

UCUENCA

Facultad de Ingeniería

Carrera de Electrónica y Telecomunicaciones

**Diseño e implementación de un dispositivo
electrónico para monitoreo y apoyo al diagnóstico
de fonoaudiología en pacientes post COVID**

*Trabajo de titulación previo a la
obtención del título de Ingeniero en
Electrónica y Telecomunicaciones*

Autores:

Mauricio Esteban Guerrero Sarmiento

C.I: 0107084436

maurguer55@gmail.com

Paúl Vladimir Llerena Orellana

C.I: 0105431985

paul.llerena8@gmail.com

Director:

Ing. Luis Ismael Minchala Ávila, PhD

C.I: 0301453486

Cuenca, Ecuador

22-Septiembre-2022

Resumen

El presente trabajo de titulación desarrolla e implementa un dispositivo de ayuda al diagnóstico de fonoaudiología en pacientes post COVID-19 mediante la obtención de parámetros fundamentales de la voz a través de un sensor colocado en la superficie del cuello del individuo. La implementación del dispositivo consta de 2 partes, hardware y software.

La parte fundamental del hardware es el módulo ESP32. Este, alimenta al circuito y controla el manejo de los datos. La segunda parte más importante del circuito es el sensor utilizado, el cual es un acelerómetro BU-21771, el cual es robusto al ruido. Para que el sensor funcione adecuadamente, se diseñó una fuente de alimentación, etapas de pre-amplificación, amplificación, filtrado y almacenamiento de datos. La parte de software, consta de un servidor web con varias funciones, una de ellas, permite recopilar los datos obtenidos a través del sensor. Estos datos se guardan en el almacenamiento disponible, lo que permite descargar, borrar y visualizar los datos mediante el servidor.

Los datos obtenidos, se procesaron mediante el software Praat, con la finalidad de obtener los parámetros principales de la voz, y así poder realizar un análisis acústico a detalle. Los principales parámetros que se obtienen son: frecuencia fundamental, intensidad, tiempo de fonación, *jitter*, *shimmer* y HNR. Para complementar el análisis en Praat, existe la posibilidad de analizar cada uno de los parámetros de forma visual con las imágenes acústicas (oscilograma y espectrograma).

El dispositivo final se probó en 21 individuos post COVID-19, 12 hombres y 9 mujeres. Pruebas iniciales comparando los resultados del dispositivo con un micrófono profesional demostraron que el *jitter*, *shimmer* y HNR fueron los parámetros que más diferencia presentaban. Las siguientes pruebas demostraron la robustez del acelerómetro frente al ruido con resultados similares a la primera prueba. Finalmente, se realizó una comparación de resultados de 10 de 21 individuos diagnosticados por fonoaudiólogos con los datos obtenidos por el dispositivo. En conclusión, el prototipo desarrollado presenta un cierto margen de error con respecto a un diagnóstico profesional pero cumple con su objetivo de ayudar y complementar al fonoaudiólogo o especialista de la voz, con el diagnóstico mediante el análisis vocal.

Palabras clave: Fonoaudiología. Parámetros. ESP32. Acelerómetro. Praat. Frecuencia fundamental. Intensidad. Tiempo de fonación. Jitter. Shimmer. HNR.

Abstract

This degree work develops and implements a device to aid in the diagnosis of speech therapy in post-COVID-19 patients by obtaining fundamental parameters of the voice through a sensor placed on the surface of the individual's neck. The implementation of the device consists of 2 parts, hardware and software.

The fundamental part of the hardware is the ESP32 module. It powers the circuit and controls the data handling. The second most important part of the circuit is the sensor used, which is an accelerometer BU-21771, which is robust to noise. For the sensor to work properly, a power supply, pre-amplification, amplification, filtering and data storage stages were designed. The software part, consists of a web server with several functions, one of them, allows to collect the data obtained through the sensor. These data are saved in the available storage, which allows downloading, deleting and visualizing the data through the server.

The data obtained were processed by Praat software, in order to obtain the main parameters of the voice, and thus be able to perform a detailed acoustic analysis. The main parameters obtained are: fundamental frequency, intensity, phonation time, jitter, shimmer and HNR. To complement the analysis in Praat, there is the possibility of analyzing each of the parameters visually with acoustic images (oscillogram and spectrogram).

The final device was tested on 21 post-COVID-19 individuals, 12 males and 9 females. Initial tests comparing the results of the device with a professional microphone showed that jitter, shimmer and HNR were the parameters that showed the most difference. Subsequent tests demonstrated the robustness of the accelerometer to noise with results similar to the first test. Finally, a comparison of the results of 10 of 21 individuals diagnosed by speech therapists with the data obtained by the device was performed. In conclusion, the prototype developed presents a certain margin of error with respect to a professional diagnosis, but it fulfills its objective of helping and complementing the speech therapist or voice specialist with the diagnosis through vocal analysis.

Keywords: Speech therapy. Parameters. ESP32. Accelerometer. Praat. Fundamental frequency. Intensity. Phonation time. Jitter. Shimmer. HNR.

Índice general

Resumen	1
Abstract	2
Índice general	3
Índice de figuras	7
Índice de tablas	10
Cláusula de licencia y autorización para publicación	11
Cláusula de licencia y autorización para publicación	12
Cláusula de propiedad intelectual	13
Cláusula de propiedad intelectual	14
Dedicatoria	15
Dedicatoria	16
Agradecimientos	17
Agradecimientos	18
Abreviaciones y acrónimos	19
1. Introducción	20
1.1. Antecedentes	20
1.2. Identificación del problema	21
1.3. Objetivos	23
1.3.1. Objetivo general	23
1.3.2. Objetivos específicos	23
1.4. Revisión del estado del arte	23
1.5. Contribuciones	25
2. Fundamentos teóricos	26
2.1. Análisis de la producción vocal	26
2.2. Patologías que causan problemas en la voz	31

2.2.1.	Patologías relacionadas con el COVID-19	32
2.2.2.	Intubación endotraqueal	33
2.3.	Análisis acústico de la voz	35
2.3.1.	Principales parámetros de la voz	35
2.3.1.1.	Frecuencia fundamental	35
2.3.1.2.	Nivel de Intensidad acústica	36
2.3.1.3.	Tiempo de fonación	36
2.3.1.4.	Timbre	37
2.3.1.5.	Ataque y filatura	37
2.3.2.	Parámetros secundarios	37
2.3.2.1.	Shimmer	37
2.3.2.2.	Jitter	38
2.3.2.3.	HNR (Harmonic-to-Noise Ratio)	38
2.3.2.4.	Energía	39
2.4.	Audición	40
2.4.1.	Audición humana y niveles de intensidad del sonido	40
2.4.2.	Audición y tono	42
2.5.	Sonido	43
2.5.1.	Ondas sonoras	43
2.5.2.	Representación gráfica del sonido	44
2.5.3.	Caracterización del ruido	46
3.	Diseño e implementación del dispositivo electrónico para monitoreo y apoyo al diagnóstico de fonoaudiología en pacientes post COVID-19	49
3.1.	Descripción del desarrollo	49
3.2.	Componentes y materiales utilizados	49
3.2.1.	ESP32	49
3.2.2.	Módulo micrófono con MAX4466	51
3.2.3.	Módulo lector de tarjeta micro SD	52
3.2.4.	Acelerómetro BU-21771	52
3.3.	Desarrollo del dispositivo electrónico	53
3.3.1.	Fuente de alimentación de 1.5V	53
3.3.2.	Circuito de pre-amplificación	54
3.3.3.	Conexión de los demás componentes y circuito final	55
3.3.4.	Diseño PCB del circuito final	56
3.4.	Script para ESP32	58
3.5.	Obtención de parámetros acústicos	59
3.5.1.	Praat	59
3.5.2.	Obtención de parámetros a través de Praat	60
3.5.2.1.	Frecuencia fundamental	60

3.5.2.2.	Intensidad	61
3.5.2.3.	Tiempo de fonación	61
3.5.2.4.	Jitter	63
3.5.2.5.	Shimmer	64
3.5.2.6.	HNR (Harmonic-to-Noise Ratio)	64
3.5.3.	Obtención de imágenes acústicas a través de Praat	65
4.	Resultados	67
4.1.	Proceso de adquisición de la muestra	67
4.1.1.	Metodología de toma de muestras por parte de especialistas de la voz	67
4.1.2.	Reactivos utilizados	67
4.2.	Prototipo del dispositivo final	69
4.3.	Consumo energético del circuito	71
4.4.	Metodología de toma de muestras con el dispositivo electrónico	72
4.4.1.	Posicionamiento del acelerómetro	72
4.4.2.	Protocolo de grabación	73
4.5.	Proceso de grabación y almacenamiento de los archivos de audio a través de la página web	74
4.5.1.	Grabar	75
4.5.2.	Descargar	76
4.5.3.	Borrar	77
4.5.4.	Directorio	77
4.5.5.	Posibles errores	78
4.6.	Adquisición de los parámetros a través de Praat	79
4.7.	Obtención de los parámetros acústicos	81
4.7.1.	Análisis estadístico	83
4.8.	Análisis y evaluación de los parámetros acústicos	91
4.8.1.	Frecuencia Fundamental (F_0)	91
4.8.2.	Intensidad	92
4.8.3.	Tiempo de fonación	92
4.8.4.	Jitter local	93
4.8.5.	Shimmer local	94
4.8.6.	HNR	96
4.9.	Comparación de los parámetros adquiridos con el dispositivo y el informe fonaudiológico	99
4.10.	Adquisición del oscilograma y espectrograma a través de Praat	100
4.11.	Análisis y evaluación de oscilograma y espectrograma	101
4.11.1.	Evaluación del oscilograma y espectrograma de un individuo con diagnóstico presuntivo (disfónico)	101

4.11.2. Evaluación del oscilograma y espectrograma de un individuo con diagnóstico presuntivo sano (eufonía)	102
5. Conclusiones y recomendaciones	104
5.1. Conclusiones	104
5.2. Recomendaciones	106
5.3. Trabajos futuros	107
A. ANEXOS	108
A. Scripts de servidor web y adquisición de datos desde el ESP32	108
A.1. Script principal del programa .ino	108
A.2. Script .h CSS para el diseño de la página web	108
A.3. Script .cpp para configuración del modo I2S	108
A.4. Script .h para configuración del modo I2S	108
A.5. Script .h para configuración de la red	108
A.6. Script .h para configuración de pin CS del lector SD	108
A.7. Script .cpp para configuración del archivo .wav de audio	108
A.8. Script .h para configuración del archivo .wav de audio	109
B. Script de Praat para obtención de parámetros	109
C. Script de Python para análisis de parámetros	117
D. Modelo de consentimiento informado	120
E. Modelo de historia fonoaudiológica	121
F. Informe fonoaudiológico general	126
G. Datasheet	132
G.1. Acelerómetro BU-21771	132
H. Diseños	134
H.1. Caja para circuito	134
H.2. Sticker 1	134
H.3. Sticker 2	135
Bibliografía	136

Índice de figuras

2.1. Esquema representativo del aparato fonador. Se muestran las principales estructuras anatómicas involucradas en el proceso de la fonación. Tomado de [1].	26
2.2. Separación (abducción) de los pliegues vocales. Tomado de [2].	27
2.3. Vista lateral. Tomado de [2].	28
2.4. Vista lateral derecha. Tomado de [2].	28
2.5. Vista superior. Tomado de [2].	29
2.6. Dibujo esquemático de la dirección del flujo de aire a través de la glotis y en contacto con la columna de aire supraglótica. Tomado de [3].	30
2.7. Producción de la voz. Adaptado de [4].	31
2.8. Tubo endotraqueal de alto volumen y baja presión. (1) Punta biselada redondeada, (2) guía de profundidad, (3) tubo de inflado, (4) balón piloto con válvula, (5) manguito endotraqueal. Tomado de [5].	33
2.9. Anatomía del oído humano. Tomado de [6].	40
2.10. Relación entre volumen en fonios y nivel de intensidad (en decibelios) e intensidad (en vatios por metro cuadrado) para personas con audición normal. Tomado de [7].	43
2.11. Onda sonora emana de una fuente, como un diapasón, que vibra a una frecuencia f . Se propaga a una velocidad v y tiene una longitud de onda λ . Tomado de [8].	44
2.12. Espectrograma de banda ancha, sonido: vocal “a”. Obtenido en Praat. Tomado de [9].	45
2.13. Espectrograma de banda estrecha, sonido: vocal “a”. Obtenido en Praat. Tomado de [9].	45
2.14. Oscilograma y Espectrograma. Tomado de [10].	46
2.15. Ruido continuo. Tomado de [11].	47
2.16. Ruido continuo intermitente. Tomado de [11].	47
2.17. Ruido impulsivo. Tomado de [11].	48
2.18. Ruido fluctuante. Tomado de [11].	48
3.1. Diagrama de bloques. Autoría propia.	49
3.2. ESP32. Tomado de [12].	51
3.3. MAX4466. Tomado de [13].	51
3.4. Lector micro SD. Tomado de [14].	52
3.5. Acelerómetro BU-21771. Tomado de [15].	53
3.6. Circuito de fuente de alimentación de 1.5V. Autoría propia.	54

3.7. Circuito de pre-amplificación. Autoría propia.	55
3.8. Circuito final. Autoría propia.	56
3.9. Circuito final. Autoría propia.	57
3.10. Circuito final en PCB. Autoría propia.	57
3.11. Circuito final en 3D. Autoría propia.	58
3.12. Interfaz del servidor web. Autoría propia.	58
3.13. Oscilograma y espectrograma en Praat. Autoría propia.	66
4.1. Prototipo del dispositivo final - vista superior. Autoría propia.	69
4.2. Prototipo del dispositivo final - vista frontal. Autoría propia.	70
4.3. Prototipo del dispositivo final - vista trasera. Autoría propia.	70
4.4. Placas impresas sin componentes soldados. Autoría propia.	71
4.5. Conjunto de puntos donde es posible colocar el acelerómetro. Tomado de [16].	72
4.6. Amplitud de cada punto. Tomado de [16].	73
4.7. Redes WiFi disponibles, incluida la del ESP32. Autoría propia.	74
4.8. Conexión a una red WiFi mediante el ESP32. Autoría propia.	75
4.9. Menú para grabar el archivo. Autoría propia.	76
4.10. Mensaje que indica que el archivo se ha grabado correctamente. Autoría propia.	76
4.11. Menú para descargar el archivo. Autoría propia. Autoría propia.	76
4.12. Descarga del archivo. Autoría propia.	77
4.13. Menú para borrar el archivo. Autoría propia.	77
4.14. Mensaje que indica que el archivo se ha borrado correctamente. Autoría propia.	77
4.15. Lista de archivos disponibles en la tarjeta micro SD. Autoría propia.	78
4.16. Error al no insertar la tarjeta micro SD. Autoría propia.	78
4.17. Ventana de 10 segundos y sub-ventana de 4 segundos en Praat. Autoría propia.	79
4.18. Prueba en cabina insonorizada simultánea con micrófono profesional y dispositivo. Autoría propia.	81
4.19. Análisis estadístico de la frecuencia fundamental. Autoría propia.	85
4.20. Análisis estadístico de la intensidad. Autoría propia.	86
4.21. Análisis estadístico del tiempo de fonación. Autoría propia.	86
4.22. Análisis estadístico del jitter local. Autoría propia.	87
4.23. Análisis estadístico del shimmer local. Autoría propia.	87
4.24. Análisis estadístico del HNR. Autoría propia.	88
4.25. Prueba fuera de la cabina insonorizada del dispositivo. Autoría propia.	88
4.26. Resultado de análisis vocal de un individuo. Autoría propia.	91
4.27. Adquisición del oscilograma completo. Autoría propia.	100

