Con todas las características definidas el dispositivo final tiene la siguiente estructura:



**Ilustración 2.8-6: Sistema intercomunicador**

Donde la A indica que ambos sistemas, el de recepción y el de emisión, son alámbricos.

### **3.1. Sistema cancelador de ruido**

Uno de los principales problemas a analizar es que los trabajadores se ven envueltos en ambientes de trabajo con niveles de ruido extremadamente altos en los cuales muchas veces se ve imposibilitada la comunicación, por lo que desarrollar un buen sistema cancelador de ruido es de vital importancia.

Muestras de niveles de presión sonora obtenidos en terreno, arrojaron resultados bastante malos para el desarrollo, pues se obtuvieron niveles superiores a los 90 dBA en la mayoría de los casos.

Existen una serie de desarrollos para cancelar ruido en audífonos, los que inspiran gran parte del desarrollo de este cancelador de ruido para micrófonos. El principio es similar al explicado en la sección 2.2.7 y para lograr el resultado se utilizan dos micrófonos: uno para captar el ruido externo y otro para captar la voz. Luego mediante un circuito electrónico se anulan las señales de ruido entregando como resultado la voz.

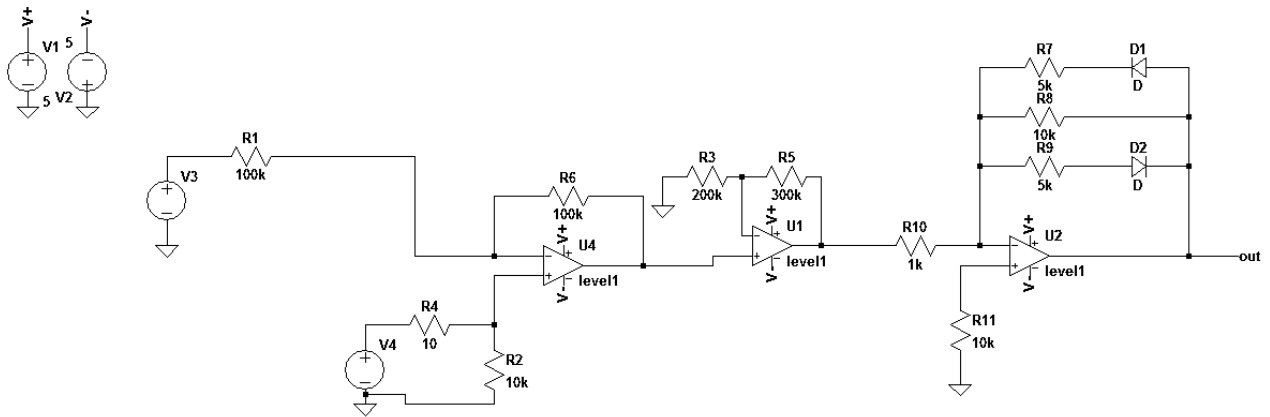
#### **3.1.1. Diseño cancelador de ruido**

En un comienzo el diseño tiene como base la cancelación activa de ruido en audífonos, se parte diseñando un circuito bastante básico para realizar simulaciones en el software LTSPICE<sup>11</sup>, el cual está compuesto de un OPAMP en configuración restador, tomando la señal de un micrófono que tan solo capta el ruido y restándosela al micrófono encargado de captar la voz y ruido. Este diseño además consta de una etapa de amplificación de la señal y otra de compresión<sup>12</sup>:

---

<sup>11</sup> <http://www.linear.com/designtools/software/#LTspice>

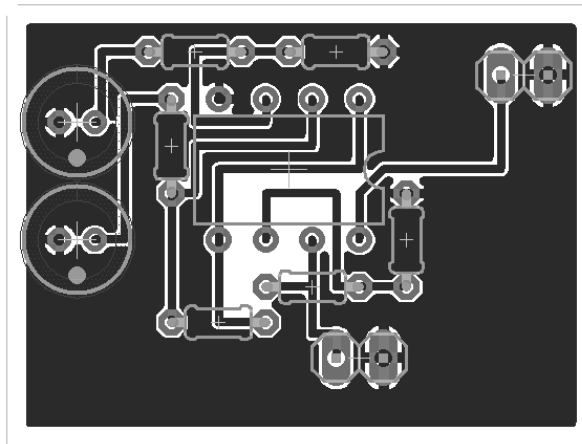
<sup>12</sup> El compresor acota el rango dinámico de la señal recortando la base de las señales con mayor nivel y enfatizando los picos de volumen de las señales mas tenues. Fuente: <http://www.rock.com.ar/especiales/sonido/entrega11.shtml>



**Ilustración 3.1-1: Circuito diseñado versión 1**

En el diagrama se observan los micrófonos como fuentes de voltajes, ya que el software permite el funcionamiento de las fuentes de voltaje como fuentes de audio.

Este circuito se diseña en PCB, para lo que se utiliza el software Eagle PCB<sup>13</sup>, obteniéndose lo siguiente:



**Ilustración 3.1-2: Diseño PCB cancelador de ruido v1**

Esta versión preliminar tiene sólo fines de aprendizaje y sirve para comenzar a trabajar con software y sistemas que posteriormente permitirán implementar otras las soluciones mas complejas.

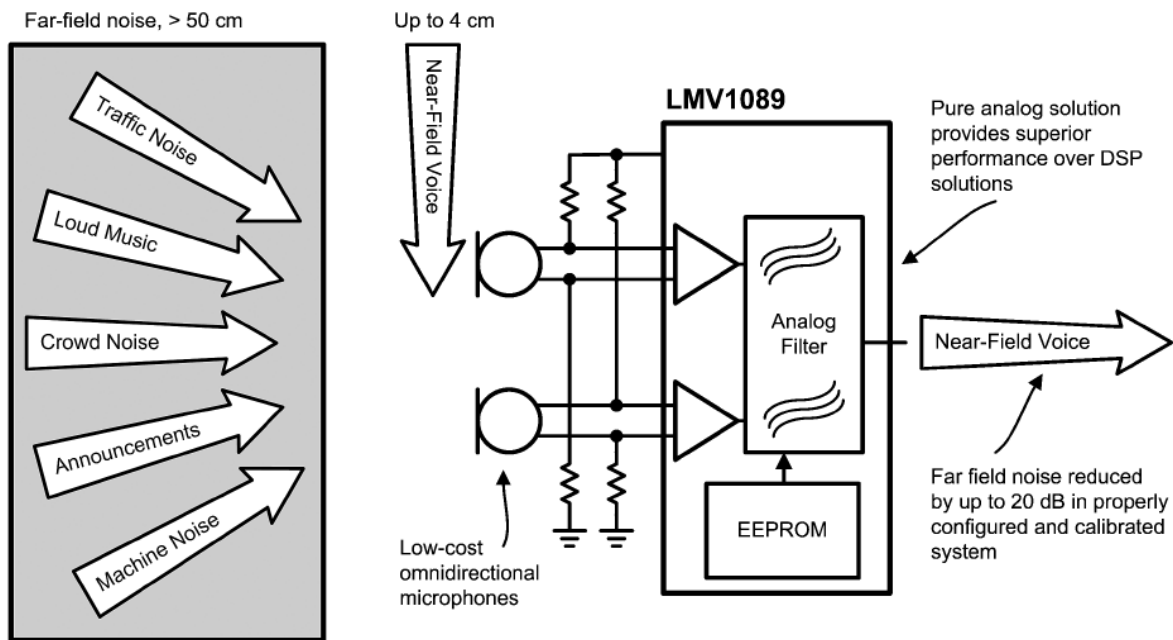
El sistema definitivo contará con el integrado LMV1089<sup>14</sup> (Dual Input, Far Field Noise Suppression Microphone Amplifier with Automatic Calibration

<sup>13</sup> <http://www.cadsoftusa.com/eagle-pcb-design-software/product-overview/?language=en>



Capability), que será el motor principal del sistema final del cancelador de ruido. Es fabricado por Texas Instruments y promete realizar la reducción de hasta 20 dB de ruido externo con un bajo consumo de energía, además si se considera su tamaño que en el empaquetado mas grande mide cerca de 1x1 cm, lo convierten en el candidato ideal para el prototipo definitivo.

El funcionamiento del integrado lo podemos observar en el siguiente diagrama de bloques:



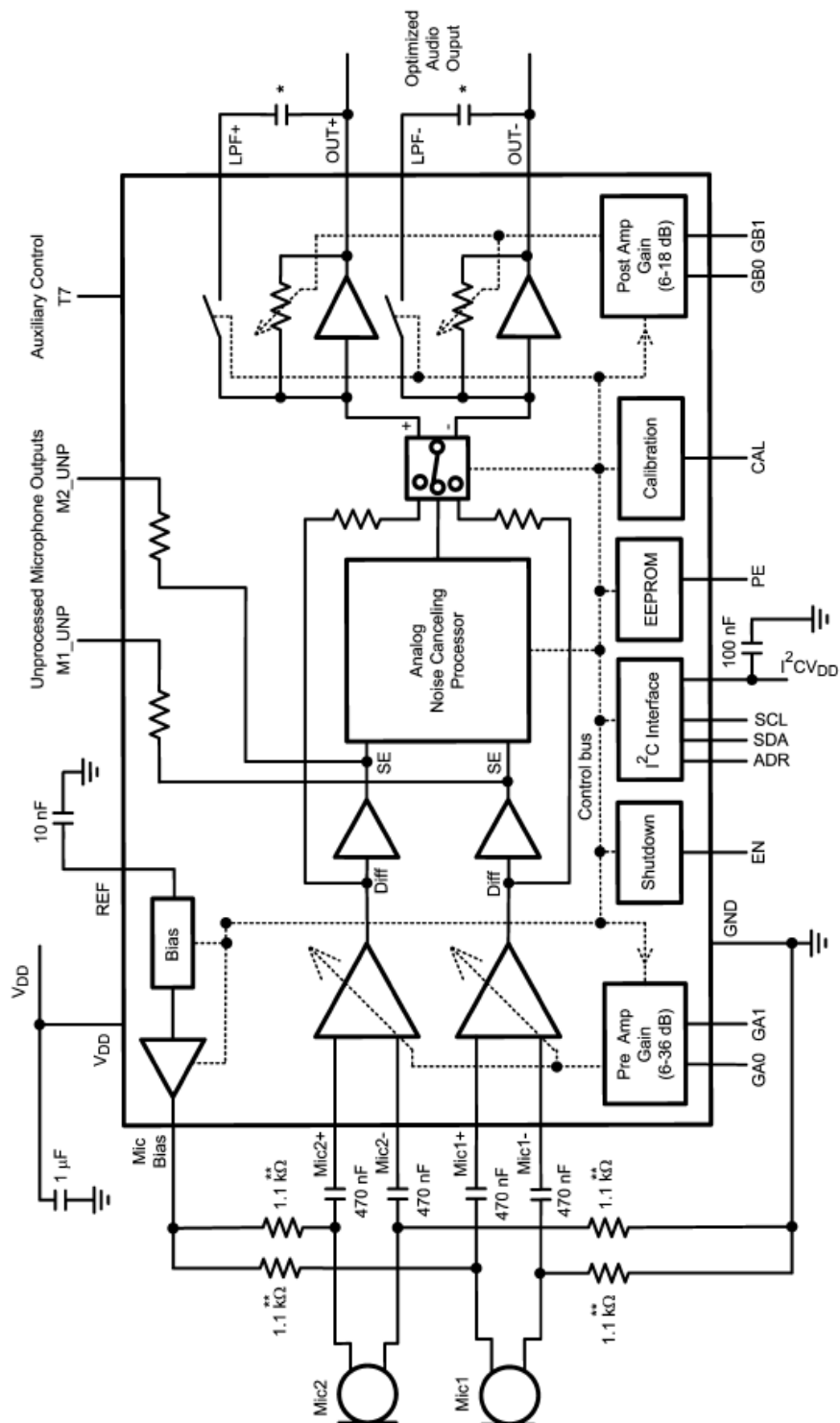
**Ilustración 3.1-3: Esquema funcionamiento LMV 1089**

Se observa que el integrado trabaja con dos micrófonos<sup>15</sup> que permiten captar el ruido cercano y lejano a la vez, el LMV1089 se encarga de diferenciar y procesar estos sonidos, amplificando el sonido cercano (hasta 4 cm) y reduciendo el lejano (desde los 50 cm), entregando como resultado una salida de voz con ruido ambiental reducido.

Si se entra en detalle, las funciones internas quedan descritas en la Ilustración 3.1-4:

<sup>14</sup> <http://www.ti.com/lit/ds/symlink/lmv1089.pdf>

<sup>15</sup> Se recomienda por el fabricante una separación de 1,5 a 2,5 cm entre ambos micrófonos



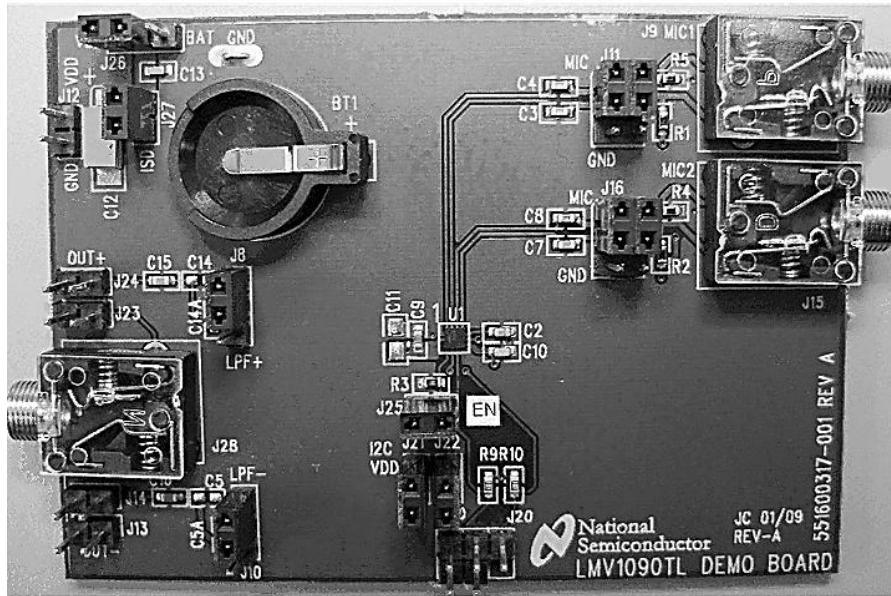
**Ilustración 3.1-4: Diagrama de bloques LMV1089**

Se distingue que las etapas del integrado son: pre amplificación de las señales de los micrófonos, procesamiento análogo de la cancelación y post amplificación.