4.28. Adquisición del oscilograma y espectrograma en un rango de 10 segundos. Autoría propia.	101
4.29. Adquisición del oscilograma y espectrograma en un paciente con diagnóstico presuntivo disfónico. Autoría propia.	102
4.30. Adquisición del oscilograma y espectrograma en un paciente con diagnóstico presuntivo sano (eufonía). Autoría propia.	103

Índice de tablas

1.1. Variables de dosimetría de voz de los principales dispositivos comercializados	25
2.1. Niveles de intensidad del sonido e intensidades	41
3.1. Características del ESP32.	50
3.2. Pines de conexión del lector microSD al ESP32.	56
3.3. Parámetros obtenidos en el software Praat.	65
4.1. Reactivos utilizados por parte de especialistas de la voz.	68
4.2. Consumo energético de los elementos del circuito.	71
4.3. Resultados de parámetros obtenidos mediante el micrófono (Beta4-080) en prueba simultánea.	82
4.4. Resultados de parámetros obtenidos mediante el dispositivo (acelerómetro) en prueba simultánea.	82
4.5. Estadísticas del dispositivo para hombres.	83
4.6. Estadísticas del dispositivo para mujeres.	84
4.7. Estadísticas del micrófono para hombres.	84
4.8. Estadísticas del micrófono para mujeres.	85
4.9. Resultados de parámetros obtenidos mediante el dispositivo en prueba no simultánea.	89
4.10. Parámetros del análisis vocal de individuos	90
4.11. Evaluación del parámetro tiempo de fonación en cada individuo.	93
4.12. Evaluación del parámetro jitter local en cada individuo.	94
4.13. Evaluación del parámetro shimmer local en cada individuo.	95
4.14. Evaluación del parámetro jitter local en cada individuo.	96
4.15. Resumen de cada individuo y el parámetro medido.	98

Cláusula de licencia y autorización para publicación en el repositorio institucional

Mauricio Esteban Guerrero Sarmiento en calidad de autor y titular de los derechos morales y patrimoniales del trabajo de titulación "Diseño e implementación de un dispositivo electrónico para monitoreo y apoyo al diagnóstico de fonoaudiología en pacientes post COVID", de conformidad con el Art. 114 del CÓDIGO ORGÁNICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÓN reconozco a favor de la Universidad de Cuenca una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Asimismo, autorizo a la Universidad de Cuenca para que realice la publicación de este trabajo de titulación en el repositorio institucional, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Cuenca, 22/09/2022



Mauricio Esteban Guerrero Sarmiento

C.I: 0107084436

Cláusula de licencia y autorización para publicación en el repositorio institucional

Paúl Vladimir Llerena Orellana en calidad de autor y titular de los derechos morales y patrimoniales del trabajo de titulación "Diseño e implementación de un dispositivo electrónico para monitoreo y apoyo al diagnóstico de fonoaudiología en pacientes post COVID", de conformidad con el Art. 114 del CÓDIGO ORGÁNICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÓN reconozco a favor de la Universidad de Cuenca una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Asimismo, autorizo a la Universidad de Cuenca para que realice la publicación de este trabajo de titulación en el repositorio institucional, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Cuenca, 22/09/2022



Paúl Vladimir Llerena Orellana

C.I: 0105431985

Cláusula de propiedad intelectual

Mauricio Esteban Guerrero Sarmiento autor del trabajo de titulación "Diseño e implementación de un dispositivo electrónico para monitoreo y apoyo al diagnóstico de fonoaudiología en pacientes post COVID", certifico que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autor.

Cuenca, 22/09/2022



Mauricio Esteban Guerrero Sarmiento

C.I: 0107084436

Cláusula de propiedad intelectual

Paúl Vladimir Llerena Orellana autor del trabajo de titulación Diseño e implementación de un dispositivo electrónico para monitoreo y apoyo al diagnóstico de fonoaudiología en pacientes post COVID”, certifico que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autor.

Cuenca, 22/09/2022



Paúl Vladimir Llerena Orellana

C.I: 0105431985

Dedicatoria

A mis abuelitos en el cielo. Hubiese querido que sean testigos de la culminación de este proceso y el cumplimiento de una meta más y poder celebrar juntos este logro. Los extraño demasiado.

Dedico el esfuerzo, la dedicación y la constancia puesta en este proyecto, a mi padre, que me brindó todo su apoyo incondicional y creyó en mí, cuando ni siquiera yo lo hacía, a ti te debo la lección más grande que he aprendido en la vida: “nunca rendirse, solo resistir”. A pesar de muchas caídas y obstáculos a lo largo de todo el proceso académico, siempre confiaste en mí y fuiste inspiración para no dejar de lado un sueño y continuar en la lucha constante para alcanzar la meta. Espero algún día compensar de alguna manera todo el esfuerzo y confianza que has puesto en mí y no decepcionarte nunca. Si algún día soy padre, espero ser por lo menos la mitad de lo que tú has sido conmigo.

Finalmente, dedico este proyecto a mi abuelita, que con cada una de sus palabras de aliento ha hecho de mí una persona fuerte y con su cariño infinito hace que todo sea un poco más fácil.

Mauricio Esteban Guerrero Sarmiento

Dedicatoria

A mis abuelitos que los extraño mucho, Mami Zoila, Mami Luz y Papi Lucho. Siempre los tendré presentes en mi corazón, gracias por cuidarme desde donde quiera que estén. Algún día volveremos a reír juntos.

A mis padres, por enseñarme lo que es el trabajo duro y la responsabilidad. Espero algún día poder devolverles todo lo que han hecho por mí.

Paúl Vladimir Llerena Orellana

Agradecimientos

Agradezco en primer lugar a Dios, por otorgarme la sabiduría para culminar este proyecto y con ello, la culminación de esta etapa académica.

A mi padre, por el constante apoyo en todos los aspectos del proyecto y por cada uno de sus valiosos consejos que me ayudaron a forjar la resiliencia necesaria para culminar esta etapa, agradezco a mi madre, por su inmensa paciencia y por la motivación brindada a lo largo de este recorrido. A mis hermanos, por el apoyo incondicional para culminar y cumplir una meta más en mi formación académica.

A mi abuelita Enma, por la paciencia, los consejos y sobre todo el inmenso afecto que ha tenido, no solo en esta etapa, si no a lo largo de toda mi existencia. Gracias, por ayudarme a no rendirme ante los obstáculos de la vida y formar a la persona que soy hoy en día.

A Paúl Llerena, amigo y compañero de este proyecto, por el compromiso, entrega y dedicación que demostró a lo largo del desarrollo del proyecto, gracias porque valió la pena cada error, cada frustración, cada desmotivación, cada retada, pero sobre todo cada aprendizaje y enseñanza que nos dejó el proyecto.

Al Ing. Ismael Minchala, por ser nuestro guía en todo el proceso de desarrollo del proyecto, de igual manera, expreso mi gratitud al Ing. Villie Morocho y a la Ing. Mabel Méndez por estar pendientes del avance en el proyecto y por los consejos brindados para la culminación satisfactoria del mismo.

A Estefanía Mejía, partícipe del proyecto, por valorar cada aspecto del proyecto y el compromiso entregado al mismo, gracias por el profesionalismo entregado en cada una de las pruebas realizadas y por el aprendizaje adquirido a lo largo del desarrollo del mismo. A cada una de las personas que formaron parte de las pruebas finales con el dispositivo, por otorgar el tiempo y disposición para concluir el proyecto.

A mis amigos y familiares, por todo el apoyo brindado y por formar parte esencial de la carga emocional que implica todo proceso académico.

“Una alegría compartida es duplicada, pero una tristeza compartida es dividida”.

Mauricio Esteban Guerrero Sarmiento

Agradecimientos

A mis padres, Milton y Janet por siempre darme ánimo y palabras de aliento a lo largo de toda la carrera. Por siempre estar para mí y procurar que nada me falte. Sin ustedes nada de esto habría sido posible, los quiero mucho.

A mis hermanos Danny y Paula por acompañarme en mis primeros años de universidad y a mis tíos Edit y José Luis, por ayudarme con cualquier cosa que necesitaba.

A mi amigo y compañero Mauricio, por esa amistad sincera todos estos años de carrera y por batallar junto a mí en la culminación de este trabajo de titulación. Seguro estoy de que los sentimientos de estima perdurarán con el tiempo.

A Evelyn, por ese inmenso amor y apoyo que me ha brindado en este tiempo. Gracias por siempre creer en mí, por siempre estar ahí y animarme a dar todo de mí. Que la vida me regale muchos momentos más a tu lado, te amo.

A Estefanía Mejía, que a pesar de no estar vinculada con nuestro trabajo y de no conocernos bien, nos ayudó tanto de manera desinteresada, con un gran profesionalismo y entrega. De igual manera a todas las personas que prestaron su tiempo para realizar las pruebas respectivas, muchas gracias. Al ingeniero Raúl Pesántez, por sus consejos y preocupación hacia nosotros y nuestro proyecto. A Panchito, por abrirnos siempre las puertas del laboratorio de la facultad de Ingeniería y depositar su confianza en nosotros.

Al ingeniero Ismael Minchala, por sus enseñanzas como docente en la universidad y por su guía como nuestro tutor. A la ingeniera Mabel Méndez, por guiarnos en los inicios del proyecto y al ingeniero Villie Morocho, por brindarnos apoyo en todo momento.

A todos los compañeros y amigos que de alguna u otra manera nos supieron apoyar y ayudar, en especial a Pablo y Nancy, cuyo apoyo al final del proyecto fue determinante.

A Lucas y Nico, por demostrarme que el amor es más que palabras.

“A veces sentimos que lo que hacemos es tan sólo una gota en el mar; pero el mar sería menos si le faltara una gota”.

Paúl Vladimir Llerena Orellana

Abreviaciones y acrónimos

AP Access Point. [74](#)

APM Ambulatory Phonation Monitor. [24](#), [25](#)

F0 frecuencia fundamental. [59](#)

GUI Graphical User Interface. [100](#)

HNR Harmonic-to-Noise Ratio. [38](#), [39](#), [64](#), [65](#), [87](#), [88](#), [91](#), [99](#), [104](#), [105](#)

LCD Liquid-Crystal Display. [49](#)

OMS Organización mundial de la Salud. [21](#), [31](#)

ORL Centro de Otorrinolaringología. [99](#)

PCB Printed Circuit Board. [56](#), [57](#), [70](#), [106](#)

PT tiempo de fonación. [24](#)

SARS-CoV-2 severe acute respiratory syndrome coronavirus 2. [21](#), [23](#)

SDRA Síndrome de Dificultad Respiratoria Aguda. [22](#)

SEORL Sociedad Española de Otorrinolaringología. [31](#)

SoC Security Operations Center. [49](#)

SPI Serial Peripheral Interface. [56](#)

SPL nivel de presión sonora. [24](#)

SSID Service Set Identifier. [74](#)

ST tiempo de habla. [24](#)

TMF Tiempo Máximo de Fonación. [99](#)

VMI ventilación mecánica invasiva. [22](#), [24](#), [32](#)

1. Introducción

La pandemia producida por el COVID-19, ha generado una gran cantidad de fallecimientos, debido a que ataca principalmente al sistema respiratorio, dando lugar a una variedad de secuelas, entre ellas las provocadas por intubamiento lo que puede causar la pérdida parcial o prolongada de la voz. Según [17], la principal afección de la voz se debe al proceso de intubación, debido a que se genera un proceso de ventilación mecánica invasiva entre la tráquea y las cuerdas vocales, generando así, cierto grado de disfonía. El daño laríngeo a corto plazo incluye edema, hemorragia y formación de granulomas. Aunque un daño más severo puede producir afonía y un daño menos severo, solo ronquera. En la aplicación de la telemedicina para fonoaudiología es necesario el uso de dispositivos que ayuden a percibir de mejor forma la realidad, variando la voz del paciente. Así mismo, resulta necesario hallar ciertas patologías laríngeas que ayuden al especialista a detectar cualquier anomalía de la voz.

En el momento de valorar los trastornos fonatorios, la cuantificación y la objetivación de diversos parámetros acústicos de manera fiable, precisa y exacta, constituyen un método de análisis sustancial para los especialistas de la voz. La cuantificación y caracterización acústica precisa de la voz, forma parte fundamental tanto desde el punto de vista de diagnóstico por parte de especialistas de la voz como para la validación o constatación de resultados de los diferentes tratamientos. Además, de las ventajas mencionadas, la caracterización acústica podría proporcionar información de patrones de riesgo al momento de prevenir anomalías o trastornos de la voz que podrían producirse en el futuro [18, 19].

1.1. Antecedentes

Desde el momento en que nacemos, los seres humanos tenemos la necesidad de comunicarnos y uno de los medios por el cual lo hacemos es mediante la voz, cuyo efecto proviene de proceso complejo, en el que interactúan diversos sistemas corporales, entre ellos el respiratorio, fonatorio, neuromuscular, auditivo, endocrino, articulatorio y resonancial, cada uno de estos procesos trabajando en conjunto dan como resultado la emisión de sonidos que posteriormente serán de utilidad al ser humano para diversos fines, principalmente para fines comunicativos, para transmisión de ideas, pensamientos, emociones, etc. La voz adopta diferentes características y pueden ser clasificadas de manera general en aspectos locutivos y no locutivos. Los primeros hacen referencia a como se produce la voz y pueden ser estudiados o analizados de manera auditiva, un

ejemplo es el parámetro de la frecuencia fundamental, esta tiene una variación relevante según el sexo de cada persona, siendo por lo general más aguda en mujeres que en hombres. Los segundos están relacionados de manera indirecta a su producción, por ejemplo, la respiración, la postura o la tonicidad [20, 21].

La voz es el resultado de la combinación de cualidades sonoras como el tono, la intensidad y el timbre que se desarrolla a través de mecanismos fisiológicos. El sistema nervioso central y periférico son de los más fundamentales en el proceso de la articulación de las palabras, refiriéndose a los componentes estructurales y funcionales de la laringe. Adicionalmente, el sistema respiratorio complementa a los dos ya mencionados, junto con el sistema auditivo que regula la intensidad de la producción vocal [22].

El desarrollo de dispositivos que facilitan la valoración y análisis acústico, han permitido ampliar en gran medida la comprensión de la función vocal. El principal objetivo de estos dispositivos es establecer un diagnóstico riguroso de la disfonía, y con esto implementar un plan para tratar específicamente cada problema. Según teoría de la historia clínica, toda evaluación comienza frecuentemente con la anamnesis; específicamente en este caso un interrogatorio relacionado con aspectos que pueden repercutir en el fenómeno vocal. De esta manera el experto profesional puede valorar en primera instancia los componentes de la voz y el habla de una persona. La primera evaluación establece un diagnóstico de existencia o no de disfonía [23].

1.2. Identificación del problema

Actualmente la pandemia del severe acute respiratory syndrome coronavirus 2s (SARS-CoV-2s), llamada COVID-19, ha originado una severa crisis económica, social y de salud, sin precedentes. Todo comenzó en China, en la ciudad de Wuhan, a finales diciembre de 2019. En este lugar se anunció a un conjunto de 27 personas con una neumonía de causas desconocidas, 7 de estas personas en estado grave. El primer caso registrado fue el 8 de diciembre de 2019, para el 24 de enero de 2020 los casos aumentaron a 835. Con el paso del tiempo se fue extendiendo a otras regiones de China e incluso a otros países. Por consiguiente la Organización mundial de la Salud (OMS), declaró en marzo de 2020 una nueva pandemia mundial [24].

Los síntomas de la enfermedad por COVID-19 varían de leves a severos (en ciertos casos incluso, puede no presentarse síntoma alguno, esto se conoce como transmisión asintomática) y pueden aparecer entre 2 y 14 días después de contraer el virus. El período entre el cual se contrae el virus y se presenta algún síntoma se conoce como período de incubación. Aunque la mayoría de personas pueden experimentar síntomas leves o moderados, existe la posibilidad de alguna complicación que puede causar la

muerte [25].

La afección más frecuente y severa en pacientes con COVID-19, es el Síndrome de Dificultad Respiratoria Agudas (SDRAs), que precisa de terapias de ventilación y oxígeno. Según [26], sugiere que entre el 9,8 y el 15,2% de los pacientes requirieron ventilación mecánica invasivas (VMIs) en el 2019 en China. Se ha demostrado que los pacientes con COVID-19 tenían una duración media de la ventilación de 17 días y una alta frecuencia de reintubación.

Los efectos perjudiciales sobre la laringe, debido a la intubación prolongada incluyen alteración de la voz, la deglución y las vías respiratorias, como parálisis de las cuerdas vocales, disfagia y estenosis laringotraqueal, que pueden afectar la morbilidad y la mortalidad del paciente [27]. Además, la mayoría de los virus en el tracto respiratorio superior pueden asociarse con laringofaringitis y se ventilan durante este pico de infección, lo que pone al paciente en mayor riesgo de estenosis glótica, subglótica o después del período de intubación [28, 29].

Según [30], para la producción de una fonación normal, es necesario el equilibrio entre las fuerzas aerodinámicas y las fuerzas mioelásticas de la laringe, no obstante, una disminución gradual de la función pulmonar puede dar lugar a una disminución del apoyo aéreo para la producción de la voz, también se puede dar el caso en que algunos pacientes compensen su escaso apoyo aéreo con la contracción laríngea durante la fonación, produciendo así una calidad vocal tensa.

Dentro del marco del COVID-19, el fonoaudiólogo es uno de los profesionales importantes del área de la salud, principalmente en ámbito de rehabilitación de pacientes sometidos a respiradores artificiales o traqueostomías, que pueden desarrollar alteraciones vocales, denominadas disfonías. Cabe señalar que el tratamiento de las afecciones o alteraciones vocales causados por problemas de salud o factores de ronquera posterior a la extubación ya existía antes de la pandemia, siendo fundamental en los casos post COVID-19 [31].

Como medida inicial para efectuar un diagnóstico, según [32], se propone la recopilación de muestras para realizar un análisis vocal. Esta muestra es una grabación digital en formato *wav* con un muestreo de 44,1 kHz. Sin embargo, es necesario un micrófono de alta calidad. El micrófono se recomienda utilizarlo dentro de una cabina insonorizada para evitar cualquier ruido exterior. El costo del micrófono junto con la cabina es alto. Otra opción para esto, se menciona en [33], que enumera 3 dispositivos disponibles para monitoreo de la voz, APM, VoxLog y VocaLog. El costo de los mismos al igual que antes, es elevado.

Se optó por desarrollar un dispositivo similar a los 3 mencionados antes, se prevee que el dispositivo sea de un costo accesible y robusto al ruido.

1.3. Objetivos

1.3.1. Objetivo general

Diseñar e implementar un dispositivo electrónico para el monitoreo y apoyo al diagnóstico de pacientes post COVID-19, que tengan la capacidad de medir y registrar: la frecuencia fundamental, la intensidad y el tiempo de fonación.

1.3.2. Objetivos específicos

El presente trabajo tiene los siguientes objetivos específicos:

- Medir y analizar los parámetros principales de la voz (frecuencia fundamental, intensidad y tiempo de fonación).
- Medir el consumo energético y posibilitar la monitorización de los parámetros de la voz durante un tiempo prudencial de forma ininterrumpida.
- Adquisición de la señales a través del microcontrolador para posteriormente realizar el procesamiento digital de la señal en una PC.
- Elaborar un resumen y banco de pruebas realizadas con individuos.

1.4. Revisión del estado del arte

La voz es una palabra que da la denominación al sonido que se produce en la laringe, que varía a través de las cavidades de resonancia. La voz es un acto voluntario, en la que interviene, el sistema nervioso central, la audición, los órganos de fonación y el estado en general [34]. Para examinar la voz existen 3 parámetros fundamentales: frecuencia fundamental, intensidad y tiempo de fonación. El tiempo de fonación y la frecuencia fundamental se pueden obtener directamente de la medición utilizando algún tipo de sensor, por otro lado la intensidad de la voz se puede estimar a través de la amplitud de las vibraciones que se registran en el dispositivo y un proceso de calibración [35–37].

Debido a varios factores, la voz puede verse afectada en distinta medida. Aunque el virus del SARS-CoV-2s, llamado COVID-19, se conoce principalmente como una infección viral que afecta al sistema respiratorio, generando signos respiratorios graves como dificultad respiratoria o incluso insuficiencia respiratoria, casi todos los sistemas u órganos del cuerpo pueden verse afectados [38]. Es decir, existen pacientes que corren el riesgo de sufrir problemas en el aparato fonador debido a la enfermedad en sí o como

consecuencia de la VMIs en casos graves de la enfermedad [39]. Según [40], en el inicio de la pandemia COVID-19, se utilizó la práctica de la evaluación médica a distancia. A pesar de ser ampliamente utilizado por los fonoaudiólogos, la falta de evaluación estática y dinámica de la laringe y el tracto vocal es un punto crítico en esta práctica.

En [41], se propone un estudio del desgaste vocal de un musulmán durante el Ramadán, específicamente en el período de oración. Utilizando un Ambulatory Phonation Monitors (APMs), que contiene un acelerómetro miniatura, dicho sensor es colocado directamente en el cuello, justo encima de la muesca esternal, para poder así realizar el monitoreo vocal. Con este estudio se concluye que el desgaste vocal severo provoca cambios en los parámetros tales como: la frecuencia fundamental, *jitter* y *shimmer*.

Referente al parámetro de la frecuencia fundamental de la voz, en [42], se realiza un estudio comparativo en donde se analiza el parámetro de la frecuencia fundamental obtenido mediante valoraciones acústicas, electroglotográficas y acelerómetros ubicados en el cuello en individuos con y sin anomalías en la voz. Los resultados de la extracción de frecuencia fundamental se compararon entre micrófono acústico, electrodos y acelerómetro para cada sujeto. Los valores de frecuencia fundamental estimados con el micrófono y el acelerómetro mostraron una correlación perfecta; sin embargo, los valores estimados por los electrodos difieren de ellos.

En [43], se desarrolla un prototipo para el monitoreo de los parámetros de la voz, se propone un dispositivo vestible, compacto y liviano con conexión inalámbrica. Se utiliza como sensor, un acelerómetro con efecto piezoeléctrico, dicho sensor es colocado en el cuello a la altura del inicio del esternón. Y un micrófono con la finalidad de detectar el nivel de presión sonoras (SPLs). Ese prototipo cuenta con un collar, en el que caben todos los dispositivos como acelerómetro, micrófono, batería y el dispositivo o módulo de conexión inalámbrica. Gracias a este estudio ha sido posible detectar diferentes parámetros vocales tales como: la frecuencia fundamental, el nivel de presión sonora y el tiempo de fonación.

Así mismo en [44], se hace referencia al dispositivo VocaLog como medio para indexar el volumen vocal a través de niveles calibrados de dB SPLs. Sin embargo, también es probable que se registre alguna actividad no fonatoria. El dispositivo proporciona una estimación razonable de tiempo de hablas (STs), pero no de tiempo de fonacións (PTs). Por otro lado en [45], se realizó un estudio del mismo dispositivo con similares resultados pero con la particularidad de que este dispositivo resulta un tanto incómodo para pacientes con la enfermedad de Parkinson.

Otro dispositivo es el VoxLog, según [33], básicamente tiene la misma finalidad de los dispositivos antes mencionados con ciertas similitudes. En cuanto a las variables de dosimetría de la voz que logra registrar, se podría decir que es bastante similar al dis-

positivo APMs. La disponibilidad comercial reciente de dispositivos para el monitoreo ambulatorio de la voz y la biorretroalimentación ofrece a los médicos e investigadores de la voz nuevas opciones para mejorar potencialmente la evaluación, el tratamiento y la investigación de los trastornos de la voz. En la tabla 3.1 se muestran las especificaciones básicas de los tres dispositivos mencionados.

Tabla 1.1: Variables de dosimetría de voz de los principales dispositivos comercializados

Variables	APM	VocaLog	VoxLog
Intensidad (dBSPL)	X	X	X
Respuesta a la intensidad	X	X	X
Frecuencia Fundamental (Hz)	X		X
Respuesta a la frecuencia	X		X
Tiempo de fonación (hh:mm:ss)	X	X	X
Ruido ambiente (dBSPL)			X
Número de ciclos	X		
Distancia recorrida por las cuerdas (m)	X		
Calibración de intensidad	Diaria	Una vez	No necesaria

1.5. Contribuciones

Actualmente, existe una gran cantidad de dispositivos cuyo propósito es registrar varios parámetros fisiológicos. Cada vez dichos dispositivos son más precisos, exactos y asequibles; sin embargo, llama mucho la atención el escaso desarrollo de dispositivos en el ámbito de registro de parámetros de la voz desde el ámbito médico o fisiológico, tomando en consideración que la voz es el medio habitual de comunicación de los seres humanos y que muchas personas lo usan como herramienta indispensable de trabajo.

El presente proyecto brinda la posibilidad de obtener los parámetros fundamentales de la voz, a través de un dispositivo electrónico, específicamente este proyecto tiene la finalidad de cuantificar y monitorizar los diferentes parámetros fisiológicos vocales que suponen un pilar fundamental de información diagnóstica para el apoyo de los especialistas del aparato fonador. El dispositivo en mención permitirá implementar técnicas y medidas de los parámetros fisiológicos de la voz de manera fiable y precisa que cuente con la debida comodidad para el paciente.

2. Fundamentos teóricos

En la siguiente investigación es importante describir algunos conceptos fundamentales. Este capítulo presenta los siguientes fundamentos teóricos: Análisis de la producción vocal, patologías que causan problemas en la voz, análisis acústico de la voz, audición y sonido.

2.1. Análisis de la producción vocal

El sistema encargado del habla está conformado por un conjunto de estructuras anatómicas que se pueden agrupar en tres estructuras principales [1]:

- Vías subglóticas, formadas por los pulmones, los bronquios y la tráquea.
- La laringe.
- Vías supraglóticas, compuestas por la faringe, cavidad oral, labios y cavidad nasal.

El diagrama esquemático del sistema responsable de la producción de voz se puede ver en la figura 2.1.

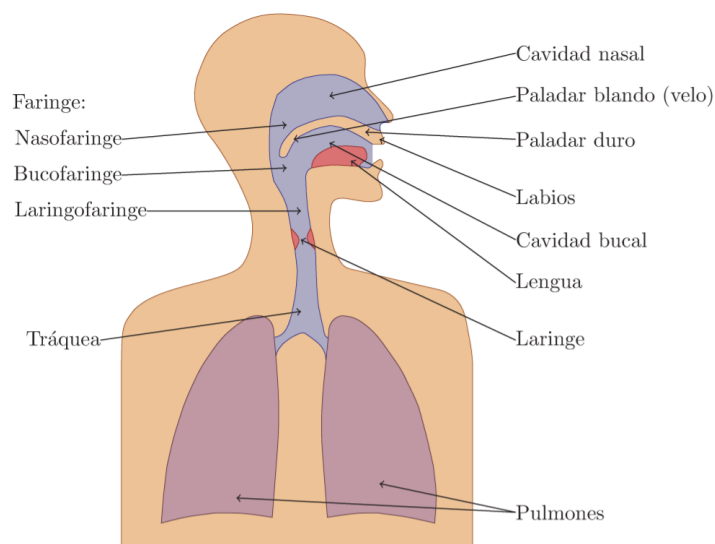


Figura 2.1: Esquema representativo del aparato fonador. Se muestran las principales estructuras anatómicas involucradas en el proceso de la fonación. Tomado de [1].