Para comprobar el funcionamiento de este chip, se utiliza una placa de desarrollo del integrado LMV1090 (Ilustración 3.1-5: Placa de desarrollo LMV1090), dado que al momento de la compra no existe stock de la placa de desarrollo del LMV1089. Sin embargo las diferencias entre ambos integrados son mínimas pues tienen la misma función pero opciones distintas, en este caso el LMV1089 tiene una opción de calibración del dispositivo que el LMV1090 no posee. Estos integrados pertenecen a una familia desarrollada por National Semiconductors dedicada a trabajar en la reducción de ruido. El detalle se observa en la Tabla 3.1-1: Familia de integrados LMV

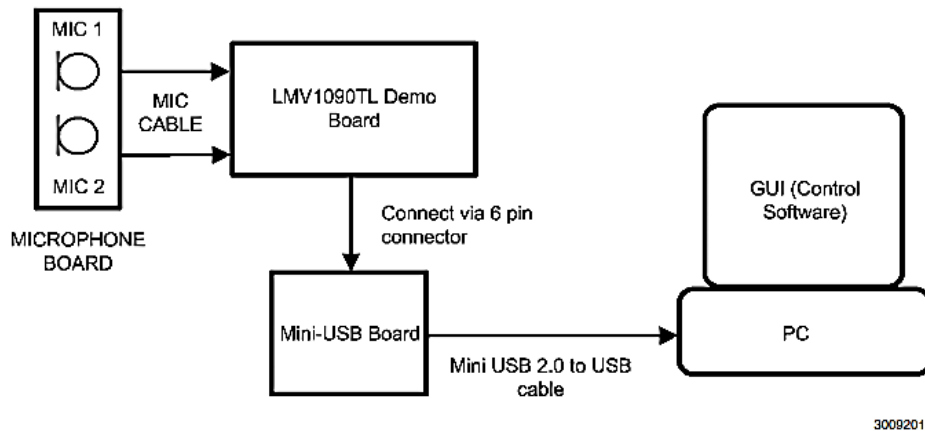
**Tabla 3.1-1: Familia de integrados LMV**

<b>Numero de parte</b>	<b>Descripción</b>	<b>Tipo</b>
<b>LMV1051</b>	Processor for Dual Microphone Adaptive Noise Cancelling with Wind Noise Alert	Device
<b>LMV1088</b>	Dual Input, Far Field Noise Suppression Microphone Amp w/ Automatic Calibration Ability	Device
<b>LMV1089</b>	Dual Input, Far Field Noise Suppression Mic Amp w/ Auto Calibration Capability	Device
<b>LMV1090</b>	Dual Input, Far Field Noise Suppression Microphone Amplifier	Device
<b>LMV1090APP-SW</b>	Software (Windows) Controls the LMV1090 and Its Operating Modes	Tool/Software
<b>LMV1090GUI-SW</b>	Software (SILAB USB/I2C board, Windows) Controls the LMV1090 and Its Operating Modes	Tool/Software
<b>LMV1090TLEVAL</b>	Dual Input, Far Field Noise Suppression Microphone Amplifier	Tool/Software
<b>LMV1091</b>	Dual Input, Far Field Noise Suppression Microphone Amplifier	Device
<b>LMV1091TMEVAL</b>	Dual Input, Far Field Noise Suppression Microphone Amplifier	Tool/Software
<b>LMV1099</b>	Uplink Far Field Noise Suppression; Downlink SNR Enhancing Mic Amp w/ Earpiece Driver	Device
<b>LMV1099APP-SW</b>	LMV1099 Software > LMV1099APP-SW	Tool/Software
<b>LMV1099TLEVAL</b>	LMV1099 Evaluation Board	Tool/Software



**Ilustración 3.1-5: Placa de desarrollo LMV1090**

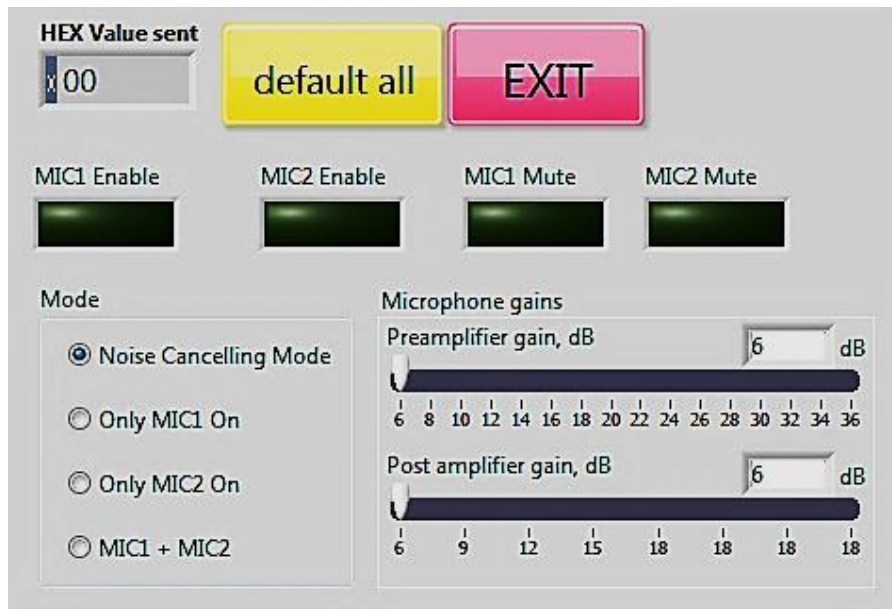
Las pruebas en la placa de desarrollo se realizan conectándola al computador a través del puerto USB y con comunicación por i2c, a través de esta se permite la configuración de diversos parámetros, que maneja las características de las etapas internas del integrado. El diagrama de bloques de la conexión lo vemos en la Ilustración 3.1-6.



**Ilustración 3.1-6: Conexión placa de desarrollo PC**

El software que controla la comunicación entre el PC y la placa de desarrollo es proporcionado por Texas Instruments y puede ser descargado libremente desde su página<sup>16</sup>. Tiene una interfaz gráfica amable y simple de controlar.

<sup>16</sup> <http://www.ti.com/tool/lmv1090app-sw>



**Ilustración 3.1-7: LMV1090 Software**

Las pruebas serán analizadas directamente en el computador, conectando la salida de la placa de desarrollo a la entrada de micrófono. El software que se utiliza en este proceso es Sony Sound Forge Pro 10, especializado en la edición de audio profesional y cuenta con una gran cantidad de herramientas necesarias para el análisis de las pruebas. Se utilizan para este trabajo, las herramientas de grabación, edición, ecualización, análisis espectral y análisis en tiempo real.



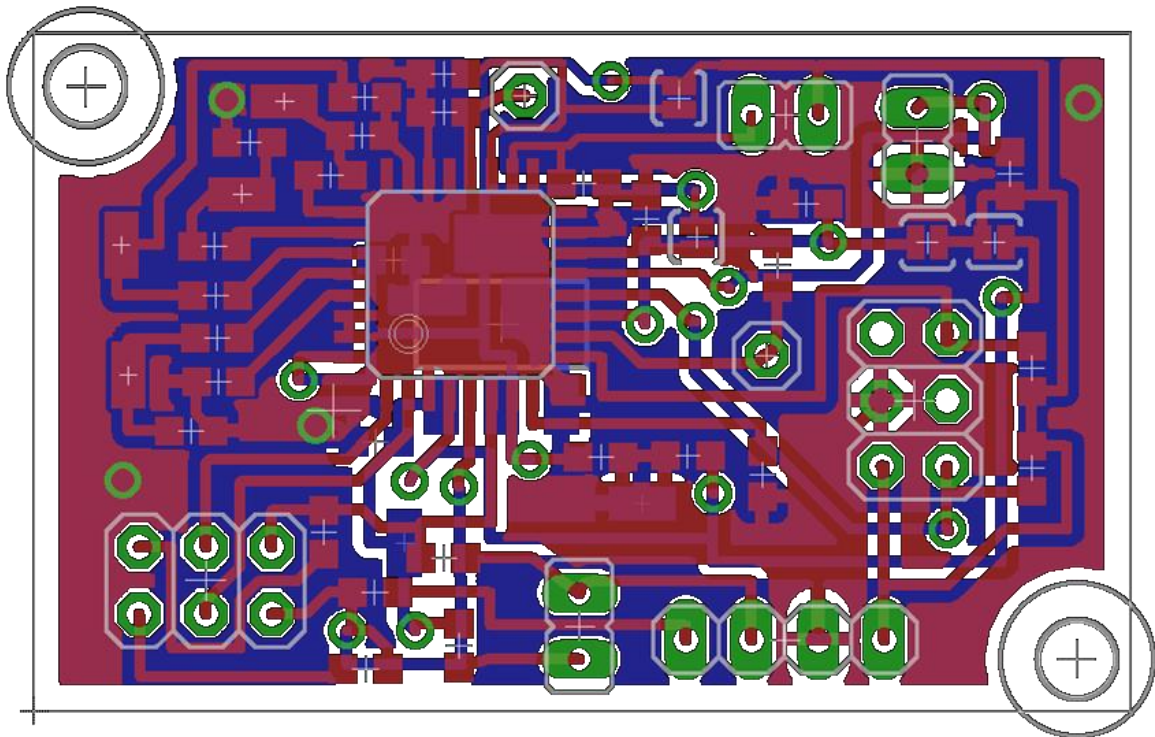
**Ilustración 3.1-8: Sony Sound Forge Pro 10**

Dado los buenos resultados, se decide comenzar el diseño de una placa que contenga al LMV1089, y que será la v1 de una futura versión comercial. Esta poseerá la mayor cantidad posible de características del integrado, siendo muy similar a la placa de desarrollo. A medida que se avance con las pruebas

y en relación a los resultados se reducirá tanto el tamaño como la cantidad de funciones que posea.

Al igual que el caso anterior se diseña en Eagle, siguiendo el siguiente diagrama esquemático propuesto en datasheet del integrado (ver anexo 7.2).

El diseño final en es el siguiente:



**Ilustración 3.1-9: PCB LMV1089 v1**

Pasando a la máscara propiamente tal, se considera realizar pruebas prácticas de posición del micrófono. Para ello se trabaja con diversos modelos de tapas, donde se ubican los micrófonos en distintas posiciones y se analiza su respuesta ante los mismos niveles de ruido.

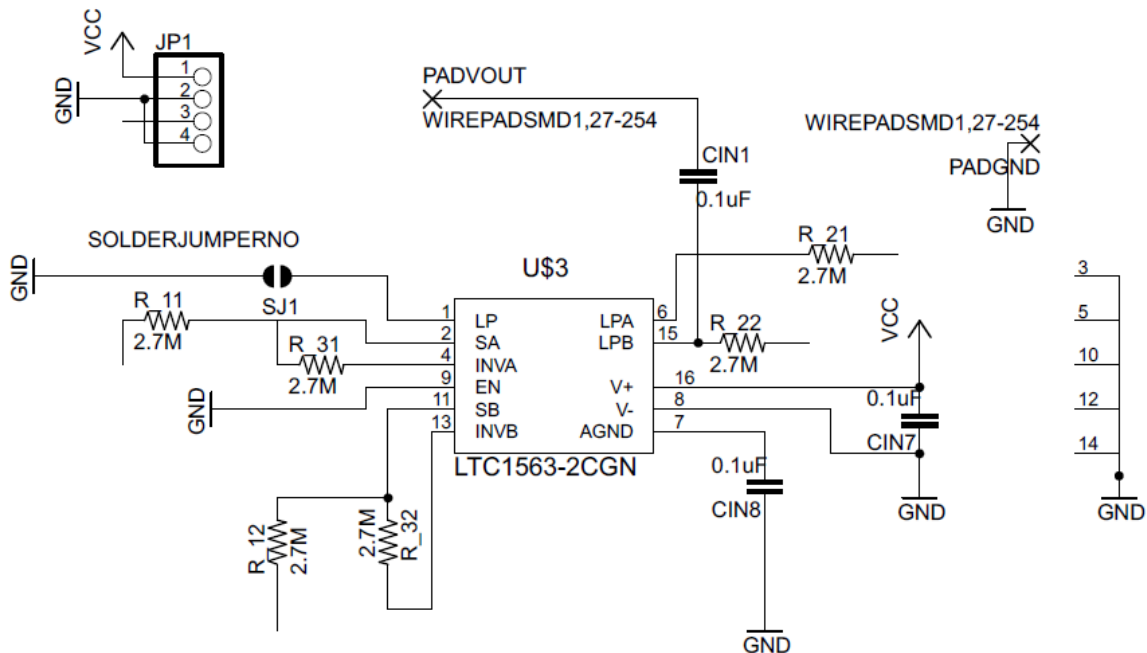


**Ilustración 3.1-10: Modelos de tapas a utilizar**

### 3.2. Diseño filtros

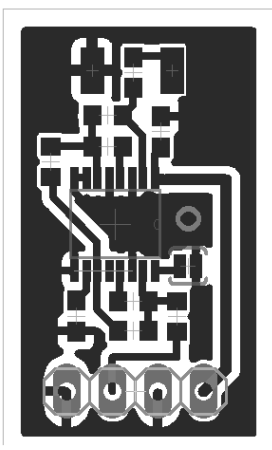
El diseño final considera filtros para lograr mejores resultados, en un principio y a partir de las muestras de ruido disponibles se opta por comenzar el diseño de un filtro pasabajos activo, para lo cual se utiliza un integrado LTC 1563, que entre sus propiedades principales esta que es un filtro de 4to orden cuyas características de frecuencia pueden ser fácilmente modificable con un set de resistencias.

El esquema a implementar se observa en la Ilustración 3.2-1, el cual corresponde a la configuración como filtro de Butterworth pasabajos, con una frecuencia de corte de 1kHz, la cual fue escogida con el fin de obtener una mejor percepción del sonido, al cancelar casi completamente los ruidos mas agudos.



**Ilustración 3.2-1: Esquemático Filtro LTC**

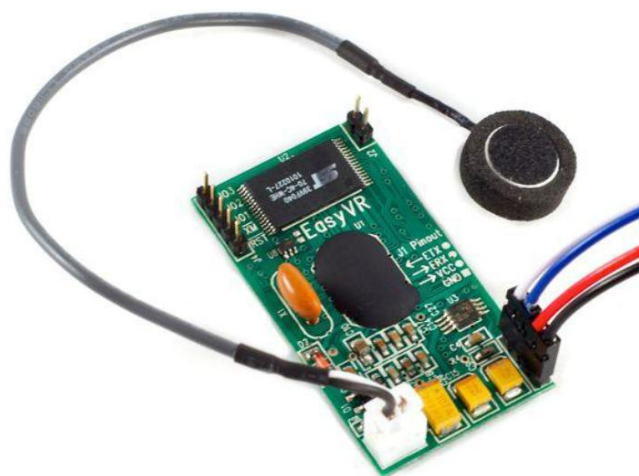
El diseño final en PCB del esquemático propuesto lo podemos ver en la Ilustración 3.2-2, en estado pre-fabricación.



**Ilustración 3.2-2: Layout filtro LTC**

### **3.3. Reconocimiento de voz**

La tercera etapa del desarrollo se trabaja con la placa de reconocimiento de voz y reproducción EasyVR de Veear. Se trata de una solución potente y de bajo costo, y que calza muy bien con el diseño propuesto.



**Ilustración 3.3-1: Easy VR**

El sistema es capaz de reconocer dos tipos de comandos, aquellos predefinidos y grabados internamente llamados voces independientes del micrófono (SI) y las grabadas y definidas por el usuario denominadas dependientes del micrófono (SD), las que deben ser entrenadas y grabadas en la memoria interna del módulo.



Existen varias aplicaciones en las que comúnmente se utiliza:

- Sistemas de control gobernados por voz
- Automatización de aplicaciones de ámbito domestico
- Control de acceso por voz
- Sistemas robóticos controlados por voz

Y entre sus características principales se encuentran:

- Contiene 26 comandos SI en ingles, italiano, español, japonés, alemán y francés.
- Admite hasta 32 comandos y sus respectivas voces definidas por el usuario, en cualquier idioma.
- Los comandos y voces pueden ser organizados en hasta 16 grupos distintos.
- Dispone de una interfaz gráfica de usuario (GUI) para Windows.
- Dispone de comunicación serie para conectar con controladores.
- Alimentación de 3.3V a 5.5V
- Contiene 3 puertos de entrada salida (E/S), que pueden ser controlados directamente mediante los comandos correctos.
- Permite la conexión de un altavoz de 8Ω mediante salida PWM de audio.
- Permite reproducir hasta 9 minutos de sonidos y/o voces dependiendo de la compresión utilizada.

Para simplificar el manejo existe una placa de desarrollo Arduino<sup>17</sup> compatible, para ser integrada con el modulo EasyVR, la cual se observa en la Ilustración 3.3-2



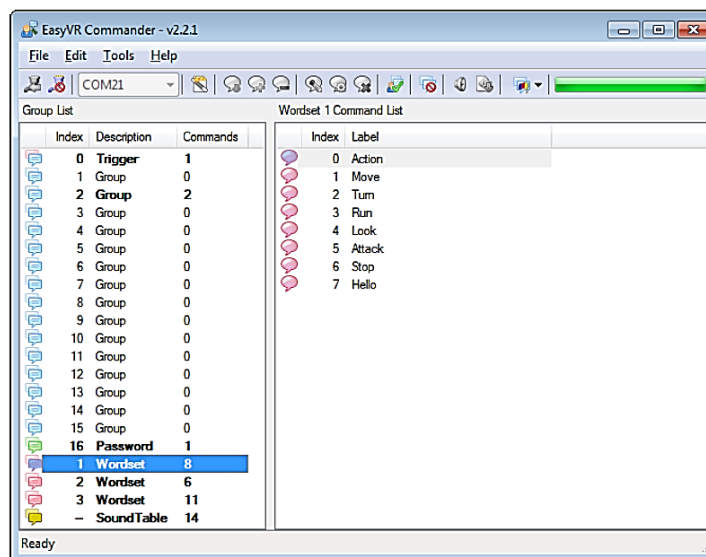
**Ilustración 3.3-2: EasyVR Shield para Arduino UNO**

<sup>17</sup> Arduino es una plataforma de electrónica abierta para la creación de prototipos basada en software y hardware flexibles y fáciles de usar. Se creó para artistas, diseñadores, aficionados y cualquiera interesado en crear entornos u objetos interactivos.