La mucosa de la laringe forma dos pares de pliegues como se muestra en la figura 2.2: un par superior representado por los pliegues vestibulares (cuerdas vocales falsas), y un par inferior compuesto por los pliegues vocales (cuerdas vocales verdaderas). El

espacio entre los pliegues ventriculares se denomina rima vestibular. Los pliegues vocales son las principales estructuras para la fonación (generación de la voz). Debajo de la mucosa de dichos pliegues se encuentran bandas de ligamentos elásticos estirados entre los cartílagos rígidos de la laringe, como las cuerdas de una guitarra. La tensión que soportan los pliegues vocales controla el tono del sonido. Cuando los músculos tensan los pliegues, estos vibran más rápido y producen un tono más alto. La disminución de la tensión muscular sobre los pliegues vocales hace que vibren con mayor lentitud y produzcan sonidos con un tono más bajo. Los pliegues vocales suelen ser más gruesos y más largos en los hombres que en las mujeres, lo que hace que vibren con mayor lentitud. Esta es la razón por la cual la voz del hombre tiene tonos más graves que de la mujer. [2].

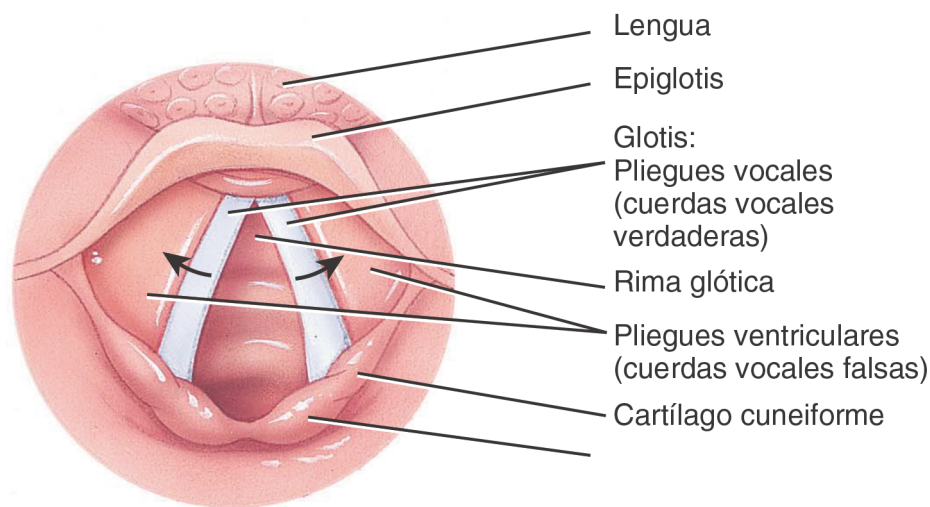


Figura 2.2: Separación (abducción) de los pliegues vocales. Tomado de [2].

El tono de la vibración de las cuerdas vocales se conoce como la frecuencia fundamental de la voz (f_0). Para los hombres el valor de referencia de la media de f_0 es igual a 100 Hz y para las mujeres es de 200 Hz [46, 47]. En un estudio más profundo se define que la frecuencia fundamental media puede adquirir cualquier valor en un rango de variación que va aproximadamente de los 75 Hz a los 300 Hz [48]. Sin embargo, la presencia de patologías en las cuerdas vocales provoca cambios en la estructura física, lo que altera las características de la señal de voz emitida por el hablante [49].

Entre los elementos que constituyen el proceso de producción de la voz se destacan los músculos cricotiroides y tiroaritenoides. El músculo tiroaritenoides se encarga de acortar las cuerdas vocales, mientras que el cricotiroides de tensarlas. En base al funcionamiento de estos músculos, es posible comprender el comportamiento físico de las personas cuando hablan en un tono más alto o más bajo, a mayor tono, mayor vibración de los pliegues y mayor tensión provocada por el cricotiroides. Las figuras 2.3, 2.4

y 2.5, ilustran los músculos intrínsecos de la laringe que incluye a los dos mencionados, cricotiroides y tiroaritenoides.

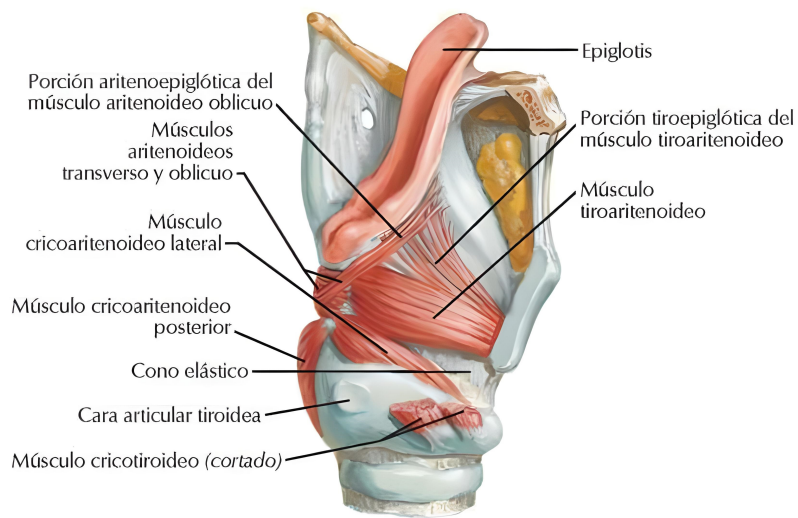


Figura 2.3: Vista lateral. Tomado de [2].

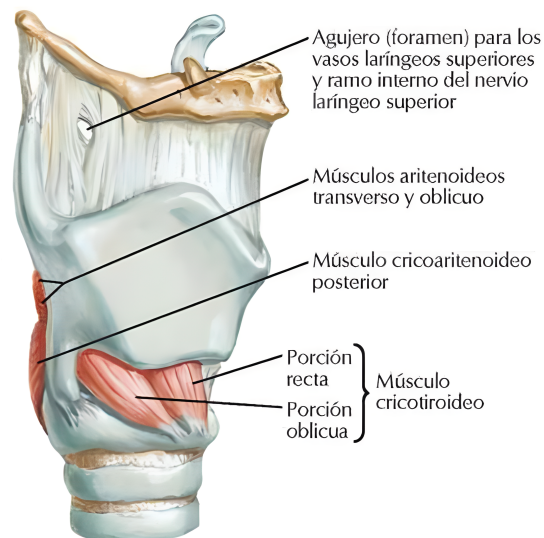


Figura 2.4: Vista lateral derecha. Tomado de [2].

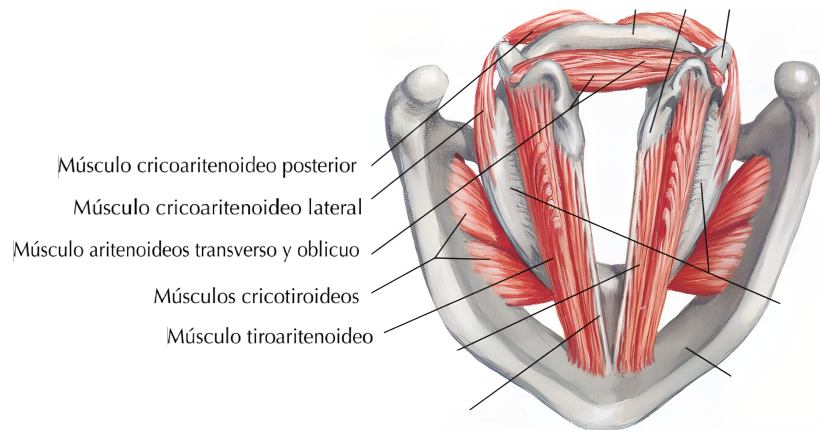


Figura 2.5: Vista superior. Tomado de [2].

Las cuerdas vocales son músculos que trabajan exhaustivamente y tienen estructuras para proporcionar tensión y al mismo tiempo elasticidad durante las tareas del habla. El aspecto mioelástico del control fonatorio se refiere principalmente al control neuromuscular de la tensión y elasticidad de las cuerdas vocales durante la fonación. El movimiento de auto-oscilación de las cuerdas vocales es proporcionado por tres principios aerodinámicos, que son:

1. El aire fluye de una región de alta presión a una región de baja presión;
2. La presión de un flujo incompresible disminuye a medida que aumenta la velocidad de las partículas del fluido, según el principio de conservación de la energía del fluido (ley de Bernoulli);
3. La velocidad de las partículas de un flujo incompresible confinado en un conducto aumenta a medida que disminuye el área de la sección transversal del conducto (ley de continuidad del flujo incompresible). Por lo tanto, al iniciar una tarea de habla, el individuo inicia la vibración de los pliegues que se encuentran en un estrechamiento en la laringe (glotis), así, la diferencia de presión entre la subglotis y la supraglotis provoca que el flujo de aire que pasa por este estrechamiento causa el efecto de succión, en condiciones normales se produce una auto oscilación [3]. La figura 2.6, ilustra el efecto causado por la diferencia de presión entre la subglotis y la supraglotis.

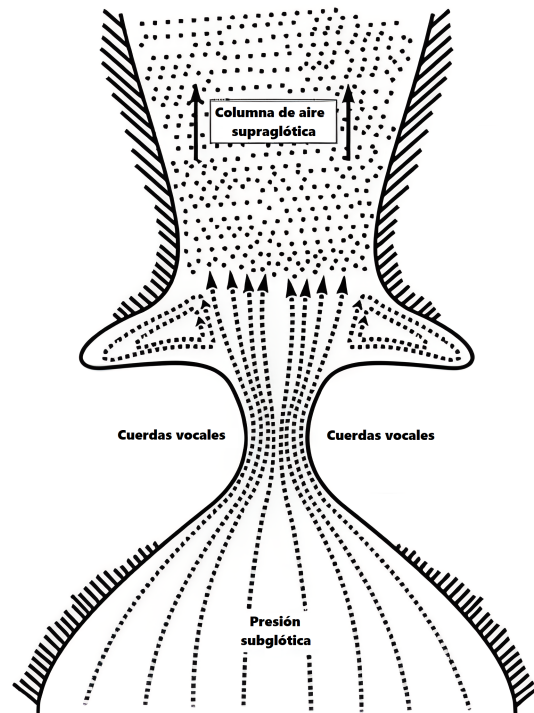


Figura 2.6: Dibujo esquemático de la dirección del flujo de aire a través de la glotis y en contacto con la columna de aire supraglótica. Tomado de [3].

En el proceso de fonación es posible detectar al menos dos señales, la acústica y la vibratoria, provenientes de las cuerdas vocales en el movimiento de auto oscilación, como se ilustra en la figura 2.7.

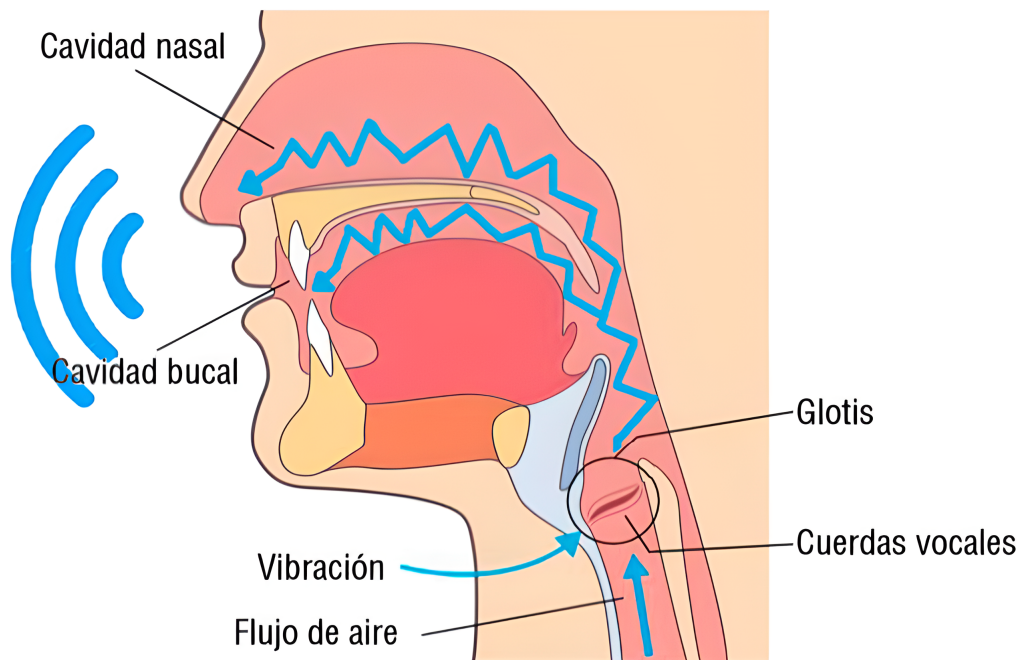


Figura 2.7: Producción de la voz. Adaptado de [4]

2.2. Patologías que causan problemas en la voz

Las patologías de la garganta suelen presentar cuadros leves que afectan a la faringe y a la laringe. La OMSs calcula que, ante cuadros leves, entre el 50 y el 60% de la población no recurre a los sistemas de salud y se automedica. Una inadecuada automedicación provoca un auténtico problema de salud pública mundial. Según la Sociedad Española de Otorrinolaringologías (SEORLs), las patologías más comunes que afectan a la garganta son producidas por infecciones virales. También son frecuentes las infecciones bacterianas y las ocasionadas por procesos inflamatorios. Podemos decir, por tanto, que el origen puede ser infeccioso y no infeccioso [50]. Dentro de este tipo de patologías, existen algunas que causan problemas en la voz.

Origen infeccioso

- Virus.
- Bacterias.
- Hongos, muy raramente, en pacientes inmunodeprimidos en tratamiento con antineoplásicos, trasplantados o enfermos de sida.

Origen no infeccioso

- Etiología alérgica.

- Etiología irritativa:
 - La tos, el goteo retrorinal o la sequedad de la mucosa, síntomas de otros procesos respiratorios, irritan la garganta.
 - Elementos como el humo, el polvo, el tabaco, los picantes, el frío o los cambios bruscos de temperatura.
- Etiología traumática:
 - El reflujo gastroesofágico que quema con su acidez.
 - Tomar alimentos demasiado calientes o fríos.
 - Cuerpos extraños.
 - Profesiones que utilizan la voz como herramienta principal de su trabajo y en las que el mal uso o abuso produce alteraciones y daños en la laringe (pólipos y nódulos).
- Etiología iatrogénica: algunos medicamentos (corticoides, antihistamínicos, broncodilatadores, vasoconstrictores nasales, etc.) que producen sequedad en la garganta.

2.2.1. Patologías relacionadas con el COVID-19

La población de pacientes que corren el riesgo de padecer un problema de voz debido a la propia enfermedad de COVID-19, o secundario a la VMI en casos graves de la enfermedad ha ido creciendo y provocando una mayor demanda de servicios. Pruebas indican que el tiempo de intubación se relaciona con la incidencia y la seriedad de las lesiones laríngeas que dan lugar a un mayor riesgo de disfonía (76 %) y disfagia (49 %), después de la extubación [51], mientras que para pacientes que no fueron intubados, existe un 28.6 % que presentaron síntomas de disfonía. Las infecciones del tracto respiratorio afectan el mismo sistema y estructura que se utilizan para la producción de la voz. Aunque la función principal del tracto respiratorio es proporcionar oxígeno al cuerpo, su función secundaria e igualmente importante es proporcionar energía y producir fonación con el fin de comunicarse con el habla [52].

Implícitamente, la voz también se ve afectada, lo que resulta en una falta de energía para producir sonido y pérdida de la voz causada por dificultad para respirar y congestión de las vías respiratorias superiores. La tos seca recurrente puede influir aún más en los cambios en las cuerdas vocales que afectan la calidad de la voz. Un estudio reciente ha informado cambios en los parámetros acústicos de la voz causados por el flujo de

aire insuficiente a través del tracto vocal como consecuencia de compromisos pulmonares y laringológicos en personas con COVID-19. Por lo tanto, todas estas afecciones respiratorias causadas por la COVID-19 pueden hacer que la voz de los pacientes sea distintiva, creando firmas de voz identificables [53].

2.2.2. Intubación endotraqueal

La intubación endotraqueal es una técnica que consiste en el paso de un tubo flexible provisto de un manguito a través de la boca (intubación orotraqueal) o nariz (intubación nasotraqueal) y laringe hasta llegar a la tráquea, con el propósito de mantener la vía aérea permeable en el proceso de ventilación. La intubación de la tráquea con un tubo endotraqueal es un procedimiento anestésico muy común. El control de la vía aérea durante los procedimientos de anestesia general requiere de intubación endotraqueal [54]. En la figura 2.8, se muestran las partes de un tubo endotraqueal que se utiliza comúnmente [55].

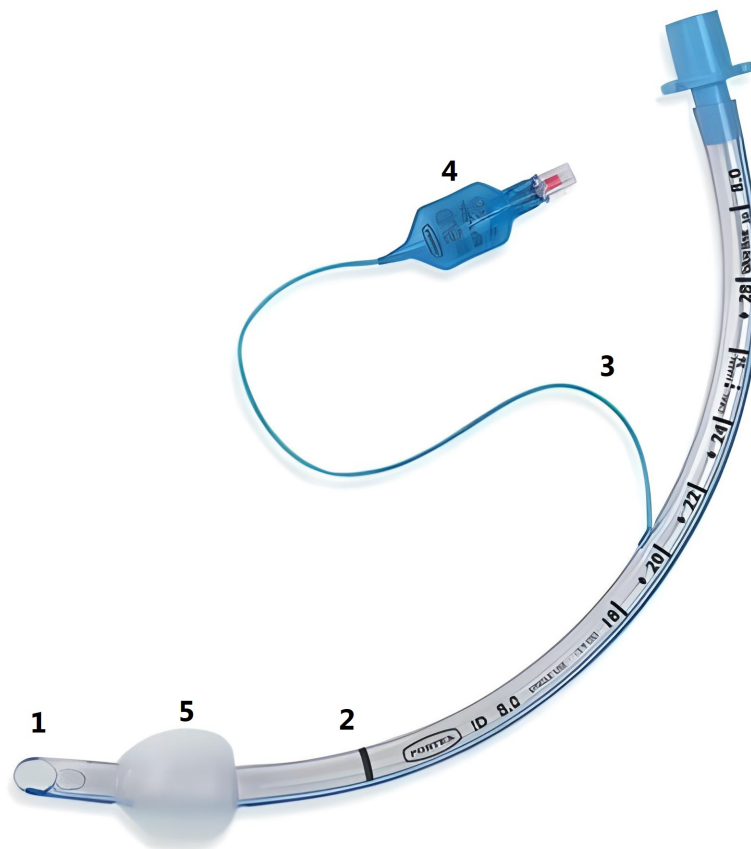


Figura 2.8: Tubo endotraqueal de alto volumen y baja presión. (1) Punta biselada redondeada, (2) guía de profundidad, (3) tubo de inflado, (4) balón piloto con válvula, (5) manguito endotraqueal. Tomado de [5]

Complicaciones a corto y largo plazo

En las complicaciones de la intubación prolongada intervienen varios factores. Es necesario no descuidar los mismos para atender a los pacientes que cumplen con estas tendencias. A continuación se enumeran los principales factores de riesgo para complicaciones post intubación prologada [56].

1. Tipo de sonda endotraqueal.
2. Características del tubo.
3. Trauma durante la intubación.
4. Duración de la intubación.
5. Enfermedades coexistentes.
6. Presión del globo de la sonda mayor a 30mmhg.
7. Septicemia.
8. Reflujo por aplicación de sonda nasogástrica.
9. Movilidad de la sonda y movilidad de región cervical.
10. Estado de conciencia y tolerancia del tubo endotraqueal.
11. Hipotensión por disminución de la perfusión.
12. Estado nutricional y metabólico del paciente.
13. Alteración del aparato mucociliar.

En un estudio a pacientes, se encontraron principalmente granulomas en el espacio interaritenoides y pared lateral y posterior de la subglotis, formaciones polipoideas en los primeros anillos traqueales y reducción notable del espacio glótico a causa de fijaciones cicatrizales de una o ambas cuerdas vocales. Todos estos pacientes tenían de común el haber permanecido intubados varios días, con insuflación completa del manguito endotraqueal durante todo el tiempo de intubación, por lo que la irritación y compresión sobre la mucosa del órgano fue considerable.

A las tres horas de intubación se encontraron marcas isquémicas en la mucosa, exactamente en los sitios de fijación del manguito endotraqueal del tubo endotraqueal, y en ambas paredes laterales de la subglotis, inmediatamente por debajo de las cuerdas vocales. A medida que el tiempo de intubación fue mayor, dichas lesiones iniciales fueron haciéndose más amplias y profundas, llegando al cabo de pocos días a convertirse en verdaderas úlceras necróticas que destruían el cartílago subyacente [57].

Los nódulos, la laringitis, los pólipos y granulomas de cuerdas vocales, son las afecciones laríngeas más comunes, unido al hecho de que en el uso y abuso de la voz intervienen factores de riesgo predisponentes, desencadenantes o agravantes como el sexo, la edad, la ocupación, los hábitos tóxicos y los trastornos respiratorios y digestivos [58].

2.3. Análisis acústico de la voz

El análisis acústico de la voz se encarga de caracterizar y cuantificar de manera objetiva y no invasiva la señal acústica obtenida a partir de la fonación. Es utilizado como apoyo del diagnóstico de la voz y para evaluar los datos de manera objetiva. El análisis se lo realiza a través de diversos parámetros vocales los cuales varían de acuerdo a las necesidades del evaluador [59–61]. Los parámetros medidos se basan principalmente en la composición de señal adquirida y al ser una señal acústica se tendrán los siguientes parámetros: periodicidad, amplitud, duración y composición espectral [62]. El principal objetivo del análisis acústico de la voz es la de realizar diferentes medidas de la voz y recoger en este proceso diferentes aspectos de las características acústicas vocales permitiendo así captar cualquier anomalía en la voz.

2.3.1. Principales parámetros de la voz

2.3.1.1. Frecuencia fundamental

La frecuencia fundamental vocal (f_0) o tono vocal, es una medida referente a la vibración de las cuerdas vocales por unidad de tiempo mientras se realiza un esfuerzo vocal, este parámetro es medido en Hertz [Hz] y está determinado por el número de ciclos de “apertura y cierre” que realizan los pliegues vocales. Este parámetro es asociado directamente con el tono de la voz, es decir, permite diferenciar tonos graves de agudos, siendo los tonos graves o bajos de baja frecuencia y los tonos agudos o altos de mayor frecuencia. El parámetro en mención cuantifica la eficiencia del aparato fonador, biomecánica laríngea y su interacción con la aerodinámica, por lo que este parámetro constituye la base para una evaluación apropiada de la laringe a nivel funcional y anatómico [62]. En [63], define la frecuencia fundamental como el número de pulsos de aire que atraviesan la glotis por unidad de tiempo (por segundo), el rango de la frecuencia fundamental en una persona adulta está en 180 a 250 Hz en el caso de las mujeres y en hombres se encuentra en el rango de 100 a 150 Hz .

Con la finalidad de identificar un proceso de fatiga vocal la media y la moda de la frecuencia fundamental pueden ser representativas. El uso excesivo de la voz causa

fatiga vocal, por consiguiente, provoca una elevación del valor medio de f_0 [64]. El cálculo de la moda representaría el valor más repetitivo de un conjunto de datos. La media y la moda se encuentran en valores muy próximos de frecuencia fundamental para casos normales de emisión de voz [65].

2.3.1.2. Nivel de Intensidad acústica

La intensidad es asociada directamente con la amplitud de la frecuencia, y depende primordialmente de la presión sonora, esta es relacionada con el volumen, es decir mediante este parámetro es posible diferenciar sonidos suaves o débiles de los sonidos intensos o fuertes. En lo que a voz humana se refiere, la intensidad es regulada mediante el control del proceso de inspiración y expiración. La escala de nivel de intensidad acústica o de presión sonora se muestra en la ecuación 2.1:

$$L_{db} = 10 \log \left(\frac{I}{I_{ref}} \right) = 20 \log \left(\frac{P_{eff}}{P_{ref}} \right) \quad (2.1)$$

Donde P_{eff} y P_{ref} representan respectivamente el valor eficaz de la presión y la intensidad acústica que corresponde al umbral de sensibilidad de la audición por un sonido puro de 1000 Hz y se muestra en las ecuaciones 2.2 y 2.3 respectivamente:

$$P_{ref} = 20 \mu Pa \quad (2.2)$$

$$I_{ref} = 10 pW/m^2 \quad (2.3)$$

Cabe mencionar, que estos valores son empíricos y resultan de medidas efectuadas sobre un gran número de individuos [66].

2.3.1.3. Tiempo de fonación

El tiempo de fonación hace referencia al tiempo hablado o fonado durante la medición, es decir el período de tiempo durante el cual las cuerdas vocales ejercen esfuerzo o están vibrando, se define también como el tiempo durante el cual ocurren períodos fundamentales detectables [67]. Es importante también mencionar la diferencia entre el tiempo de fonación y el tiempo de locución. El tiempo de fonación es aquel que las cuerdas vocales emiten vibraciones y el tiempo de locución es el tiempo en que además de las vibraciones de las cuerdas vocales se incluyen los tiempos de silencio además de los segmentos sordos y sonoros [68].

2.3.1.4. Timbre

El Timbre de voz es el espectro característico y específico de la voz. Depende de la morfología de los resonadores y como se aproximan las cuerdas vocales. La calidad del timbre es mayor cuando existe una mayor adhesión de las cuerdas vocales cuando se juntan entre sí. Cada individuo tiene una voz única debido a su timbre de voz que depende de las dimensiones del tracto vocal y la configuración laríngea. El timbre de voz consta de diversas cualidades como: color, volumen, espesor, mordiente y vibrato [69–71].

2.3.1.5. Ataque y filatura

El ataque es el nombre que se otorga al inicio de la fonación y es clasificado en brusco (duro), soplado y normal [69–71]. La Filatura es el nombre que se otorga al final de la fonación o emisión de las cuerdas vocales, determina si la fonación culmina en cadencia, anticadencia o monótona [72].

2.3.2. Parámetros secundarios

Una vez analizados los parámetros principales, es posible calcular y obtener otros parámetros vocales a partir de los tres primeros anteriores (frecuencia fundamental f_0 , nivel de Intensidad acústica I , tiempo de fonación t). A continuación, se describen los parámetros de relevancia.

2.3.2.1. Shimmer

El *shimmer* es el parámetro que mide la variabilidad de la amplitud de una señal de voz, indica la diferencia de amplitud que existe entre los ciclos glóticos vecinos. Dicho parámetro está relacionado con la reducción de la resistencia glótica, siendo así un indicativo de la presencia de ruido en la fonación, mejor conocido como ronquera y también de la respiración entrecortada o aire sin sonar [73]. Un parámetro muy utilizado para el análisis de las señales de voz es el *shimmer* (local), que es expresada como la diferencia de la media absoluta entre las amplitudes de dos períodos consecutivos de la señal, dividida por la amplitud media, tal como se muestra en la ecuación 2.4 [62].

$$shimmer(local) = \frac{\frac{1}{N_g - 1} \sum_{i=1}^{N_g - 1} |A_i - A_{i+1}|}{\frac{1}{N_g} \sum_{i=1}^{N_g} A_i} * 100 \quad (2.4)$$

Donde, A_i representa la amplitud del impulso glotal i , y N_g el número de impulsos glotales.

2.3.2.2. Jitter

El *jitter* es el parámetro que mide la variabilidad de la frecuencia, específicamente el *jitter* es un índice de perturbación de f_0 a corto plazo, indicando así la variabilidad de la frecuencia fundamental medida entre los ciclos glóticos vecinos. Dicho índice muestra la diferencia entre el período antecesor y el período sucesor de una señal de voz, si ambos períodos fuesen iguales el sistema sería perfecto, resultando un valor de *jitter* de cero. Este parámetro se relaciona con la aspereza en la voz y también es un indicativo de falta de control en la vibración de las cuerdas vocales [73].

Otra forma muy utilizada para el análisis de la voz es el *jitter* (local), que se representa como la media absoluta de la diferencia de medias entre dos períodos consecutivos, dividida para el período promedio, tal como se muestra en la ecuación 2.5 [62].

$$jitter(local) = \frac{\frac{1}{N_p-1} \sum_{i=1}^{N_p-1} |T_i - T_{i-1}|}{\frac{1}{N_p} \sum_{i=1}^{N_p} T_i} * 100 \quad (2.5)$$

Donde, T_i representa la duración de cada período de f_0 y N_p el número de períodos de f_0 .

2.3.2.3. HNR (Harmonic-to-Noise Ratio)

Harmonic-to-Noise Ratios (HNRs) o relación armónico-ruido es un índice que vincula el componente armónico con el componente de ruido de una señal acústica. Este parámetro es considerado como uno de los mejores en aplicaciones clínicas, se relaciona directamente con la calidad vocal, debido a que cuantifica anomalías vocales. El valor promedio viene expresado en decibelios y es menor en hombres que en mujeres [73].