Entre sus ventajas esta que dispone de una interfaz de conexión adaptada a módulos Arduino facilitando la programación a los desarrolladores. También cuenta con LEDs indicadores de actividad.

Como se mencionó anteriormente el módulo posee un software que puede ser usado para configurar fácilmente. En este caso se conecta al PC a través de la placa Arduino compatible mediante el puerto USB. En el pueden ser definidos los comandos, generando las plantillas básicas de código para posteriormente ser modificadas con las acciones requeridas.

Estas plantillas contienen todas las librerías necesarias para el correcto funcionamiento de los procesos de reconocimiento o reproducción generados a partir de la grabación y entrenamiento de los comandos respectivos.



**Ilustración 3.3-3: EasyVR Commander**

Para el desarrollo del prototipo se han definido los comandos para apertura y cierre del canal, como sigue:

- UNO: Comando que pone en alerta de comunicación el módulo EasyVR.
- ATENTO: Comando de apertura de canal, pone en estado HIGH (5V) pin DigitalOut 11 de Arduino.
- CAMBIO: Comando de cierre del canal, ponen en estado LOW (0V) pin DigitalOut 11 de Arduino.

El comando UNO se encuentra en el índice 0 Trigger, ATENTO y CAMBIO, por su parte se encuentran en el índice 1 Group 1.

Un extracto del código ejecutado y almacenado en la memoria de EasyVR es el que se muestra a continuación:

```

void action()
{
  switch (group)
  {
    case GROUP_0:
      switch (idx)
      {
        case G0_UNO:
          group = GROUP_1;
          break;
        }
      break;
    case GROUP_1:
      switch (idx)
      {
        case G1_ATENTO:
          group = GROUP_1;
          digitalWrite(11, HIGH);
          break;
        case G1_CAMBIO:
          group = GROUP_0;
          digitalWrite(11,LOW);
          break;
        }
      break;
    }
  }
}

```

Para las primeras pruebas de implementación se trabajará con un LED, que se encenderá al nombrar el comando ATENTO y se apagará con el comando CAMBIO. Posteriormente y una vez verificado el correcto funcionamiento de esto, se procederá a unir con la cuarta etapa de desarrollo correspondiente a la activación del canal de comunicación. Cabe destacar que esta definición de comandos es absolutamente arbitraria y eventualmente podría ser modificada con el fin de mejorar los resultados.

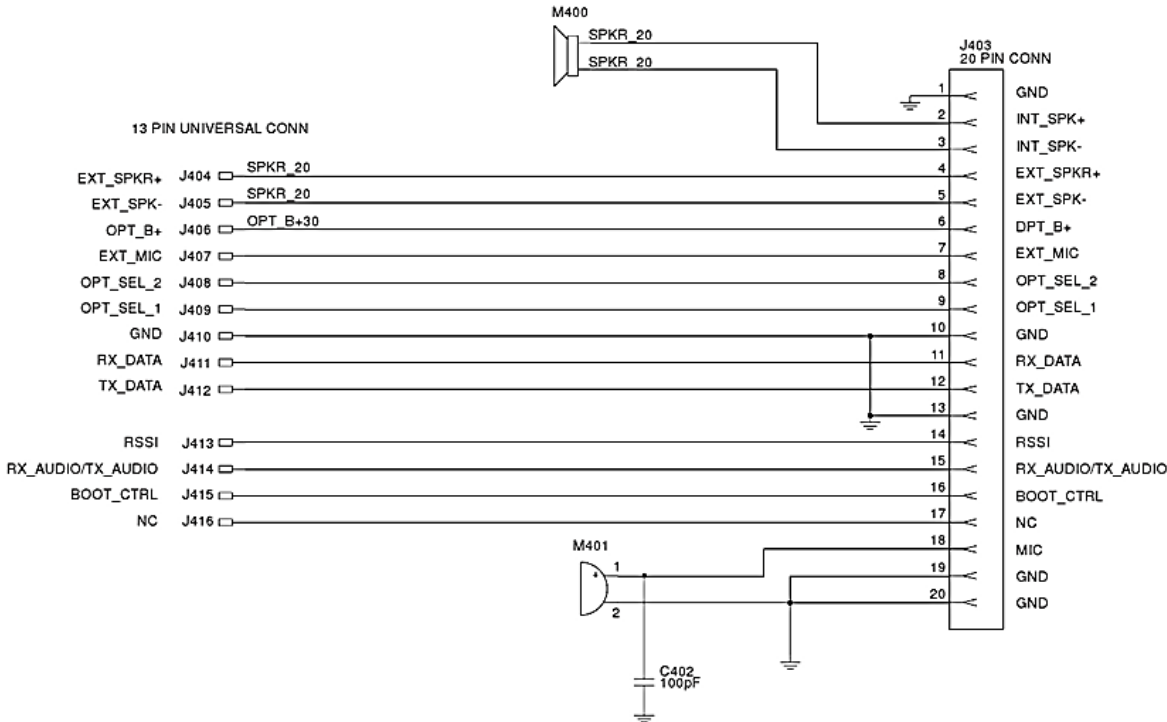
### **3.4. Activación canal de comunicación**

La última etapa del proceso es la activación del canal de comunicación, en este caso abrir la comunicación de una radio portátil Motorola PRO5150. Para ello se debe cumplir que los circuitos de control proveniente del detector de comandos y el circuito de la radio estén aislados, dado que existe desconocimiento de las características eléctricas de la radio y se podría producir un mal funcionamiento del sistema y eventualmente el daño definitivo de las etapas anteriores.

Para ello se propone el uso de un relé que es un dispositivo electromecánico que funciona como interruptor controlado por un circuito eléctrico. Este proceso se lleva a cabo por medio de una bobina y un electroimán que accionan una serie de contactos que permiten abrir o cerrar un circuito.

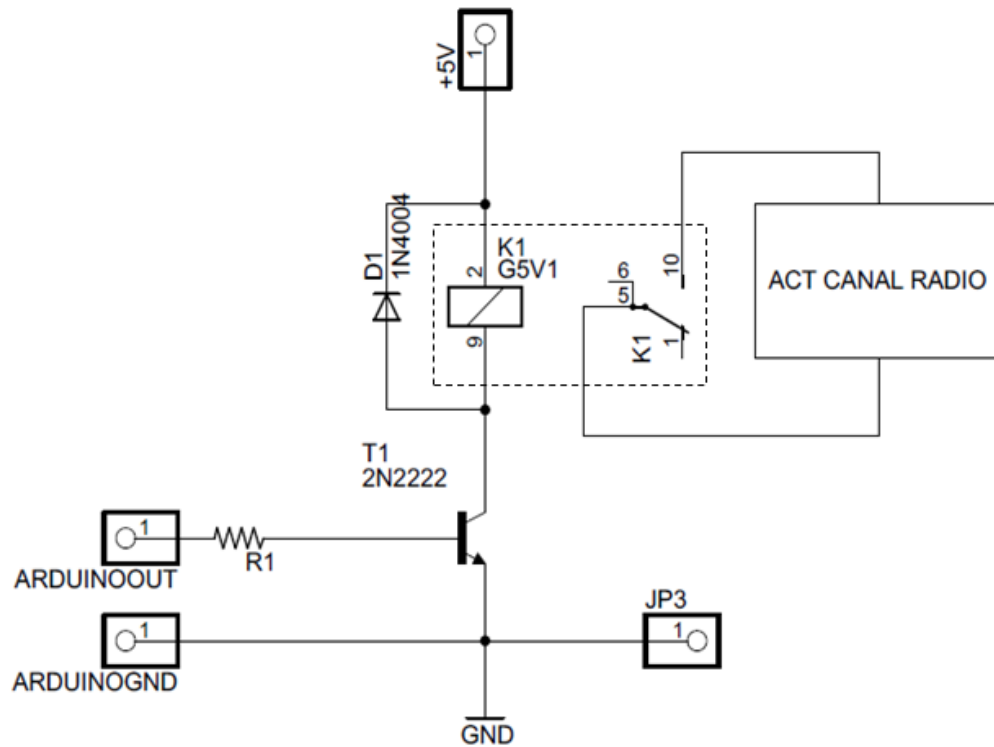
Como base del circuito esta el módulo Arduino UNO utilizado para el reconocimiento, este proveerá la alimentación de la bobina del relé, para lo cual además se utilizará un transistor 2N2222A por si la potencia entregada por el módulo no fuese suficiente. El relé a utilizar será el modelo G5V-1-DC5 de OMRON, de 5V que es lo que entrega la placa mencionada, además se utilizará un diodo 1N4007 como medio de protección ante las corrientes residuales de la bobina del relé.

Por el otro lado estará conectada la radio a través de un adaptador que permitirá acceder a los pines de control, OPT\_B+ y GND, según la siguiente descripción:



**Ilustración 3.4-1: Descripción pines de acceso radio Motorola pro5150**

El circuito final a desarrollar se muestra en la Ilustración 3.4-2:



**Ilustración 3.4-2: Esquema conexión sistema activación del canal de comunicación**

Este circuito utiliza como entrada (ARDUINOOUT) proveniente de la salida DigitalOut 11 del módulo Arduino y que fue definida en el apartado anterior con el estado HIGH para el comando ATENTO y LOW para el comando CAMBIO.

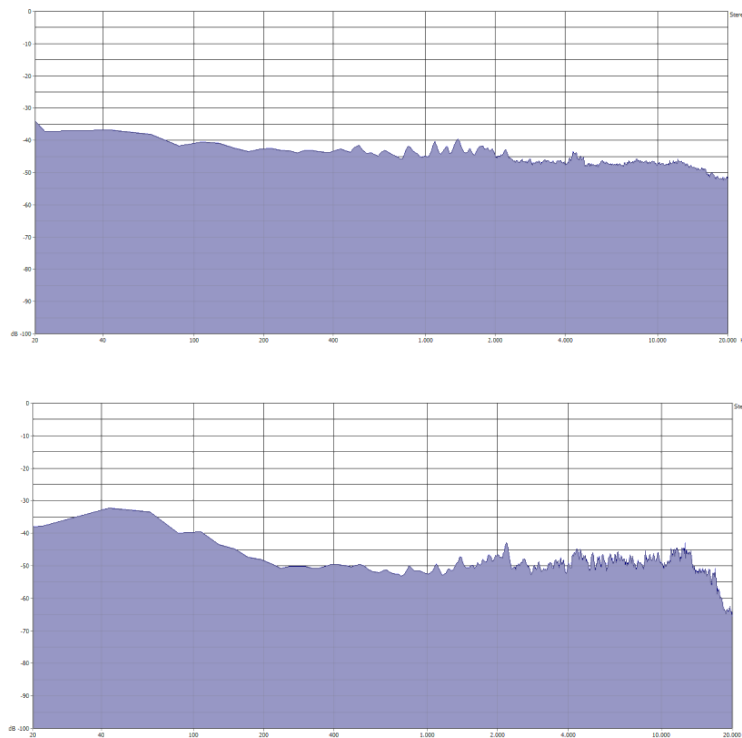
## 4. Discusión de resultados

Con el fin de entender mejor los resultados obtenidos, y lograr mayor profundidad en los análisis, éstos se realizarán por etapas; las mismas y en el mismo orden en que fueron definidas para la etapa de implementación.

### 4.1. Resultados cancelación de ruido

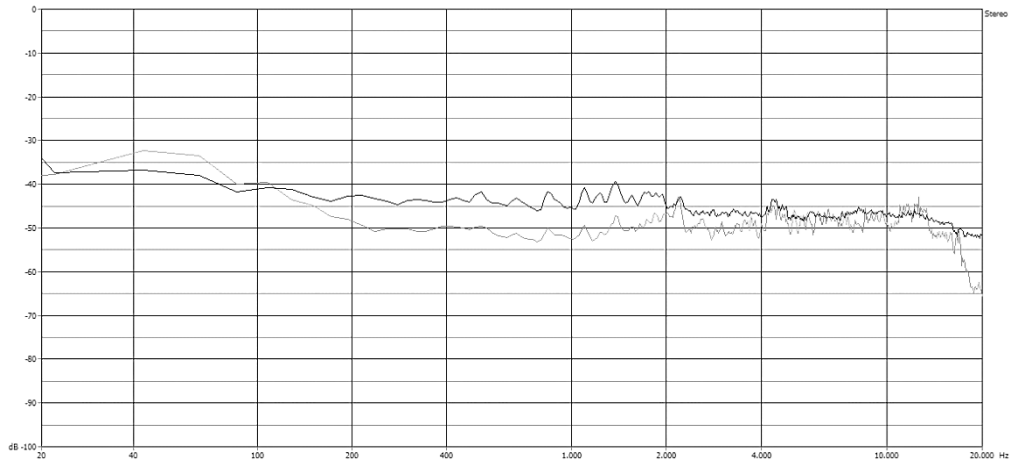
Las pruebas se realizaron en un laboratorio implementado con un equipo de sonido adaptado para llegar a niveles de ruido cercanos a 100 dBA con peaks de hasta 107 dBA. Estos no son elegidos al azar, sino que corresponden a los niveles máximos de ruido medidos la zona concentradora de la División Chuquicamata de Codelco, que es el parámetro de referencia a la hora de definir casos límites.

Las primeras pruebas están diseñadas para comprobar el correcto funcionamiento de la placa de desarrollo (ver Ilustración 3.1-5: Placa de desarrollo LMV1090), se utilizan muestras puras de ruido blanco y ruido rosa, cuyas distribuciones espectrales se encuentran descritas en la Sección 2.2. Una vez activado el ruido se habla sobre el mismo y finalmente se activa la cancelación mediante el software, las gráficas siguientes muestran las diferencias entre cada medición:



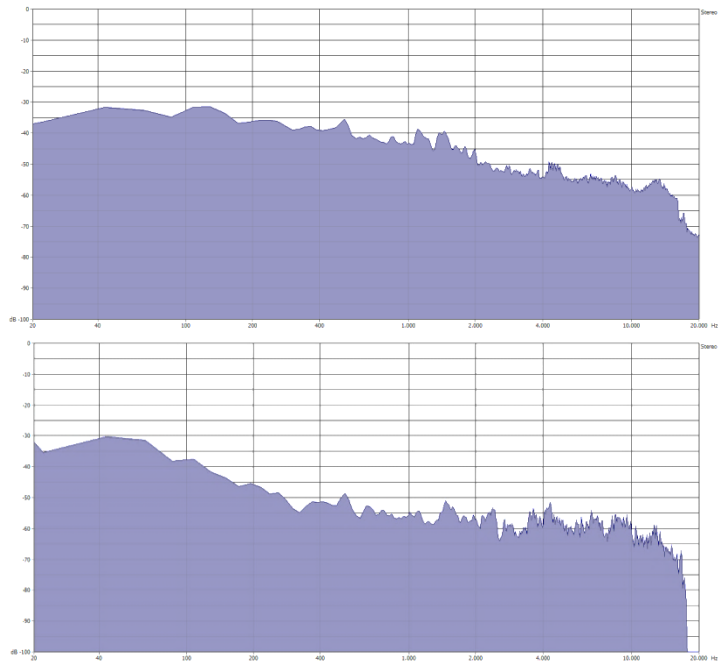
**Ilustración 4.1-1: Arriba ruido blanco con voz; abajo ruido blanco con voz y cancelación de ruido**

Dado que no resulta fácil la comparación a simple vista de las muestras, se superponen ambas para notar las diferencias:

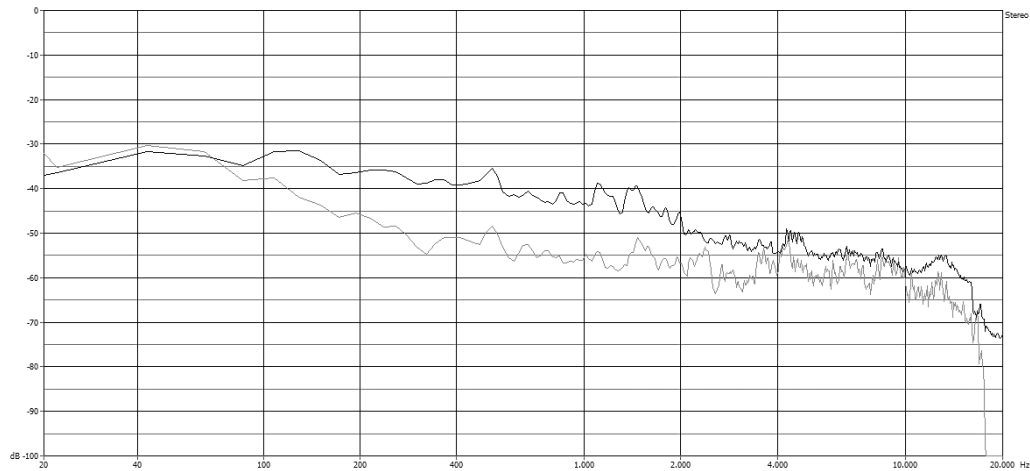


**Ilustración 4.1-2: Comparación ruido blanco voz con y sin cancelación de ruido; en negro sin cancelación; en gris la muestra con cancelación.**

Realizando el mismo ejercicio para las muestras de ruido rosa:



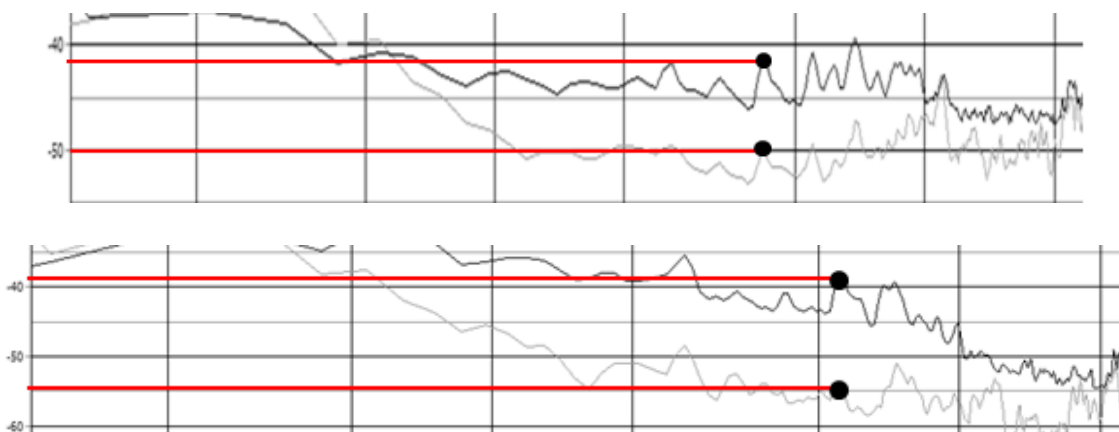
**Ilustración 4.1-3: Arriba ruido rosa con voz, abajo ruido rosa con voz y cancelación de ruido**



**Ilustración 4.1-4: Comparación ruido rosa voz con y sin cancelación de ruido; en negro sin cancelación y en gris la muestra con cancelación.**

Se observa en ambos casos la notoria disminución de ruido en especial en la banda que va de los 80 a los 4000 Hz que es en donde se concentra la voz, no así en las bandas fuera de ese rango donde prácticamente no se ve reducción.

Haciendo un acercamiento y tomando como referencia uno los puntos donde se produce reducción, se encuentra que es cercana a los 10dB para ruido blanco y a los 15dB en el ruido rosa, esto no se aleja mucho de lo que se ofrece en el datasheet del integrado LMV 1090 que son los 20 dB en las mejores condiciones.



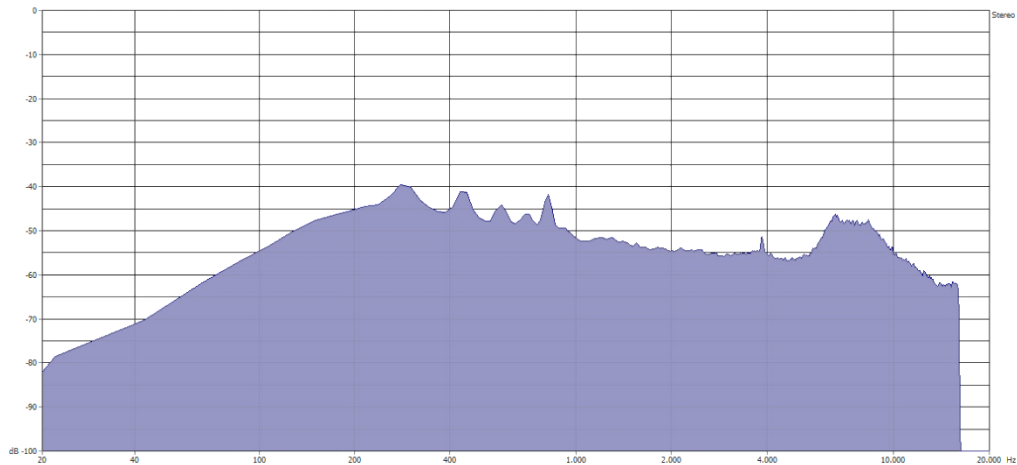
**Ilustración 4.1-5: Arriba reducción de ruido para ruido blanco; abajo reducción de ruido para ruido rosa**

Si bien estos resultados son alentadores corresponden a condiciones distintas a las que se tendrían en el dispositivo final, pues este será sometido a un tipo de ruido que no necesariamente es equivalente a alguno de los dos



anteriores y además se utilizará con los micrófonos adosados a la mascarilla cuyas características físicas también podrían influir.

Para acercarse a las condiciones reales se utiliza una muestra de ruido, desde ahora "MR1", obtenida de la zona A0 de la concentradora en la División Chuquicamata de Codelco mediante un dispositivo de grabación de audio no profesional (teléfono celular) y cuya distribución espectral es la siguiente:



**Ilustración 4.1-6: Espectro de frecuencias muestra de ruido n°1**

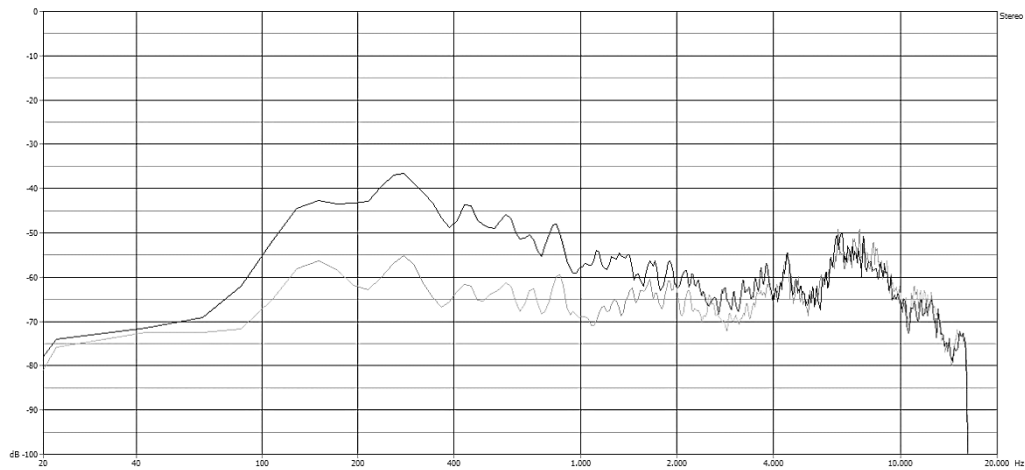
Se observan diversos peaks en frecuencias dentro del rango de 200 a los 1000 Hz., un notorio aumento en torno a los 10000 Hz., una disminución constante hacia las frecuencias bajas y por motivos que pueden ser del dispositivo usado para la grabación un corte para las frecuencias cercanas a los 20 kHz.

Utilizando MR1 se realizaron al menos 50 pruebas distintas según las condiciones dadas en la Sección 3.1.1, pero para efectos de un análisis mas focalizado se omitirán aquellas que puedan resultar redundantes y se entregan los resultados mas ilustrativos.

Unas de las pruebas importantes, modifica las pre y post amplificaciones que la placa permite, mediante el software que la controla a través del puerto I2C.

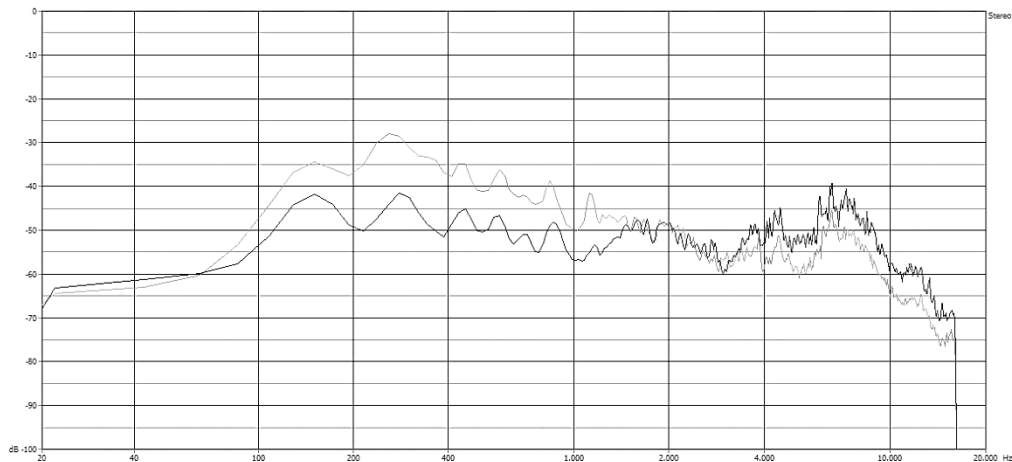
En un principio se comienza variando la pre amplificación desde los 6 hasta los 36dB, subiendo cada vez 2dB, y luego la post amplificación para cada caso desde los 6 a los 18dB subiendo cada vez 3dB.

La Ilustración 4.1-7 muestra la comparación entre los resultados obtenidos con y sin cancelación para el caso de 18dB de pre amplificación y 6dB post amplificación con los micrófonos al aire y sin máscara.



**Ilustración 4.1-7: Comparación MR1 18dB pre amplificación 6dB post amplificación sin máscara, en negro sin cancelación de ruido en gris con cancelación.**

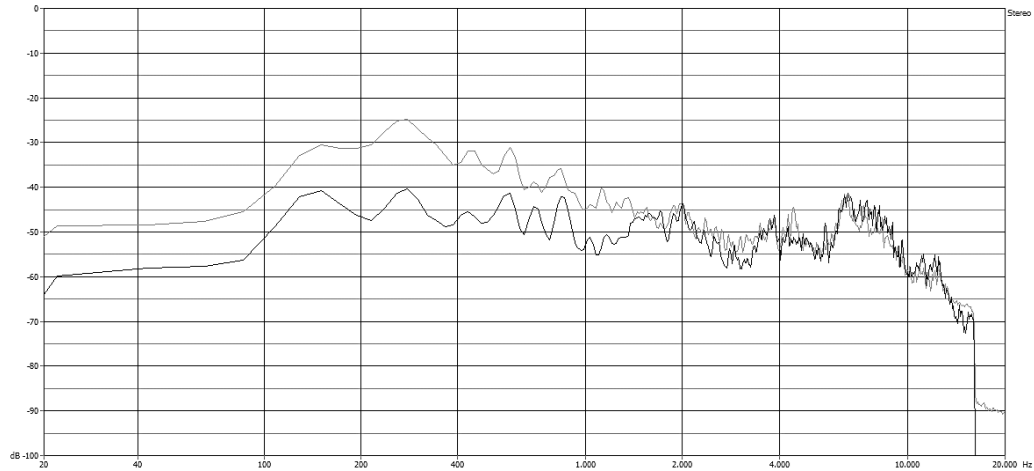
En comparación con los casos anteriores el resultado es algo mejor, dado que se obtiene una reducción de hasta 20 dB en la banda de interés, sin embargo el rango en donde se produce la mayor efectividad del sistema sigue siendo entre los 80 y los 4000 Hz.



**Ilustración 4.1-8: Comparación MR1 36dB pre amplificación 6dB post amplificación sin máscara; en negro sin cancelación de ruido; en gris con cancelación.**

Ahora si se observa el caso con 36 dB de pre-amplificación (Ilustración 4.1-8) la amplitud de la reducción de ruido es menor a los casos anteriores llegando tan solo a las 10dB y el rango en el que se produce también es menor estando entre los 100 y lo 2000 Hz aproximadamente.

Para ver que sucede si se modifica la post amplificación se tiene el caso mostrado en la Ilustración 4.1-9.

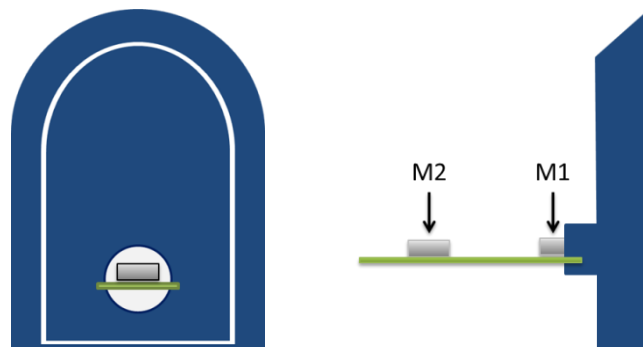


**Ilustración 4.1-9: Comparación MR1 36dB pre-amplificación 12dB post amplificación sin máscara; en negro sin cancelación de ruido; en gris con cancelación.**

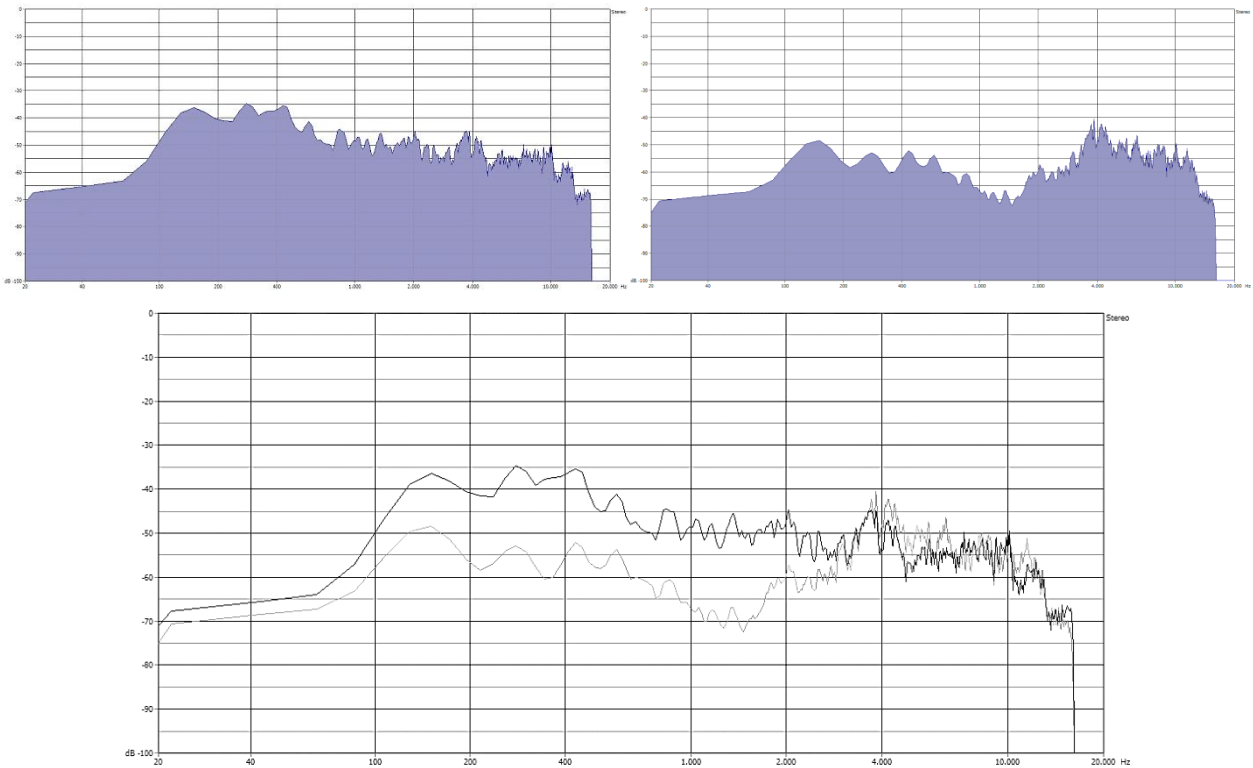
Se observa que la reducción de ruido sigue estando en los niveles anteriores, sin embargo el rango en el que se produce se acerca más a las frecuencias bajas estando ahora entre los 20 y 2000 Hz.