Un método desarrollado por [74], relaciona la medición de la energía acústica del armónico y la energía del componente de ruido, expresadas en las ecuaciones 2.6 y 2.8 respectivamente.

$$H = n_s \cdot \int_0^T f_A^2(\tau) dt \quad 0 \leq \tau \leq T ; \quad (2.6)$$

$$f_a(\tau) = \sum_{i=1}^{n_s} \frac{f_i(\tau)}{n_s} ; \quad (2.7)$$

$$N = \sum_{i=1}^{n_s} \int_0^{T_i} [f_i(\tau) - f_A(\tau)]^2 dt \quad 0 \leq \tau \leq T \quad (2.8)$$

Donde,

H es la energía del componente armónico;

N es la energía del componente del ruido;

f_i es la porción de señal a lo largo de la duración del período fundamental;

T_i es la duración del i -ésimo período;

T es el valor máximo que T_i puede asumir;

n_s es el número de ventanas segmentadas.

Para expresar la relación en decibelios se utiliza la ecuación 2.9.

$$\frac{H}{N}(dB) = 10 \log_{10} \left(\frac{H}{N} \right) \quad (2.9)$$

El HNRs es una relación entre los valores de H y N , es decir, la proporción de armónicos basados en el ruido $\frac{H}{N}$ [74]. Para poder determinar la proporción del ruido basados en los armónicos de la señal, es necesario calcular la relación inversa de HNR representado por $\frac{N}{H}$, dicha relación lleva el nombre de NHR (Noise-to-Harmonic Ratio).

2.3.2.4. Energía

La Energía en un intervalo de tiempo, aporta información importante sobre una señal. Para garantizar que la señal sea estacionaria se necesita segmentarla en intervalos de tiempo entre 16 y 32 ms , para poder así obtener la energía para cada uno de los segmentos de la señal. La ecuación 2.10 muestra el cálculo de la energía por cada segmento [75].

$$E_{seg} [n] = \sum_{n=0}^{N_A-1} [s(n)]^2 \quad (2.10)$$

Donde,

E_{seg} es el valor de la energía por segmento;

$s(n)$ es la señal con muestras de la señal de vibración en un segmento;

N_A es el tamaño de la ventana (bloque de muestras de la señal).

La ecuación 2.11 muestra la energía expresada en decibelios.

$$E_{seg}(dB) = 10 \log [E_{seg}] \quad (2.11)$$

2.4. Audición

La audición es uno de los cinco sentidos que tenemos los seres humanos, este sentido está funcionando todo el tiempo, debido a que en cualquier momento se perciben los sonidos producidos. La audición es la percepción de las ondas sonoras que viajan primero a través de la oreja para posteriormente llegar a los conductos auditivos externos y chocar con el tímpano que vibra con ellas y finalmente el cerebro es el encargado de interpretar las ondas sonoras y su significado, discerniendo de ruido o un sonido determinado [76].

2.4.1. Audición humana y niveles de intensidad del sonido

El oído humano es un transductor que se encarga de convertir las ondas sonoras en impulsos nerviosos eléctricos de manera mucho más precisa y sofisticada que un micrófono actual. La figura 2.9, muestra la anatomía del oído [76, 77].

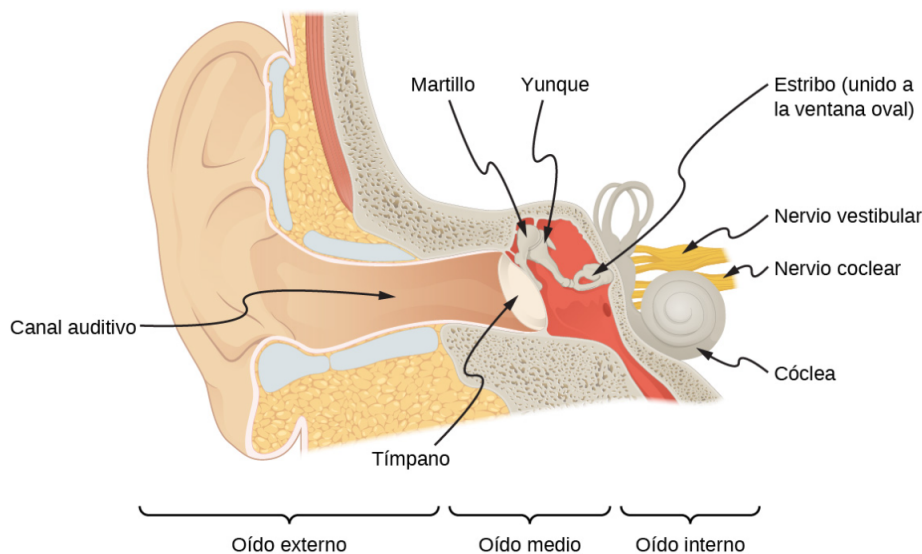


Figura 2.9: Anatomía del oído humano. Tomado de [6].

El canal auditivo o también llamado oído externo, es el encargado de llevar el sonido hasta el tímpano. la columna de aire formada dentro del canal auditivo, resuena y es parcialmente responsable de la sensibilidad del oído a los sonidos en el rango de 2000 y 5000 Hz . El oído medio es el encargado de convertir el sonido en vibraciones mecánicas y estas vibraciones son aplicadas a la cóclea [77].

El rango de intensidades que el oído humano puede percibir estará relacionado con la frecuencia del sonido, por lo general este rango es muy amplio. La intensidad mínima

del umbral audible es $I_0 = 10^{-12} \text{W/m}^2 = 10 \text{ pW/m}^2$. El dolor es experimentado cuando se presentan intensidades de $I_{dolor} = 1 \text{W/m}^2$ [76, 77].

La intensidad expresada en decibelios ha sido descrita en la ecuación 2.1. En la tabla 2.1, se presenta la medida en decibelios (dB) y en vatios por metro cuadrado (W/m^2) de la intensidad del sonido para varios sonidos o efectos conocidos [76, 77].

Tabla 2.1: Niveles de intensidad del sonido e intensidades

Nivel de intensidad del sonido L_{dB}	Intensidad del sonido I (W/m^2)	Ejemplo/efecto
0	1×10^{-12}	Umbral de audición a 1000 Hz
10	1×10^{-11}	Crujido de hojas
20	1×10^{-10}	Susurro a 1m de distancia
30	1×10^{-9}	Hogar tranquilo
40	1×10^{-8}	Hogar regular
50	1×10^{-7}	Oficina regular, música suave
60	1×10^{-6}	Conversación normal
70	1×10^{-5}	Oficina ruidosa, tráfico congestionado
80	1×10^{-4}	Radio con volumen alto, conferencia en un aula
90	1×10^{-3}	Dentro de un camión pesado; daños por exposición prolongada
100	1×10^{-2}	Fábrica ruidosa, sirena a 30m; daños por exposición de 8h al día
110	1×10^{-1}	Daños por una exposición de 30 minutos al día
120	1	Concierto de rock ruidoso; trituradora neumática a 2m; umbral de dolor
140	1×10^2	Avión a reacción a 30 m; dolor intenso, daños en segundos
160	1×10^4	Estadillo de los tímpanos

2.4.2. Audición y tono

El oído humano es muy sensible y tiene un gran alcance, puede dar una gran cantidad de información sencilla, como el tono, volumen y la dirección u orientación del sonido [7, 76].

La percepción de la frecuencia se denomina tono. Habitualmente, los seres humanos tienen un tono relativo excelente y pueden diferenciar entre dos sonidos si sus frecuencias difieren entre sí en un 0.3% o más. Ejemplificando, si se tiene una frecuencia de 500 Hz y 500,5 Hz son notablemente diferentes. Un aspecto adicional a mencionar es que las notas musicales son sonidos a una frecuencia determinada y pueden ser producidos por la mayoría de instrumentos musicales conocidos [7, 76].

La percepción de la intensidad acústica se denomina volumen. A una frecuencia determinada es posible identificar diferencias de al rededor de 1 dB y un cambio de 3 dB es muy notorio. Es importante destacar que el volumen no está relacionado únicamente con la intensidad acústica, ya que la frecuencia también forma parte esencial en el volumen de un sonido [7].

A la percepción de la combinación de frecuencias e intensidades se la denomina calidad del tono o más conocida como timbre del sonido. El timbre es la forma de onda que surge de las varias reflexiones, resonancias y superposiciones de un instrumento [7].

Numéricamente, es posible expresar el volumen mediante la unidad denominada *fonio*. Los fonios se diferencian de los decibelios ya que el fonio es una unidad de percepción del volumen y el decibelio es una unidad de intensidad física. La figura 2.10, muestra la relación de volumen e intensidad (o nivel de intensidad) y frecuencia para personas con audición normal. Las curvas fueron determinadas de manera empírica haciendo que un gran número de personas compararan el volumen de los sonidos a diferentes frecuencias y niveles de intensidad del sonido. Como se observa en la figura 2.10, a una frecuencia de 1000 Hz, los fonios son iguales numéricamente a los decibelios [7, 76].

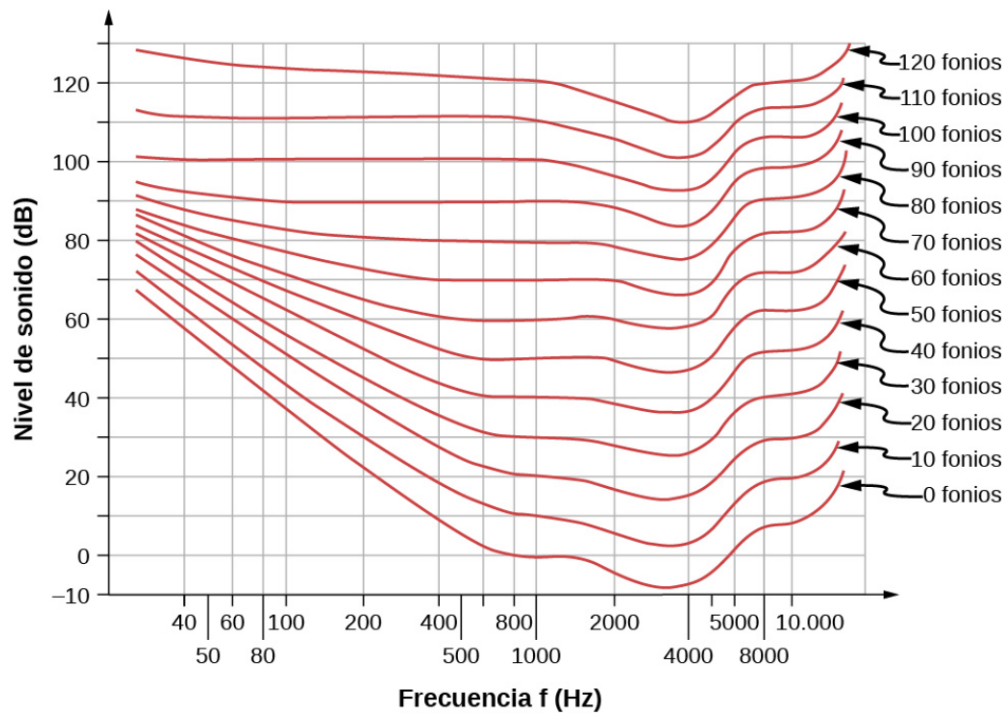


Figura 2.10: Relación entre volumen en fonios y nivel de intensidad (en decibelios) e intensidad (en vatios por metro cuadrado) para personas con audición normal. Tomado de [7].

2.5. Sonido

El sonido es un fenómeno físico, es una onda mecánica longitudinal propagada a través de un medio elástico [76]. El sonido, como todas las ondas, se desplaza a una cierta velocidad y tiene las propiedades de frecuencia y longitud de onda. En el aire, que es el medio que habitualmente nos referimos, el fenómeno se propaga por la vibración de las moléculas de aire situada cerca de un elemento vibrante. La velocidad de propagación en este medio, en condiciones normales de presión y temperatura, es de aproximadamente 340 m/s [7].

2.5.1. Ondas sonoras

Las ondas sonoras son ondas longitudinales que se propagan en un medio elástico y transmiten lo que se asocia como sonido. La velocidad de la onda está relacionada con su frecuencia y longitud de onda mediante la ecuación 2.12 [76, 78].

$$v = f\lambda \quad (2.12)$$

Donde, v es la velocidad de la onda, f es su frecuencia que es el número de ondas que pasan por un determinado punto por unidad de tiempo y λ es la longitud de onda que es la distancia entre partes secuenciales idénticas de una onda, por ejemplo, entre compresiones secuenciales como se muestra en la figura 2.11.

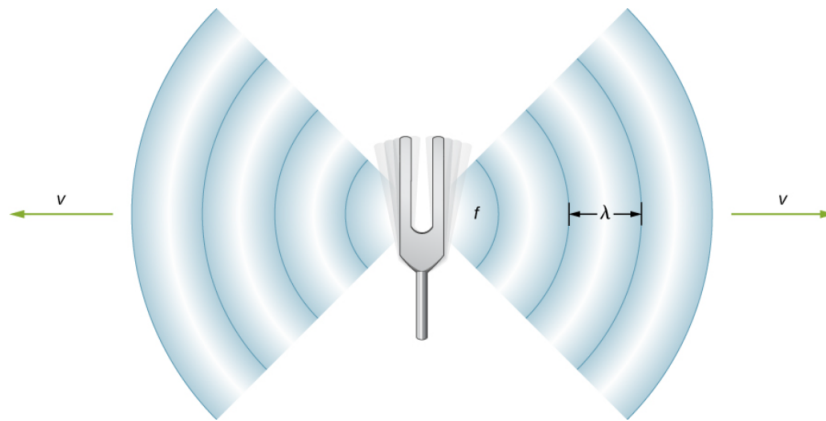


Figura 2.11: Onda sonora emana de una fuente, como un diapasón, que vibra a una frecuencia f . Se propaga a una velocidad v y tiene una longitud de onda λ . Tomado de [8].

2.5.2. Representación gráfica del sonido

Actualmente, los equipos modernos de análisis digital de la voz y varios programas de análisis de la misma, permiten obtener una extensa variedad de imágenes para representar el sonido, ruidos, silencios, etc. Aunque de manera fundamental y básica los equipos y programas se basan en dos tipos de gráficos: oscilogramas y espectrogramas [79].

- **Oscilograma:** Es la imagen acústica que capta directamente los sonidos, ruidos y silencios del habla, esta tiene forma de ondas que oscilan, de ahí el nombre de oscilograma. Es posible observar y medir la amplitud en el eje de las ordenadas y el tiempo en el eje de las abscisas.
- **Espectrograma:** Es la imagen acústica que capta el sonido en su evolución frecuencial en un lapso de tiempo, es decir, desde otra perspectiva que el oscilograma, mostrando la frecuencia en el eje de las ordenadas y el tiempo en el eje de las abscisas. Existen dos tipos de espectrogramas según el ancho de banda del filtro que se utiliza. El espectrograma de banda ancha y banda estrecha. En el caso de los espectrogramas de banda ancha es posible obtener una gran resolución temporal. En cambio, en los espectrogramas de banda estrecha se obtendrá una buena resolución frecuencial.

- **Filtro de banda ancha (300Hz):**

Este tipo de filtro dentro del espectrograma, aumenta la resolución en tiempo, pero a su vez disminuye la resolución en frecuencia. La principal ventaja de este filtro es que mejora la percepción de los formantes (líneas horizontales relacionadas con la forma y tamaño de las cavidades resonantes del tracto vocal), pero perjudica en alguna medida la identificación de los armónicos individuales y su distinción del ruido. En la figura 2.12, se puede observar un ejemplo de una vocal “a” aplicando este filtro.

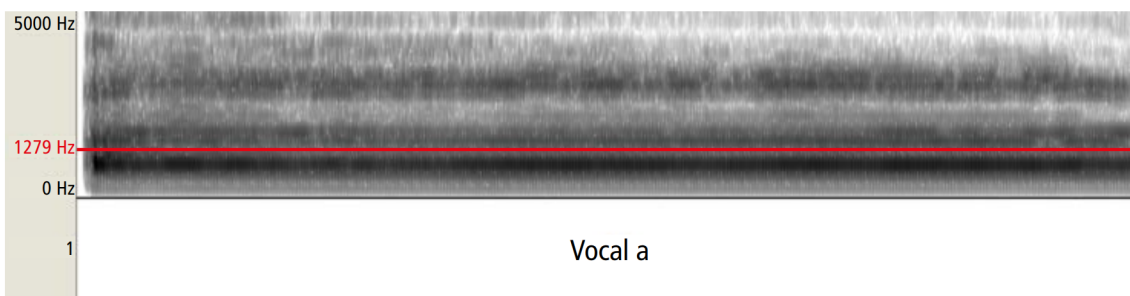


Figura 2.12: Espectrograma de banda ancha, sonido: vocal “a”. Obtenido en Praat. Tomado de [9].

- **Filtro de banda ancha (45Hz):**

Este filtro dentro del espectrograma, permite la eliminación de frecuencias muy próximas, pero un aspecto negativo es que su respuesta es lenta. Este tipo de espectro se utiliza más en análisis de la voz, ya que distingue entre armónicos y ruido subyacente, y refleja cambios de frecuencia en un período de tiempo, como el vibrato, aunque la distinción de los formantes se ve perjudicada. En la figura 2.13, se puede observar un ejemplo de una vocal “a” aplicando este filtro.

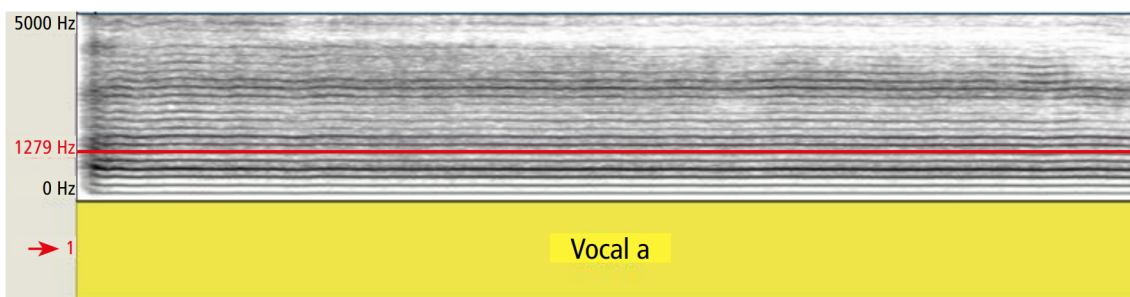


Figura 2.13: Espectrograma de banda estrecha, sonido: vocal “a”. Obtenido en Praat. Tomado de [9].

En la figura 2.14, en la parte superior se muestra un ejemplo de un oscilograma y

en la parte inferior un ejemplo de un espectrograma.

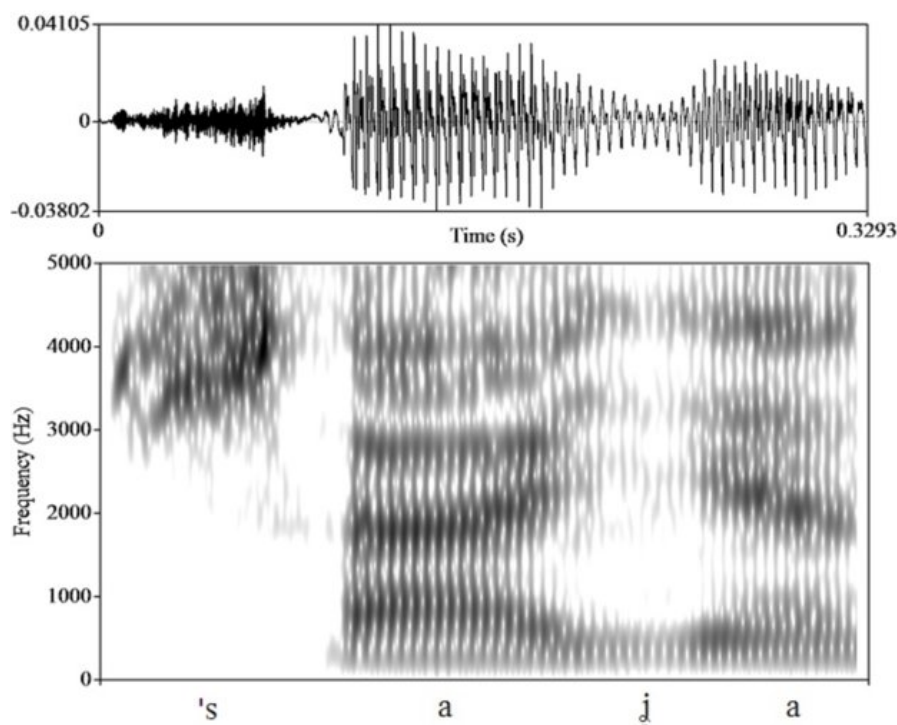


Figura 2.14: Oscilograma y Espectrograma. Tomado de [10].

2.5.3. Caracterización del ruido

El ruido es la sensación auditiva no deseada y corresponde generalmente a una variación aleatoria de la presión a lo largo del tiempo. Es una emisión de energía originada por un fenómeno vibratorio el cual es percibido por el oído humano y al ser escuchado provoca una sensación de molestia [80].

Existen cuatro principales tipos de ruido, listados a continuación:

Ruido continuo: Es producido de forma constante, un ejemplo podría ser una maquinaria trabajando sin interrupción, el ruido de un motor o sistemas de calefacción. En la figura 2.15, se muestra una representación de este tipo de ruido [80, 81].

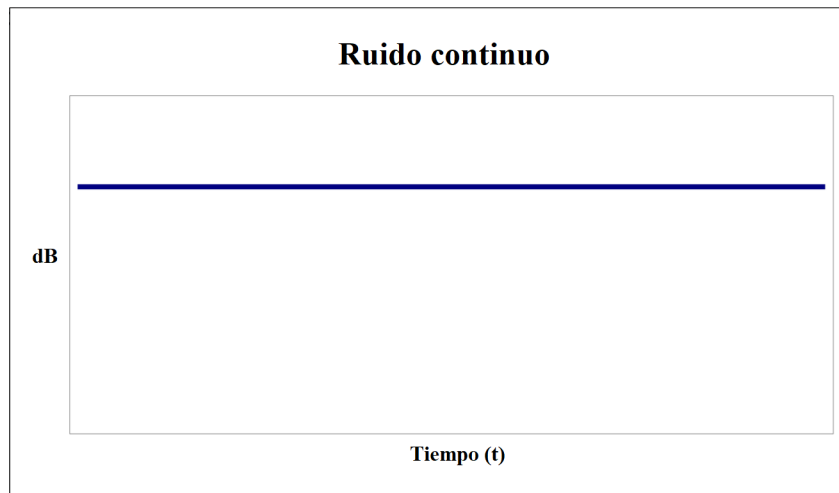


Figura 2.15: Ruido continuo. Tomado de [11].

Ruido continuo intermitente: Nivel de ruido que crece y disminuye de manera rápida, por ejemplo el sonido de un avión despegando, equipos funcionando por ciclos, etc. En la figura 2.16, se muestra una representación de este tipo de ruido [80, 81].

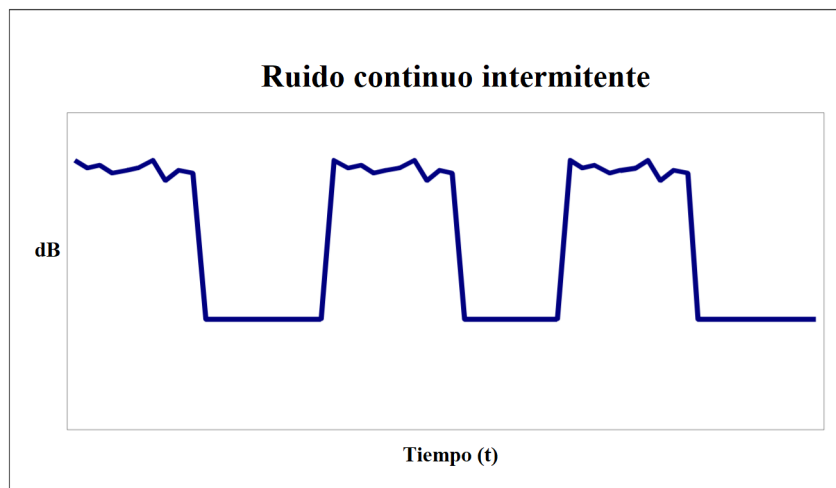


Figura 2.16: Ruido continuo intermitente. Tomado de [11].

Ruido impulsivo: Este tipo de ruido es asociado generalmente a la industria de la construcción y la demolición. Son repentinos estadios de ruido que generan molestia por su naturaleza rápida y sorprendente, el ejemplo más común son los ruidos producidos por algún tipo de explosión. En la figura 2.17, se muestra una representación de este tipo de ruido [80, 81].

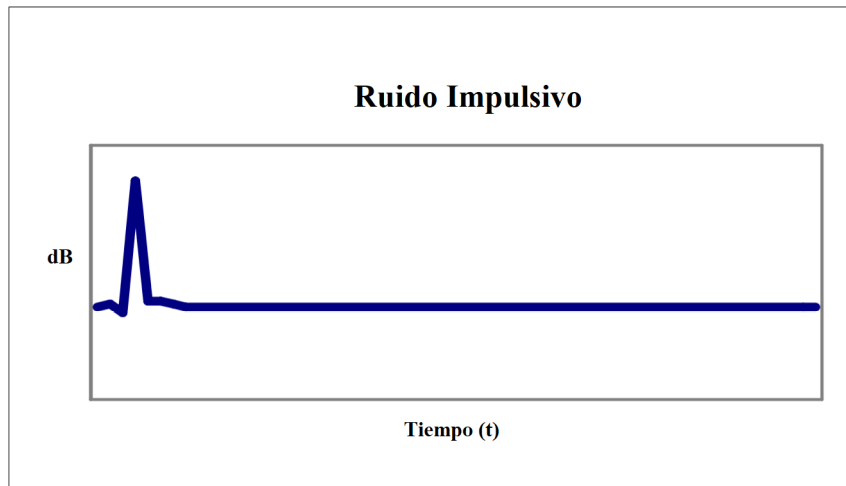


Figura 2.17: Ruido impulsivo. Tomado de [11].

Ruido fluctuante: Es el ruido que forma parte de la cotidianidad del ser humano, ya sea un zumbido de fondo de una central eléctrica cercana, o el sonido de grandes motores diésel, de modo que es algo que está presente de manera constante en la vida diaria. En la figura 2.18, se muestra una representación de este tipo de ruido [80, 81].

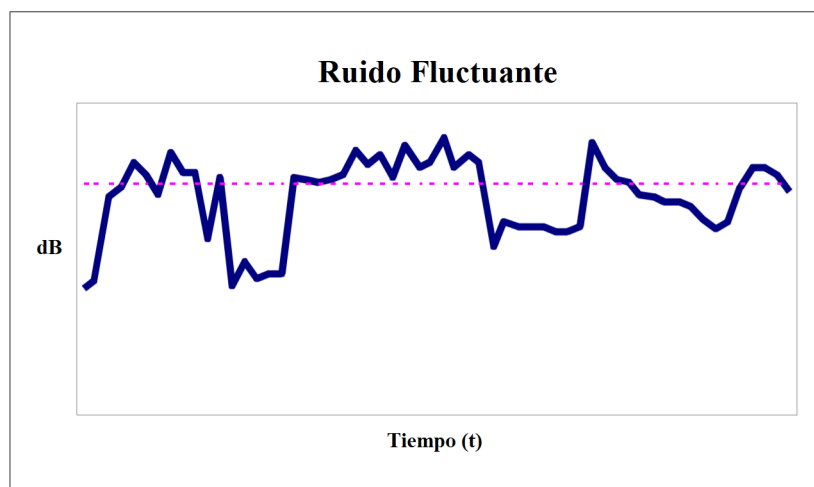


Figura 2.18: Ruido fluctuante. Tomado de [11].

3. Diseño e implementación del dispositivo electrónico para monitoreo y apoyo al diagnóstico de fonoaudiología en pacientes post COVID-19

Este capítulo describe el diseño e implementación del dispositivo electrónico propuesto, el mismo será utilizado para monitoreo y apoyo al diagnóstico en patologías de la voz. Para esto se especifica el procedimiento adecuado para la realización, desde la elección de los materiales hasta la programación para el análisis de los datos.

3.1. Descripción del desarrollo

En la figura 3.1, se muestra el diagrama correspondiente a las actividades a realizar en el desarrollo de este trabajo de titulación.