Las pruebas con MR1 a pesar de ser mas cercanas a lo realidad, el dispositivo final se debe utilizar con los micrófonos integrados a la mascarilla y el ruido al cual será sometido, en la mayoría de los casos, no será rosa ni blanco. Es por ello que la siguiente etapa en las pruebas se realiza con los micrófonos ubicados en distintas posiciones con diferentes modelos de tapas como se observa en Ilustración 3.1-10.

Una configuración, desde ahora "MT1", se muestra en detalle en la Ilustración 4.1-10



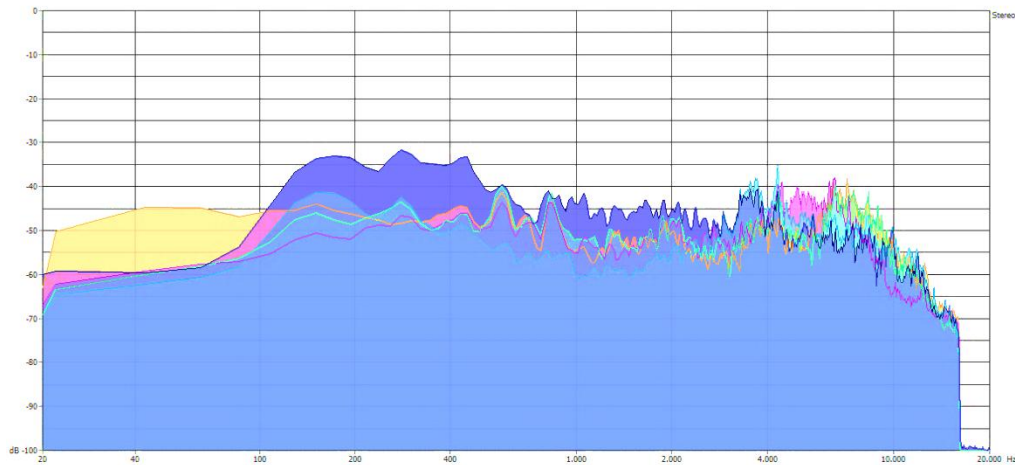
**Ilustración 4.1-10: Vista frontal y lateral ubicación micrófonos muestra tapa n°1**



**Ilustración 4.1-11: Prueba MR1 con tapa MT1 22 pre amplificación y 6 post amplificación; arriba izquierda sin cancelación; arriba derecha con cancelación; abajo comparación, en gris con cancelación en negro sin cancelación**

Cabe destacar que esta es la posición de micrófonos propuesta para el integrado LMV1090 por el fabricante y los resultados a partir de ello debiesen ser los mejores que se puedan obtener al ir variando los modelos de tapas. Ellos aparecen en la Ilustración 4.1-11, y se ve que es una mezcla de los resultados obtenidos para las pruebas anteriores, dado que se tiene la amplitud de reducción máxima lograda hasta el momento así como el rango más amplio de frecuencias en el que se produce reducción, en otras palabras se logra reducir hasta 20 dB dentro del rango que va de los 20 a los 4000 Hz. Por lo que se comprueba que la posición actual es la más exacta para lograr un mejor resultado.

Para despejar dudas se realizan pruebas modificando tanto el modelo de tapa como la posición, los resultados son muy similares para las distintas configuraciones sin embargo MT1 sigue siendo la mejor, a continuación se muestra el resultado espectral de las reducciones de ruido para cada una de las muestras realizadas.

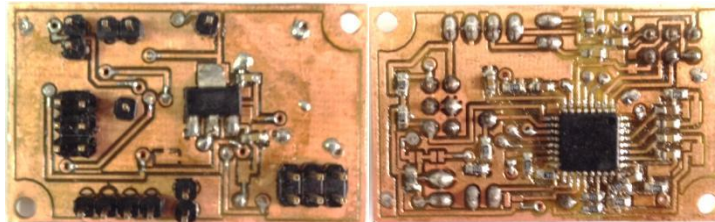
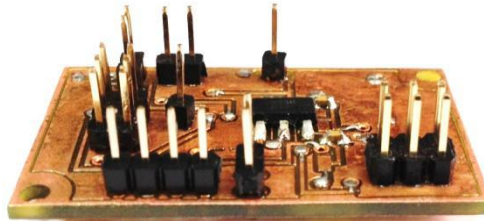
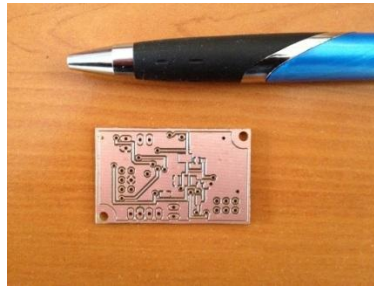


**Ilustración 4.1-12: Pruebas distintas posiciones micrófonos y modelos tapas**

Se observa en la Ilustración 4.1-12 en azul oscuro la muestra sin cancelación con pre amplificación de 36 dB y post amplificación de 6 dB, en los otros colores las distintas pruebas con cancelación. Claramente todos los resultados con cancelación están en un nivel mas bajo a excepción de a muestra en naranja cuyas frecuencias bajas están a un nivel más alto que el resto. Sin embargo ninguna llega a los 20 dB de reducción ni tampoco al amplio rango de frecuencias con reducción como en el caso MT1 por lo que se comprueba que la mejor solución es con los micrófonos en posición frontal y con la muestra de tapa vista en la Ilustración 4.1-10.

Los resultados anteriores fueron obtenidos con la placa de desarrollo, sin embargo y como se menciona en el capítulo 3, se fabricó un prototipo de placa basada en el chip LMV1089 (gracias a la facilidad de construcción que entrega uno de sus empaquetamientos superficiales, capaz de ser integrado en una pcb de tan solo 2 capas). El proceso se llevó a cabo en el Laboratorio de Electrónica de la Facultad donde se encuentran los implementos necesarios para la producción de una placa de mas alto nivel.

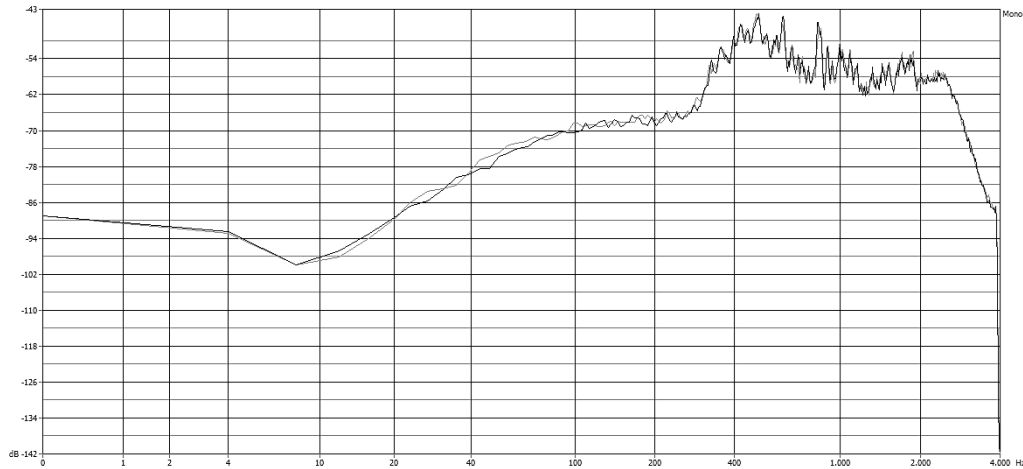
El resultado de este trabajo se observa en la Ilustración 4.1-13



**Ilustración 4.1-13: Placa LMV1089 fabricación propia**

A pesar de que la placa se diseñó intentando conservar las mismas características funcionales que posee la placa de desarrollo del LMV1089 necesarias para el correcto funcionamiento del dispositivo, a la hora de realizar las pruebas, no se obtuvo buenos resultados, si bien la placa se mantenía funcionando y se obtenía señal a la salida, no se logró comprobar si realmente se produce la reducción de ruido, pues no se establecieron claras diferencias en los resultados con y sin cancelación. Pudiendo estar entre las causas, fallas a la hora de la fabricación, dado que gran parte del proceso es manual y no es posible comprobar fehacientemente que todos los componentes integrantes funcionan de manera correcta.

Una muestra de la salida del circuito se muestra en Ilustración 4.1-14, donde se observa que la frecuencia está limitada en rango, pues sobre los 4000 Hz se pierde por completo y además no se observan diferencias claras entre cancelación y sin cancelación.

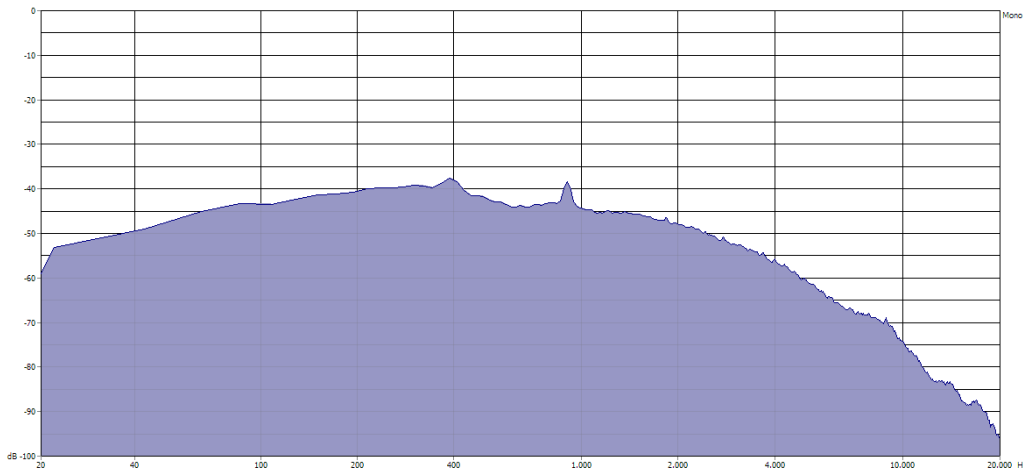


**Ilustración 4.1-14: Salida placa diseño propio**

La última etapa de pruebas de cancelación se lleva a cabo bajo las condiciones reales de la mina, para ello se viajó a Chuquicamata y en lo que el tiempo (1 hora en el lugar) y las condiciones permitía, se realizó una serie de pruebas similares a las anteriores y los resultados se analizan a continuación.

Para comenzar se analiza que tan distinta es la muestra inicial de ruido MR1 de la muestra real de ruido que a partir de ahora será "MRM" (muestra ruido mina).

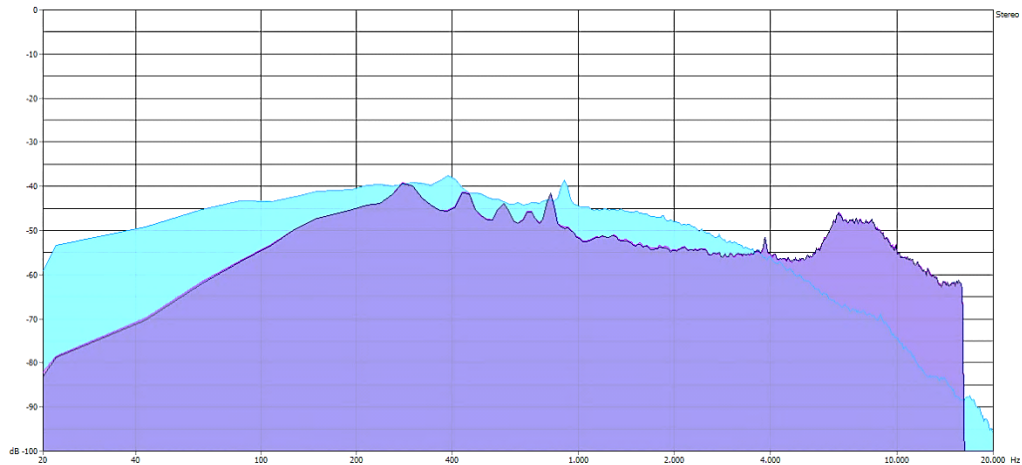
Para obtener la muestra de ruido mina se utilizó un micrófono especial de características profesionales, cuya distribución en frecuencias es prácticamente constante lo que entrega resultados muy cercanos a los reales y bastante independientes de las características del dispositivo de grabación.



**Ilustración 4.1-15: Muestra ruido mina**

Es posible destacar de la Ilustración 4.1-15 que el espectro es distinto al ocupado para las pruebas anteriores, destacando un alza en los bajos, y con un peak en las frecuencias medias que se encuentran en el rango de los 200 a los 1000 Hz. Esto auditivamente hablando significa una mejora dado que la percepción de las bajas frecuencias no es tan molesta, como la de las altas.

La comparación entre ellas se muestra en la Ilustración 4.1-16:

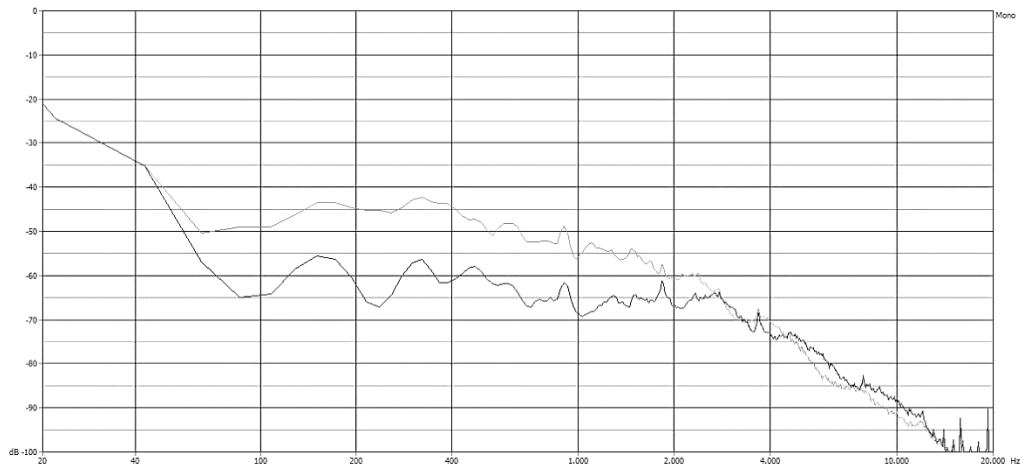


**Ilustración 4.1-16: Comparación MR1 con MRM; en azul claro MRM, en azul oscuro MR1.**

Los análisis realizados con esta muestra de ruido se realizaron al igual que los casos anteriores grabando las pruebas directamente al computador, lo que eventualmente podría introducir pequeños cambios en las muestras.

A continuación se observa una comparación de una prueba realizada con los micrófonos al aire, esto es, al igual que las primeras pruebas, sin máscara, primero grabando sin cancelación y luego activando la cancelación de ruido.

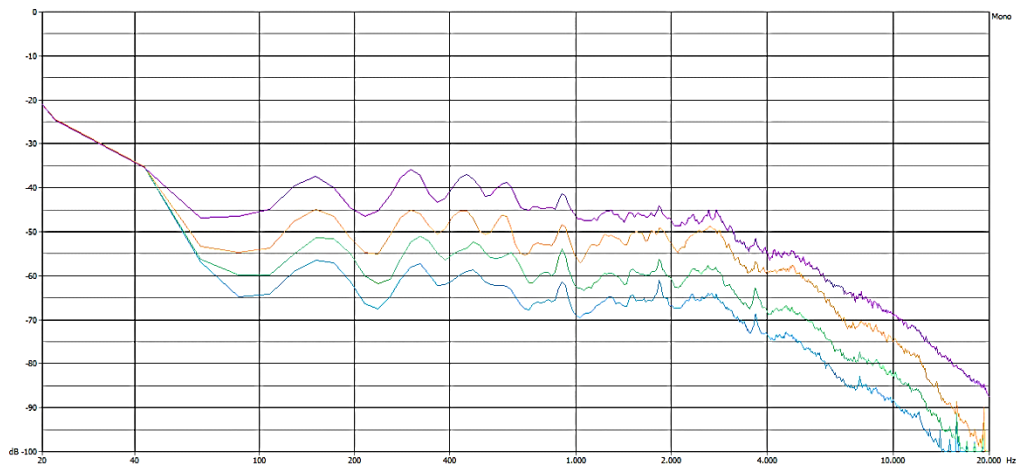




**Ilustración 4.1-17: Comparación MRM 6dB pre amplificación 6dB post amplificación sin máscara; en negro con cancelación de ruido; en gris sin cancelación.**

Nuevamente se ve que la cancelación se produce en un amplio rango en este caso de los 40 a los 4000 Hz y con una amplitud máxima de reducción de 20 dB lo que se encuentra en el mejor de los casos, sin embargo existe una situación además favorable que es la mencionada reducción del ruido en las altas frecuencias.

Además podemos comparar que pasa si se aumenta la pre-amplificación de los micrófonos:



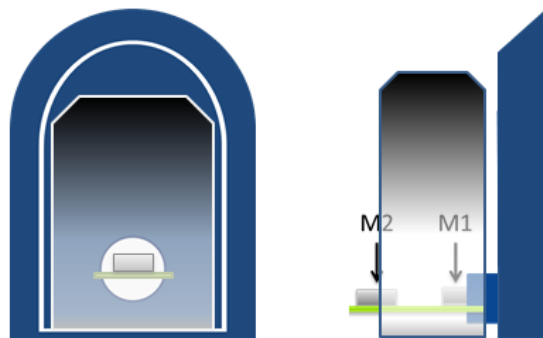
**Ilustración 4.1-18: Comparación MRM con cancelación; de abajo hacia arriba 6, 12, 18 y 22 dB**

Se ve que coincidentemente con el aumento de pre amplificación se produce el aumento de los niveles audio, pero sin cambiar las frecuencias, por lo que si el tema final de la reducción del ruido es mejorar la comunicación y en el marco teórico se estableció que a mayores niveles de audio esta empeoraba,

la propuesta es trabajar con pre amplificaciones que no estén tan bajas como para no escuchar ni tan altas como para molestar (en el rango de los 18 a los 24 dB).

Además es posible notar que los resultados no se alejan mucho de los obtenidos con "MR1" por lo que estos últimos se validan para efectos de simulación.

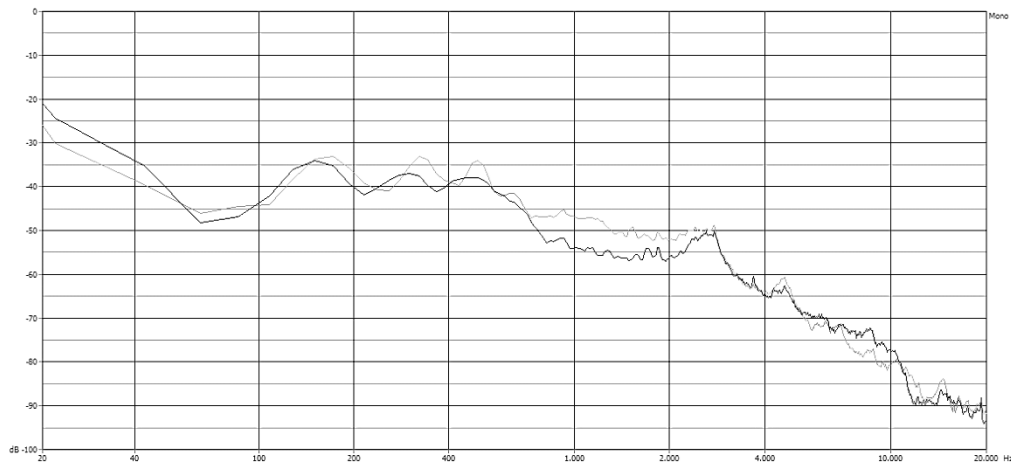
El último paso de esta discusión es revisar los resultados de hacer las pruebas con los micrófonos en la mascarilla, esto es, en la misma posición mostrada en Ilustración 4.1-10, pero con una pieza diseñada especialmente para ello, que deja los micrófonos en la misma posición pero protegidos de agentes externos como polvo o humedad.



**Ilustración 4.1-19: Vistas esquemáticas frontal y lateral nueva pieza**

Sin embargo los resultados de las pruebas no son buenos ya que no se observa una reducción considerable del ruido, lo que puede deberse a que en esta posición los 2 micrófonos no están escuchando lo mismo, por lo que el sistema no tiene la posibilidad de analizar las diferencias entre ellos.

El ruido no es escuchado de la misma forma por ambos y por ende la resta de este no es correcta, lo que podemos comprobar gráficamente al observar la comparación entre el sistema con cancelación y sin cancelación.



**Ilustración 4.1-20: Comparación MRM con pieza modificada, 6 pre amplificación 6 post amplificación; en gris sin cancelación, en negro con cancelación.**

De la Ilustración 4.1-20 es claro que los resultados se alejan de los anteriores y a pesar de existir una leve reducción de los niveles con cancelación no alcanza para considerarlos como una prueba exitosa.

## 4.2. Filtros

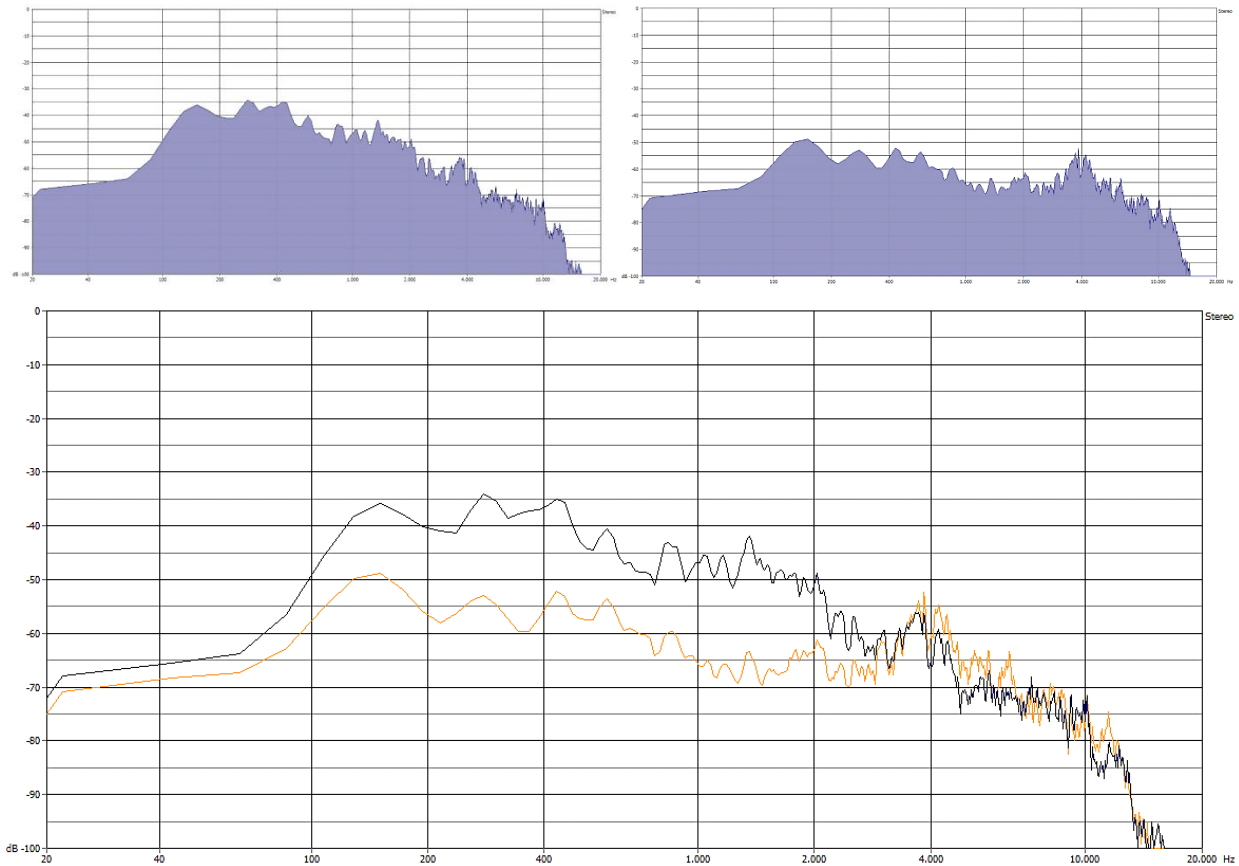
La siguiente etapa es analizar la inclusión de un filtro al sistema con el fin de reducir los ruidos molestos que pudiesen quedar post-cancelación. Se estima conveniente no introducir un filtro pre-cancelación, ya que al ser un arreglo de dos micrófonos, se debería tener un filtro para cada uno ellos pudiendo ocasionar desfases de las señales a la entrada del sistema de cancelación reduciendo la efectividad.

Se trabaja principalmente con filtros pasabajos pues se considera necesario eliminar o reducir sólo las frecuencias altas pues niveles por sobre los 2000 Hz resultan mas molestos. Además si se considera la muestra de ruido nº1, que es con la que se realizaron las pruebas principales, esta tiene un peak de ruido sobre los 4000 Hz, de modo que un filtro pasabajos a 1 kHz eliminaría por completo estas frecuencias.

Del capítulo anterior se sabe cuales son los casos con mejores resultados, a partir de esto y para no redundar se analizaran sólo ellos. Se considerará favorable el que tiene una pre amplificación de nivel medio, esto es mayor a 18 dB y menor a 24 dB, y una post amplificación de 6 dB.

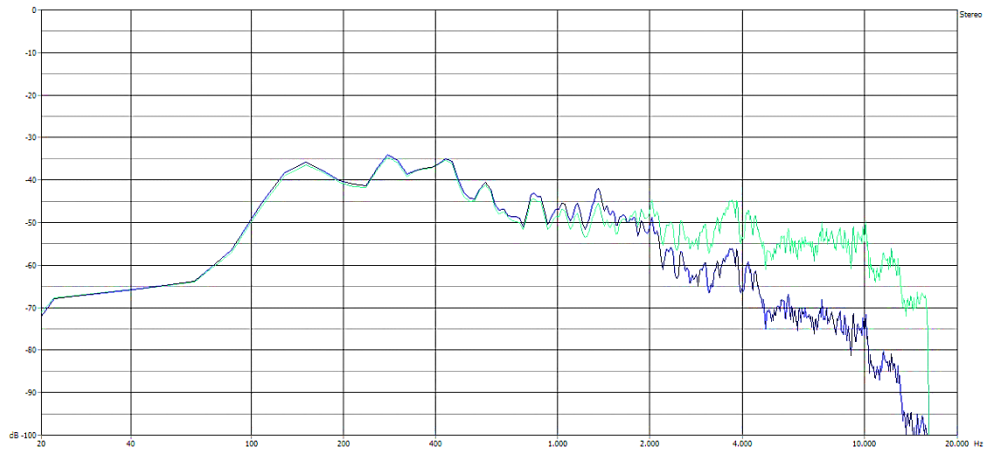
El primer caso es con MR1 y con los micrófonos al aire. Los resultados de cancelación están en los rangos esperados, esto es reducción de 20 dB en el máximo y dentro del rango que va de los 20 a los 4000 Hz. El efecto de introducir el filtro pasabajos es el esperado de reducir las frecuencias más altas. Es así como a partir de los 4000 Hz observamos una continua

decaencia en los niveles de audio de estas, gráficamente es lo que se ve en la Ilustración 4.2-1.



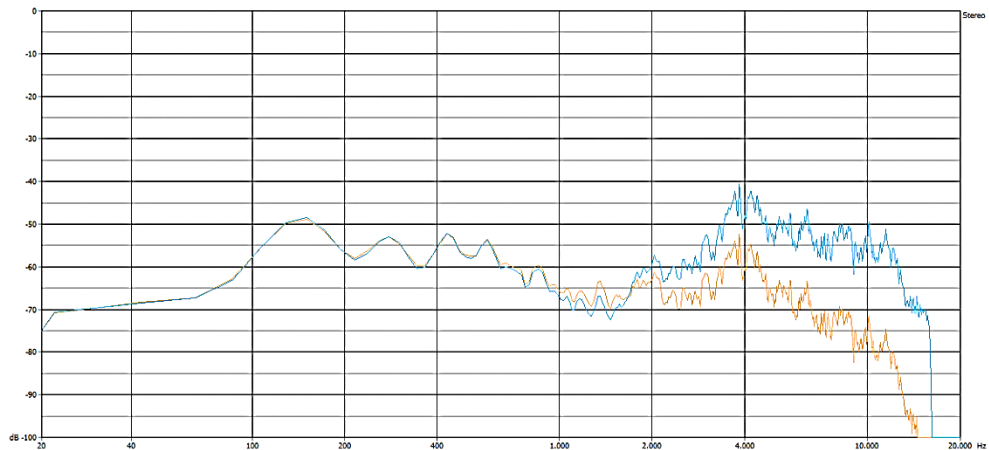
**Ilustración 4.2-1: MR1 con filtro pasabajos 1Khz pre amplificación 22dB post amplificación 6 dB; arriba izquierda sin cancelación; arriba derecha con cancelación; abajo comparación**

Un caso importante a considerar es cuando se ve que sucede al introducir el filtro directamente sin realizar la cancelación, los resultados son claros al notar que tan solo las frecuencias altas se ven afectadas mejorando la percepción del ruido. La aplicación del filtro no mejora la "claridad" de la voz, pues no modifica las frecuencias entre los 100 y 2000 Hz, sin embargo la convierte en una voz mas "nasal", lo que podría considerarse una condición no ideal, para lo que se espera del sistema.



**Ilustración 4.2-2: MR1 con filtro pasabajos 1Khz pre amplificación 22dB post amplificación 6 dB; verde sin cancelación; azul sin cancelación y con filtro pasa bajos.**

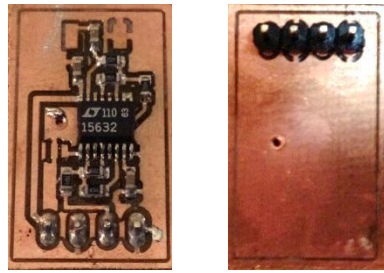
El mismo efecto se produce al insertar el filtro posterior a la cancelación, sin embargo el resultado es distinto, pues se logra disminuir el ruido en la banda de la voz y además se reducen los niveles en las frecuencias mas altas provocando la mejora perceptiva, es decir el mejor resultado hasta el momento. La única situación que se puede considerar contraria es que al eliminar gran parte de las frecuencias el sonido de la voz se vuelve más nasal que en el caso sólo con cancelación, pero para temas de recepción del mensaje no es gran problema dado que la cancelación influye en la mejora de la inteligibilidad.



**Ilustración 4.2-3: MR1 con filtro pasabajos 1Khz pre-amplificación 22dB post-amplificación 6 dB; azul claro sin cancelación; naranja sin cancelación y con filtro pasa bajos.**

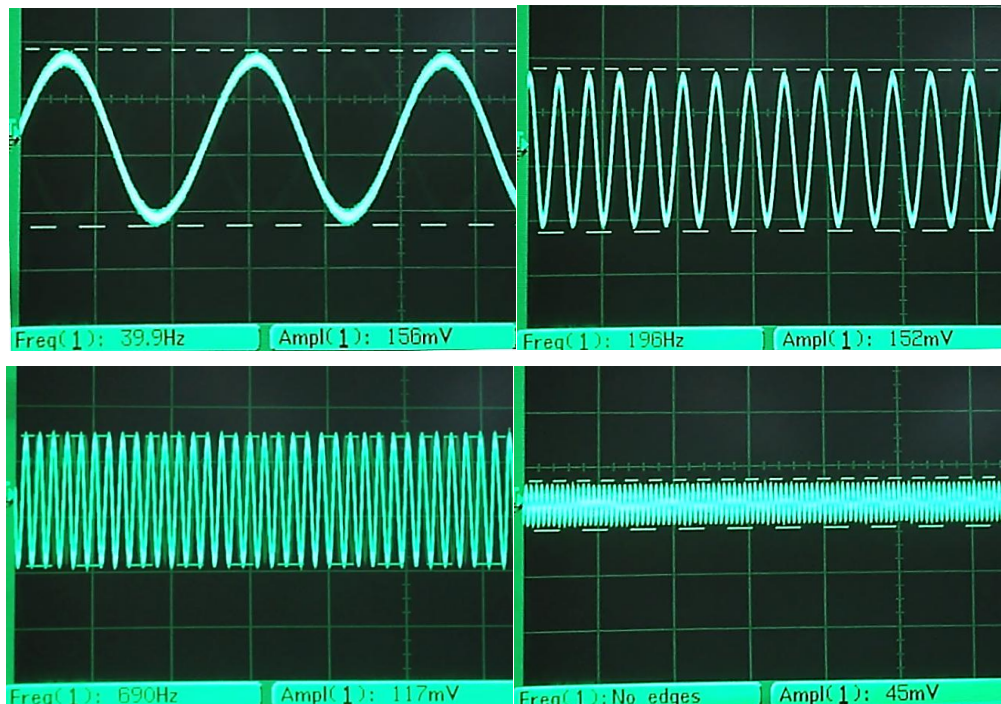
El siguiente paso fue el de la fabricación del filtro basado en LTC 1563 a partir del esquemático y layout mostrados en el capítulo 3 (Ilustración 3.2-1 e Ilustración 3.2-2), bajo las mismas condiciones en que se fabricó la placa

para el cancelador de ruido. Los resultados se muestran en las siguientes imágenes:



**Ilustración 4.2-4: Filtro LTC**

Las primeras pruebas de este resultaron correctas dado que se comprobó que reduce las frecuencias por sobre los 1000 Hz hasta casi hacerlas desaparecer desde los 2000 Hz en adelante. Esto se comprobó ingresándole una señal sinusoidal de 156 mV peak to peak (nivel similar a la señal de salida de la placa de cancelación) desde un generador de señales y observando la salida en un osciloscopio, las capturas de este muestran que la señal por sobre los 1000 Hz llegan a niveles de aproximadamente 22 mV peak to peak.



**Ilustración 4.2-5: Resultados filtro pasa bajos; arriba izquierda 39,9 Hz 156 mV; arriba derecha 196 Hz 152 mV; abajo derecha 690 Hz 117 mV; abajo derecha 1050 Hz 45 mV**

Lamentablemente el buen resultado obtenido en el laboratorio, no pudo ser comprobado directamente en terreno pues a la hora de ser conectado a la

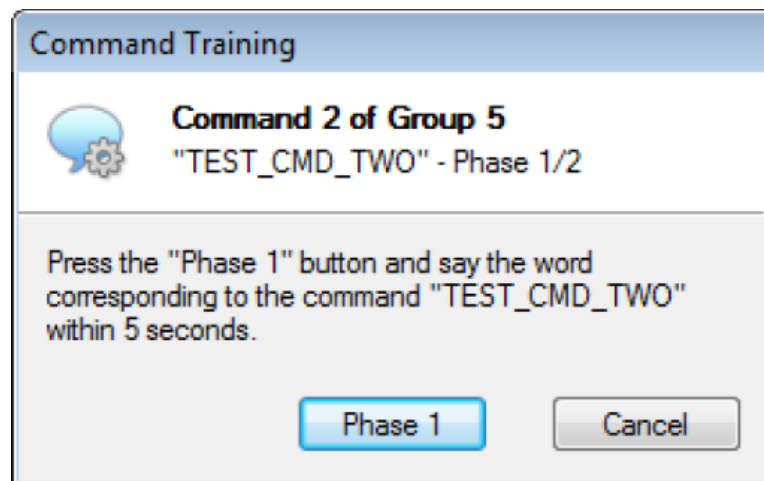
radio el filtro se sobre calentó y se quemó, además los plazos de realización no permiten la fabricación otra placa. Ésta es directamente una consecuencia del desconocimiento del circuito interno y detalles de funcionamiento del comunicador, por lo que resulta interesante plantearse a futuro una nueva fabricación y prueba del dispositivo, conocidos los parámetros internos de la radio portátil.

### 4.3. Reconocimiento

El proceso de reconocimiento se llevó a cabo, trabajando de la forma recomendada por el fabricante del dispositivo EasyVR. Esto incluye realizar los respectivos procesos de creación y entrenamiento de parámetros de voz, comprobación de detección, generación de código y pruebas de implementación.

La comunicación entre la placa de reconocimiento y el computador se realiza mediante la conexión USB a través del puerto serial COM3, razón necesaria para instalar controladores específicos para este dispositivo.

El primer paso consiste en la creación de un nuevo grupo de comandos en el dispositivo, este debe considerar al menos los dos necesarios y que se han definido en la implementación ATENTO y CAMBIO (pertenecientes al grupo\_0), además se definió un comando especial UNO (perteneciente al grupo\_trigger), que será el encargado de permitir el paso hacia los comandos de apertura y cierre de canal, este proceso se realiza mediante el programa EasyVR commander, en dos etapas (Phase 1 y Phase 2), donde primero el sistema graba el comando y luego lo comprueba.



**Ilustración 4.3-1: Entrenamiento comando**

Una vez grabados y entrenados los comandos se genera el código y se modifica en función de las salidas esperadas para cada caso, para este caso

levantando o bajando el nivel lógico del DigitalPin11, para cada comando, la secuencia en que se realiza en el proceso de reconocimiento, con sus respectivos tramos de código es:

- Usuario: UNO
- Dispositivo: Escucha, sale de modo espera y permite el dictado un comando del grupo\_0

```
Código: case GROUP_0:  
    switch (idx)
```

- Usuario: ATENTO
- Dispositivo: Escucha, pasa el pin DigitalOut11 del dispositivo de modo bajo (0V) a modo alto (5V) y permite el dictado del siguiente comando del grupo\_1

```
Código: case G1_ATENTO:  
    group = GROUP_1;  
    digitalWrite(11, HIGH);  
    break;
```

- Usuario: CAMBIO
- Dispositivo: Escucha, pasa el pin 11 de modo alto a modo bajo y pasa el dispositivo a modo espera.

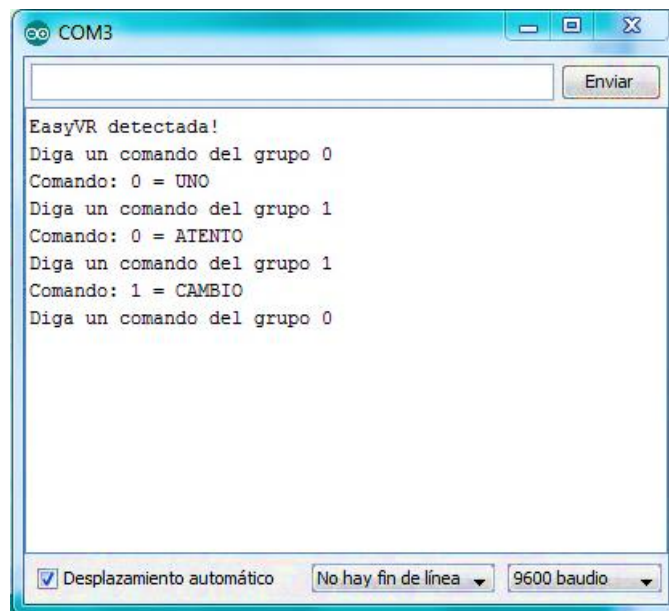
```
Código: case G1_CAMBIO:  
    group = GROUP_0;  
    digitalWrite(11, LOW);  
    break;
```

Si bien los resultados del reconocimiento son correctos, el entrenamiento no es un proceso simple, muchas veces la comparación que realiza entre la fase 1 y 2 resulta incorrecta sólo por decir de una forma un poco distinta uno de los comandos misma situación si se realiza con ruido externo, lo que impide un correcto funcionamiento del sistema general pues si bien el sistema anterior reduce el ruido no lo logra por completo.

El posterior proceso de reconocimiento se produce correctamente siempre que el usuario que dé los comandos sea el mismo que entrenó el sistema, o en su defecto alguien con un tono de voz similar, esto se transforma en un problema pues en lo ideal debiese ser una sistema robusto y que en el mejor de los caso tenga que ser entrenado una sola vez y permitir el reconocimiento de todos los eventuales usuarios del sistema.

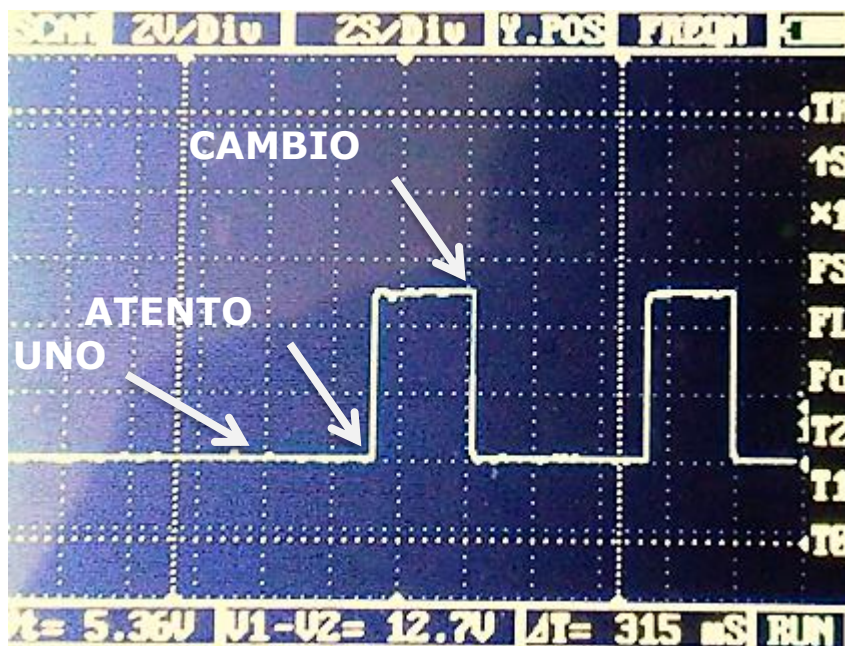
La Ilustración 4.3-2 muestra una captura del monitor serial de Arduino con la secuencia detectada de forma correcta por parte de la placa de reconocimiento:





**Ilustración 4.3-2: Resultados reconocimiento**

El proceso de activación de las salidas digitales posteriores al reconocimiento se observa en las siguientes capturas de osciloscopio en que aparecen claras las diferencias de voltaje entre una etapa y otra, indicando el correcto funcionamiento del sistema.

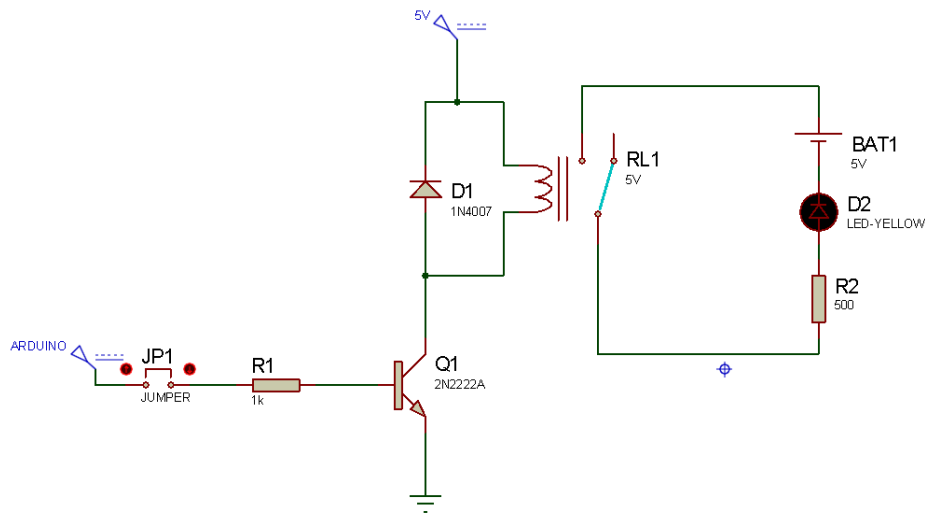


**Ilustración 4.3-3: Captura osciloscopio activación y desactivación salida DigitalOut11**

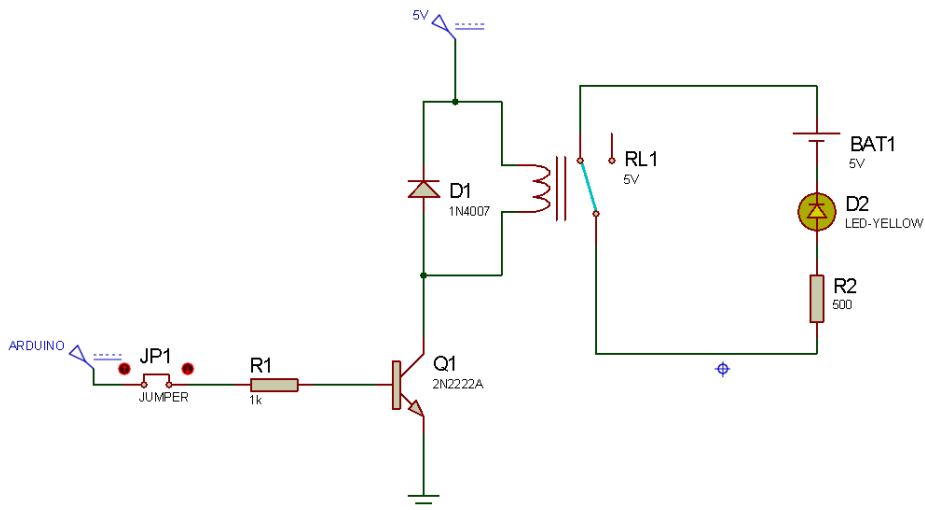
#### 4.4. Activación Canal

El proceso de activación del canal es el último en diseñar y desarrollar, dado que al menos se necesita tener resuelta la etapa de reconocimiento del comando. Sin embargo se puede tratar como una etapa independiente pues sólo basta con enviarle un pulso para realizar la prueba de activación del relé.

El circuito diseñado se simuló en el programa PROTEUS y los resultados se exponen a continuación

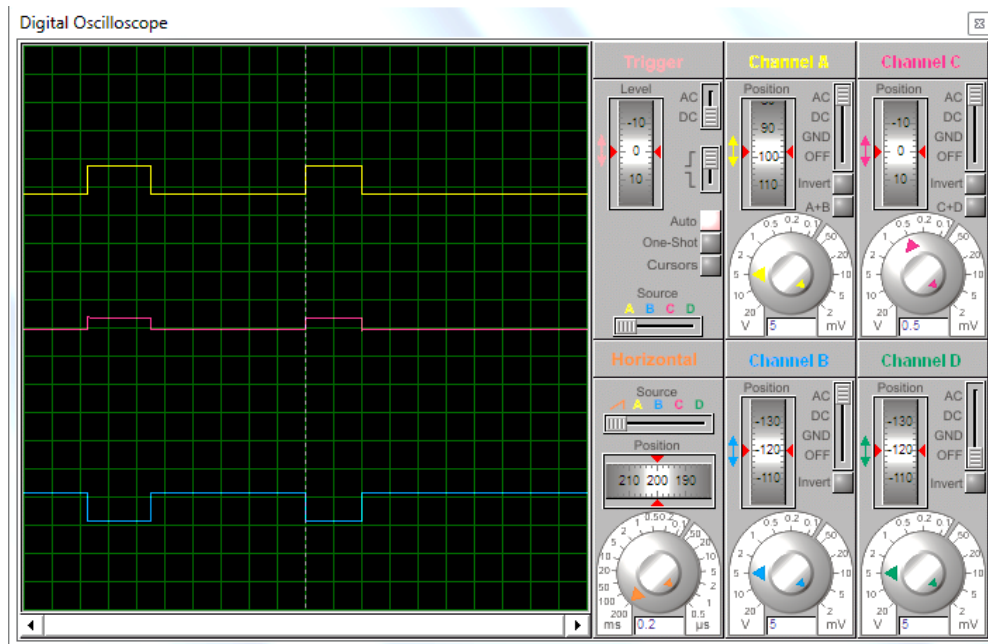


**Ilustración 4.4-1: Circuito abierto, canal desactivado**



**Ilustración 4.4-2: Circuito cerrado, canal abierto**

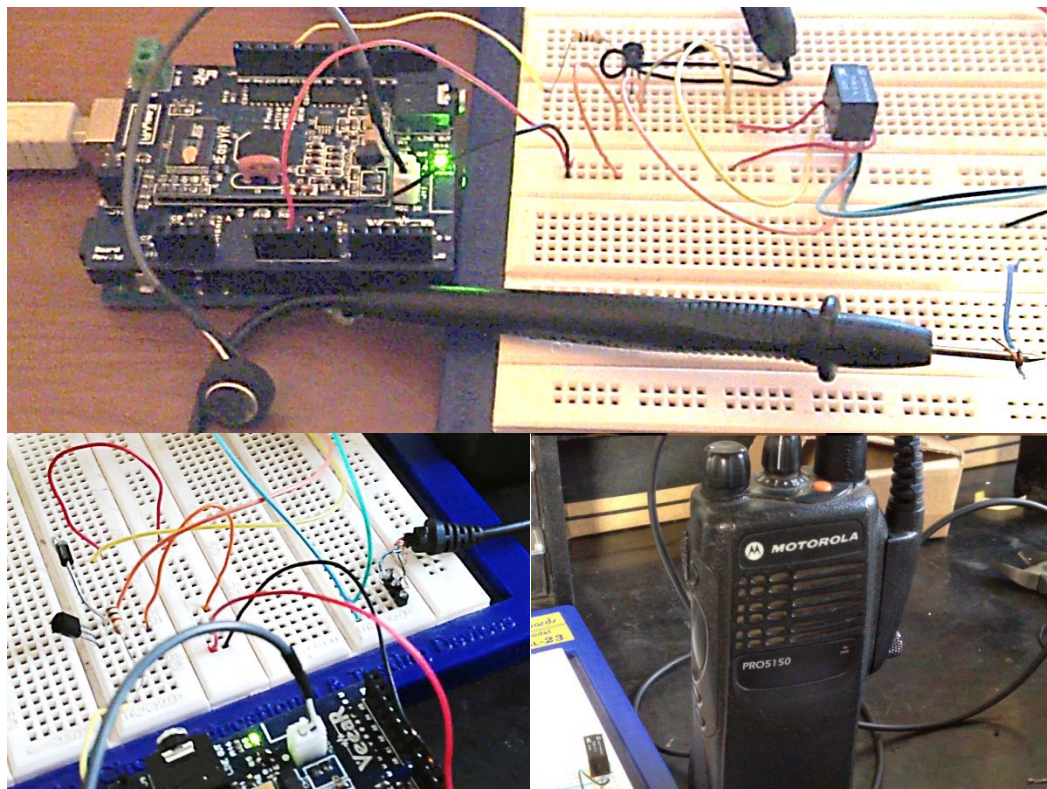
Para temas propios de la simulación, se representa la radio por el sistema resistencia-led-batería, y se observa que el led se enciende al cerrar el jumper JP1 que equivale a activar el pin DigitalOut11 del sistema de reconocimiento.



**Ilustración 4.4-3: Resultados simulación activación canal**

En la Ilustración 4.4-3 se observa las diferentes respuestas ante la activación de la salida de la placa de reconocimiento, en amarillo el voltaje a la entrada en el colector del transistor, en azul la activación del circuito del led, y en rosado el voltaje a la entrada en la base.

Al desarrollar el prototipo se obtienen los mismos resultados que en la simulación, logrando con éxito activar/desactivar el canal, sin embargo por temas propios del trabajo, donde se debe escuchar el resultado, es imposible reproducirlo de manera escrita. El montaje del sistema real lo podemos observar en la Ilustración 4.4-4.



**Ilustración 4.4-4: Montaje circuito activación**

## 5. Conclusiones

Un dispositivo que permite comunicarse por voz a través de la máscara respiratoria sin la necesidad de uso de un botón "push-to-talk" le agrega una característica tecnológica al elemento de protección personal y además aumenta los niveles de seguridad del operario, al liberarlo de la necesidad de tener que utilizar sus manos, en esta tarea, convirtiendo el dispositivo en un verdadero sistema manos libres.

Una reducción considerable de los niveles de ruido a la hora de comunicarse permite aumentar la calidad del trabajo realizado, pues puede disminuir la cantidad de veces en que se oye defectuosamente un mensaje o instrucción, convirtiendo a la máscara mas que en un elemento de protección personal en un sistema de comunicación indispensable para ambientes de ruido extremo en los cuales la comunicación normal no es posible

El sistema de cancelación logra reducir los niveles de ruido ambiente hasta en 20 dBA, lo que es bastante aceptable. Si se considera que normalmente la mina esta sometida a niveles de 100 dBA, esta reducción permite pasar de un estado donde la voz es completamente enmascarada por el ruido a uno donde los niveles de voz provenientes desde los micrófonos superan a los del ruido, mejorando la calidad de la comunicación

El sistema de filtrado de la señal es capaz de mejorar los resultados de la etapa de cancelación, produciendo una disminución de los niveles de ruido en aquellas bandas en que en principio no se logra. Además entrega una flexibilidad a la hora de trabajar con el sistema completo, pues permite eliminar frecuencias de ruido específicas para cada ambiente, por ejemplo en una faena minera se podría eliminar la banda de ruido que incluye la trituración de rocas o en una empresa metalúrgica el sonido del corte de una barra de fierro. Sin embargo el sistema se debe utilizar posterior a la etapa de cancelación de ruido, para evitar efectos de desfases entre las señales que podrían reducir la efectividad. Un punto importante a destacar es que la sola aplicación de la etapa de filtrado al sistema, esto es, sin aplicar cancelación de ruido, no otorga mejorías notorias y además provoca un efecto de volver la voz mas nasal.

El reconocimiento de voz implementado a pesar de ser un sistema simple y robusto sólo logra buenos resultados en condiciones ideales, esto es con bajos niveles de ruido, por lo que no es adecuado para el dispositivo diseñado dado que normalmente encontrará condiciones de ruido adversas que podrían significar un mal funcionamiento, impidiendo con ello la posterior activación del canal de comunicación. Esto si bien puede resultar un retroceso, debe considerar que por restricciones de costos los elementos utilizados son no necesariamente los óptimos, de modo que si se avanza en el desarrollo, se debiese encontrar un sistema que reconozca ante estos

niveles de ruido o el sistema capaz de reducir el ruido a un nivel tal que logre ser reconocido posteriormente. A pesar de existir las soluciones con micrófonos de cuello o de transmisión por hueso, que permiten mejorar la comunicación en ambientes ruidosos, no resultan superiores al sistema realizado, desde el punto de vista de la usabilidad, dado que son un accesorio nuevo e incómodo para los operarios, a diferencia de la mascarilla con micrófono que sigue estando integrada al implemento de seguridad.

La activación del canal entrega como principal resultado, que es posible realizar el proceso sin grandes complicaciones, además un circuito como el diseñado en este trabajo entrega la suficiente independencia como para ser integrado a cualquier dispositivo que consiga buenos niveles de reconocimiento de voz, pues se establece como un sistema independiente. Al utilizar como dispositivo de activación un relé también se abre la solución a que no sólo se pueda activar el canal de una radio sino que cualquier otro sistema que se encuentre integrado a la comunicación, ya sea una luz al interior de una mina o la apertura de una puerta dado que él por si mismo permite aislar circuitos que puedan trabajar a voltajes distintos o con regímenes de corriente diferentes.

El prototipo final del sistema se puede clasificar como "prueba de concepto", sin embargo logra el objetivo principal, de desarrollar el sistema de comunicación a través de la mascarilla, permitiendo activar y desactivar el canal de la radio sin la necesidad de quitársela o utilizar un botón "push-to-talk". Además entrega una solución, integrada y robusta de bastante bajo costo, con claras posibilidades de ser comercializada. Además partes del desarrollo podrían a futuro ser patentadas, dado que invenciones en esta área no son comunes.

A modo general es claro que la construcción de un dispositivo capaz de mejorar tanto las características de protección como las de comunicación es posible, sin embargo es necesario mejorar algunos de los sistemas que en este trabajo se realizaron, por lo que a futuro se propone el perfeccionamiento de la etapa de reconocimiento de voz, ya sea utilizando algún sistema superior al Easy Veear como el SmartVR que permita además el reconocimiento ante condiciones mas adversas o aumentando los niveles de reducción de ruido. También se vuelve deseable para continuar con el trabajo, obtener mayores conocimientos del sistema de comunicación al que se acoplará el dispositivo, esto para evitar problemas como los del filtro, en que por desconocimiento se quema una de las etapas. Importante es a futuro considerar otros tipos de filtros ya sea variando la frecuencia a la que afectan o bien cambiando los análogos por digitales.

## 6. Bibliografía

- [1] M. Villayandre Llamazares, Fonética y Fonología. Tema 2. Fonética Articulatoria, Universidad de León, 2009.
- [2] S. F. G. R. García-Tapia Urrutia, Fisiología de la fonación, Universidad de Navarra, 2006.
- [3] E. G. P. Ralph Jones, Inteligibilidad del habla, Meyer Sound, 2009.
- [4] S. Alice H, Enciclopedia de salud y seguridad en el trabajo, Madrid: Ministerio de Trabajo y Asuntos Sociales, Subdirección General de Publicaciones, 1998.
- [5] I. N. d. S. e. H. e. e. Trabajo, El ruido en el ambiente laboral, Murcia: Comunidad Autónoma de la Región de Murcia, 1999.
- [6] D. d. F. U. d. Córdoba, Escalas de ponderación, Córdoba: Universidad de Córdoba, 2011.
- [7] S. f. C. L. Ron Kurtus, Active Noise Cancellation, 2011.
- [8] M. F. J.I. Escudero, Filtros, 2008.
- [9] J. A.Sanz, Diseño de filtros activos, Zaragoza: Universidad de Zaragoza.
- [10] E. V. F. Casacuberta, Reconocimiento automático del habla, Barcelona, España: Marcombo, 1987.
- [11] Casacuberta Nolla, Francisco, Alfonso Galipienso, María Isabel, M. J. Castro Bleda, Marzal Varo, Andrés, Pérez Cortés, Juan Carlos, Sánchez, Andrés, Benedí Ruiz, José Miguel, Galiano Ronda, Isabel Remedios y Vidal Ruiz, Enrique, Articulatory-acoustic correlations in coarticulatory processes: a cross-language investigation (Proyecto ACCOR), Sociedad Española para el Procesamiento del Lenguaje Natural, 1991.
- [12] <http://www.tetra-association.com/about/page/12030>, 08-03-2012, 03:41.
- [13] <http://www.p25.com/>, 08-03-2012, 03:38.

## 7. Anexos

### 7.1. Código Arduino

```
#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"
EasyVR easyvr(port);

//Groups and Commands
enum Groups
{
  GROUP_0 = 0,
  GROUP_1 = 1,
};

enum Group0
{
  G0_UNO = 0,
};

enum Group1
{
  G1_ATENTO = 0,
  G1_CAMBIO = 1,
};

EasyVRBridge bridge;

int8_t group, idx;

void setup()
{
  // bridge mode?
  if (bridge.check())
  {
    cli();
    bridge.loop(0, 1, 12, 13);
  }
  // run normally
  Serial.begin(9600);
  port.begin(9600);

  if (!easyvr.detect())
  {
    Serial.println("EasyVR no detectada!");
    for (;;)
  }

  easyvr.setPinOutput(EasyVR::IO1, LOW);
```



```

Serial.println("EasyVR detectada!");
easyvr.setTimeout(10);
easyvr.setLanguage(4);

group = EasyVR::TRIGGER; //<-- start group (customize)
}

void action();

void loop()
{
  easyvr.setPinOutput(EasyVR::IO1, HIGH); // LED on (listening)
  Serial.print("Diga un comando del grupo ");
  Serial.println(group);
  easyvr.recognizeCommand(group);
  do
  {
    // can do some processing while waiting for a spoken command
  }
  while (!easyvr.hasFinished());

  easyvr.setPinOutput(EasyVR::IO1, LOW); // LED off

  idx = easyvr.getWord();
  if (idx >= 0)
  {
    // built-in trigger (ROBOT)
    // group = GROUP_X; <-- jump to another group X
    return;
  }
  idx = easyvr.getCommand();
  if (idx >= 0)
  {
    // print debug message
    uint8_t train = 0;
    char name[32];
    Serial.print("Comando: ");
    Serial.print(idx);
    if (easyvr.dumpCommand(group, idx, name, train))
    {
      Serial.print(" = ");
      Serial.println(name);
    }
    else
      Serial.println();
    easyvr.playSound(0, EasyVR::VOL_FULL);
    // perform some action
    action();
  }
  else // errors or timeout
  {
    if (easyvr.isTimeout())
      Serial.println("Tiempo agotado, intente nuevamente...");
    int16_t err = easyvr.getError();
    if (err >= 0)
    {
      Serial.print("Error ");
      Serial.println(err, HEX);
    }
  }
}

```

```

    }
}

void action()
{
    switch (group)
    {
        case GROUP_0:
            switch (idx)
            {
                case G0_UNO:
                    // write your action code here
                    group = GROUP_1;// <-- or jump to another group X for composite
commands
                    break;
            }
            break;
        case GROUP_1:
            switch (idx)
            {
                case G1_ATENTO:
                    // write your action code here
                    group = GROUP_1;// <-- or jump to another group X for composite
commands
                    digitalWrite(11, HIGH);
                    break;
                case G1_CAMBIO:
                    // write your action code here
                    group = GROUP_0;// <-- or jump to another group X for composite
commands
                    digitalWrite(11,LOW);
                    break;
            }
            break;
    }
}
}

```

## 7.2. Esquemático placa de desarrollo LMV1089

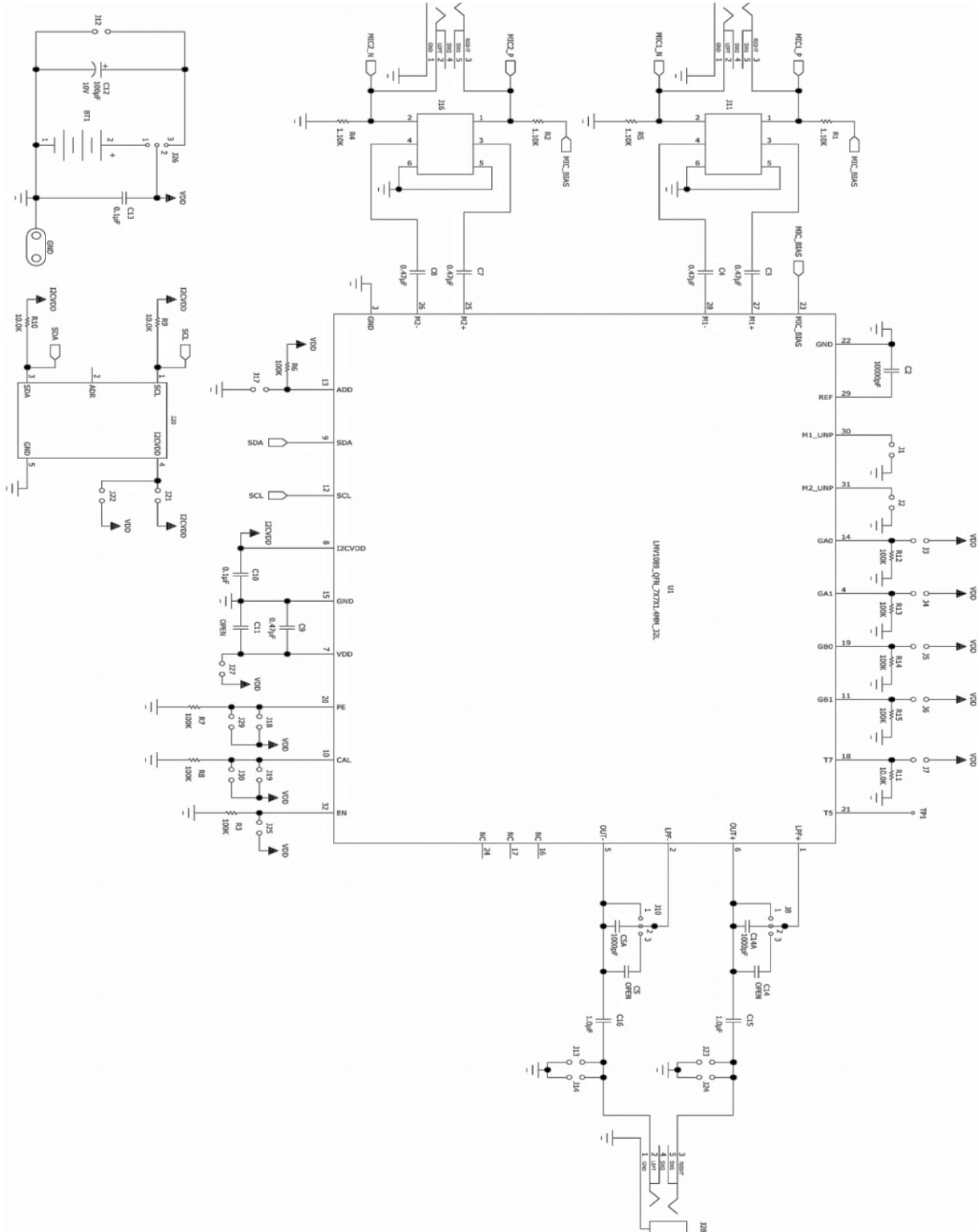


Ilustración 7.2-1: Esquemático cancelador