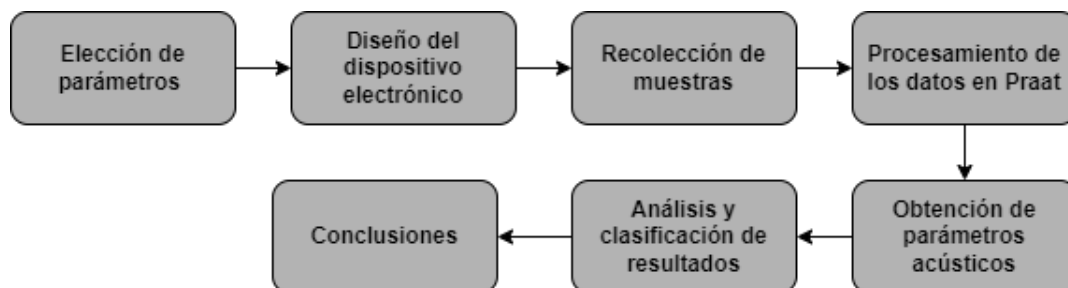


Figura 3.1: Diagrama de bloques. Autoría propia.

3.2. Componentes y materiales utilizados

3.2.1. ESP32

El chip ESP32 ha sido desarrollado por la empresa *Espressif Systems*. Actualmente ofrece varias versiones ESP32 del Security Operations Centers (SoCs) en forma de *ESP32 Developer Kit*, el *ESP32 Wrover Kit*, que también incluye una tarjeta SD y una pantalla Liquid-Crystal Displays (LCDs) de 3,2" y por último el kit ESP32 Azure IoT con USB Bridge y otros sensores incorporados. Además de *Espressif Systems*, otros

productores se dedican a estos chips como: SparkFun con ESP32 Thing DB, WeMoS con su TTGO, D1, Lolin32 y Lolin D32, Adafruit (con Huzzah32), DF Robot (ESP32 FireBeetle). En este contexto, resulta importante mencionar que muchos fabricantes a veces ofrecen dispositivos clonados que no poseen las mismas características de los equipos mencionados anteriormente [82]. La tabla 3.1, muestra una breve descripción de las características principales del chip ESP32 [83], mientras que la figura 3.2, muestra el dispositivo en su totalidad.

Tabla 3.1: Características del ESP32.

Características	ESP32
Microprocesador	Xtensa Dual-Core 32-bit LX6 con 600 DMIPS
Wi-Fi (802.11 b / g / n)	HT40
Bluetooth	Bluetooth 4.2 y BLE
Frecuencia de operación (valor típico)	160 MHz
SRAM	448 KB
Flash	520 KB
GPIO	34
PWM (hardware)	No posee
PWM (software)	16 canales
SPI	4
I2C	2
I2S	2
UART	2
ADC	12-bits de resolución
CAN	Sí
Interfaz MAC Ethernet	Sí
Sensor de tacto	Sí
Sensor de temperatura	Sí (solo las versiones antiguas)
Sensor de efecto hall	Sí
Temperatura de trabajo	-40°C to 125°C

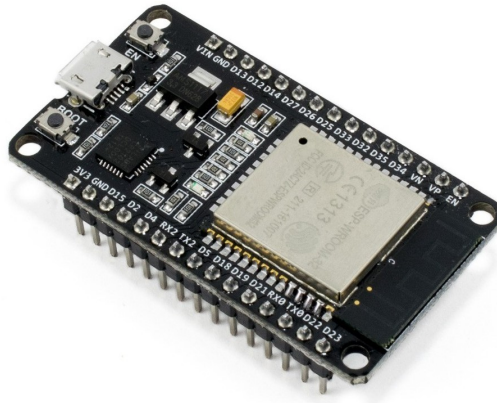


Figura 3.2: ESP32. Tomado de [12].

3.2.2. Módulo micrófono con MAX4466

Este módulo incorpora un micrófono electret de 20KHz amplificado por el MAX 4466, un chip diseñado específicamente para la tarea de preamplificar micrófonos. Este amplificador operacional tiene excelente rechazo de ruido por fuente de poder. En la parte posterior cuenta con un pequeño potenciómetro para ajustar la ganancia de 25x a 125x. La salida del amplificador operacional es riel a riel, por lo que a máxima ganancia puede llegar a 5Vpp. Tiene las siguientes características [84]:

- Voltaje de Operación: 2.4 - 5.5V DC
- Micrófono electret
- Ratio de rechazo de fuente de poder: 112dB
- Salida riel a riel

La figura 3.3, muestra un gráfico del módulo.



Figura 3.3: MAX4466. Tomado de [13].

3.2.3. Módulo lector de tarjeta micro SD

Un lector SD es un dispositivo que permite emplear como almacenamiento una tarjeta SD. Las tarjetas SD y micro SD se han convertido en un estándar, desplazando a otros medios de almacenamiento de datos debido a su gran capacidad y pequeño tamaño. Por este motivo han sido integradas en una gran cantidad de dispositivos, siendo en la actualidad componentes frecuentes en ordenadores, tablets y smartphones, entre otros. Dentro del mundo de Arduino, es posible encontrar lectores de bajo coste tanto para tarjetas SD como micro SD. Los primeros en aparecer fueron los lectores SD y posteriormente los micro SD. En general, los módulos con micro SD son modelos más modernos que los de SD. En ambos tipos de lectores, la lectura puede realizarse a través de bus SPI. Aunque pueden disponer de otros interfaces, como bus I2C o UART, normalmente es preferible emplear SPI por su alta tasa de transferencia.

La tensión de alimentación es de 3.3V, pero en la mayoría de los módulos se incorpora la electrónica necesaria para conectarlo de forma sencilla a Arduino, lo que frecuentemente incluye un regulador de voltaje que permite alimentar directamente a 5V [85]. La figura 3.4, muestra una imagen del dispositivo.

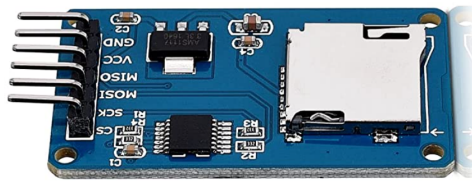


Figura 3.4: Lector micro SD. Tomado de [14].

3.2.4. Acelerómetro BU-21771

El acelerómetro piezocerámico de la serie BU tiene una salida de 20 Hz a más de 10 kHz, y se utiliza principalmente como micrófono de contacto para entornos con mucho ruido. También puede utilizarse como acelerómetro con estructuras ligeras. Como micrófono de contacto, el acelerómetro (BU-21771), puede llevarse en la garganta o tocando una parte ósea de la cabeza (parte superior, sien o hueso mastoideo). La serie (BU) se ha integrado en diseños de cascos para la comunicación por radio de los servicios militares y de emergencia (refiérase al anexo G.1). Para la obtención ideal de los parámetros acústicos es de vital importancia que la respuesta del dispositivo sea lineal en el rango de frecuencias a trabajar (20-2KHz). Esto permite correlacionar la misma amplitud con la misma intensidad relativa a pesar de existir una variación en la frecuencia. Por lo que la señal del acelerómetro tendrá una proporcionalidad lineal a la intensidad de la oscilación que mide. La figura 3.5, muestra una imagen del dispositivo.



Figura 3.5: Acelerómetro BU-21771. Tomado de [15].

Se ha escogido el acelerómetro frente a un micrófono por una razón principal, la cual es, la relación señal a ruido que tiene el acelerómetro, debido a que no se ve afectado por señales externas, más que las vibraciones de las cuerdas vocales; sin embargo el micrófono recoge señales del ambiente que rodea.

3.3. Desarrollo del dispositivo electrónico

El dispositivo que se pretende implementar consta de un microcontrolador ESP32 que junto con un lector micro SD graba los datos recolectados por un acelerómetro. La realización del dispositivo electrónico conllevó una gran cantidad de pruebas, resultado de estas pruebas se tiene el circuito final que se explicará a continuación sección por sección.

3.3.1. Fuente de alimentación de 1.5V

Con el fin de alimentar el acelerómetro según la hoja de datos técnicos el dispositivo BU-21771 el voltaje de alimentación es de 1.5V, entonces se crea la primera parte del circuito. En este caso se realiza un divisor de tensión con los 3.3V que aporta el ESP32 regido por la ecuación 3.1.

Ahora, se impone una resistencia de referencia, en este caso se ha escogido una resistencia R2 de 1KΩ, por lo que la resistencia R1 se calcula y aproxima a la resistencia comercial más cercana como se muestra en la ecuación 3.2, en este caso resulta ser R1=1.2KΩ de esta manera, se obtiene un voltaje de salida $V_{out} = 1.5V$.

$$V_{out} = V_{in} * \frac{R2}{R1 + R2} \quad (3.1)$$

$$\frac{1.5V}{3.3V} = \frac{1K\Omega}{R1 + 1K\Omega}$$

$$0.455R1 + 454.455 = 1000 \quad (3.2)$$

$$0.455R1 = 545.545$$

$$R1 = 1.2K\Omega$$

En la figura 3.6, se muestra dicho circuito implementado en Proteus.

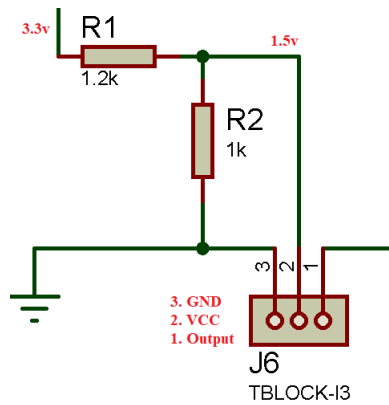


Figura 3.6: Circuito de fuente de alimentación de 1.5V. Autoría propia.

3.3.2. Circuito de pre-amplificación

La señal que produce el acelerómetro es sumamente pequeña en amplitud, por ende, se implementó una etapa de amplificación de la señal. En primer lugar, a la entrada y salida del primer amplificador se colocó capacitores (C1, C4) de 100 μ F cada uno, con la finalidad de atenuar señales de frecuencias altas que no son relevantes para el proyecto. La señal que otorga el acelerómetro ingresa al amplificador operacional LM358 por el capacitor C4 y posteriormente, la ganancia de amplificación se obtiene a partir de la ecuación 3.3:

$$V_0(t) = -\frac{R_2}{R_1} V_i(t) \quad (3.3)$$

Donde los valores de RV1 y R3 corresponden a R2 y R1 respectivamente. La resistencia R3 se fija en un valor de 10K Ω . Para el caso de RV1 se utilizó un potenciómetro de 1M Ω , para variar la ganancia del circuito de pre-amplificación, en teoría se podría amplificar hasta 100 veces la señal de ingreso. Debido a que existen diferentes tipos de voz y diferentes intensidades de la misma, se encontró necesario el hecho de ocupar el potenciómetro y así ajustar la sensibilidad dependiendo del tipo

de voz, es decir si existen voces con un nivel de intensidad muy bajo lo que se haría es aumentar el nivel de amplificación, por el contrario, en voces con alta intensidad, se disminuirá el nivel de amplificación debido a que existe una saturación en la amplificación.

Las resistencias R4 y R7 en conjunto con el capacitor C2 cumplen la función de establecer un *offset*, es decir, que la señal esté por encima del valor de referencia, ya que es necesario contar con la señal completa para analizarla y no perder los valores negativos de la misma.

La configuración del LM358 está establecido en amplificador inversor, es decir, la señal se invierte al momento de realizar la amplificación, por esta razón, en la segunda etapa del amplificador es necesario una configuración de seguidor con la finalidad de invertir la señal nuevamente. En la figura 3.7, se muestra el circuito respectivo.

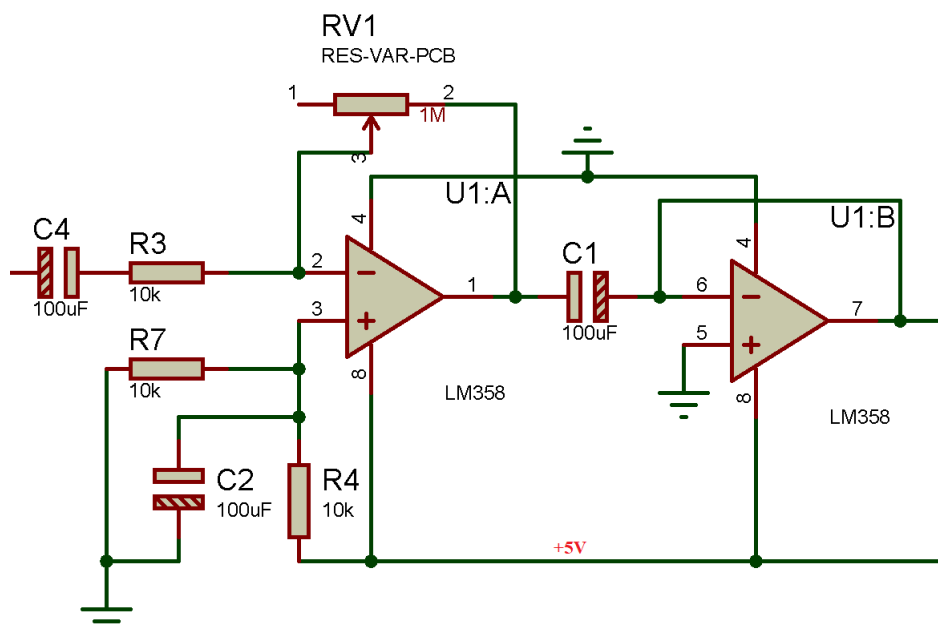


Figura 3.7: Circuito de pre-amplificación. Autoría propia.

3.3.3. Conexión de los demás componentes y circuito final

De la etapa de pre-amplificación se obtiene la señal que ingresará al módulo MAX4466, la cual, alcanza una respuesta en frecuencia de 20 Hz a 20 KHz . Además, cuenta con gran robustez al ruido ocasionado por el voltaje de alimentación. Este módulo funciona para sensores sensibles al sonido y cuenta con un potenciómetro para ajustar la ganancia de 25x a 125x. Por otro lado, cabe mencionar que el módulo tiene incorporado un micrófono Electret. No obstante, resulta importante mencionar que en este caso se optó por retirar dicho micrófono con el objetivo de acoplar la señal del acelerómetro

amplificada. La salida del MAX4466 va al pin VP del ESP32, esto se explicará más adelante como parte del código desarrollado. Para conectar el lector microSD al ESP32 se usaron los canales de comunicación Serial Peripheral Interfaces (SPIs) como se observa en la tabla 3.2 y figura 3.8.

Tabla 3.2: Pines de conexión del lector microSD al ESP32.

Pin lector microSD	Pin ESP32
CS	5
SCK	18
MOSI	23
MISO	19
VCC	5V
GND	GND

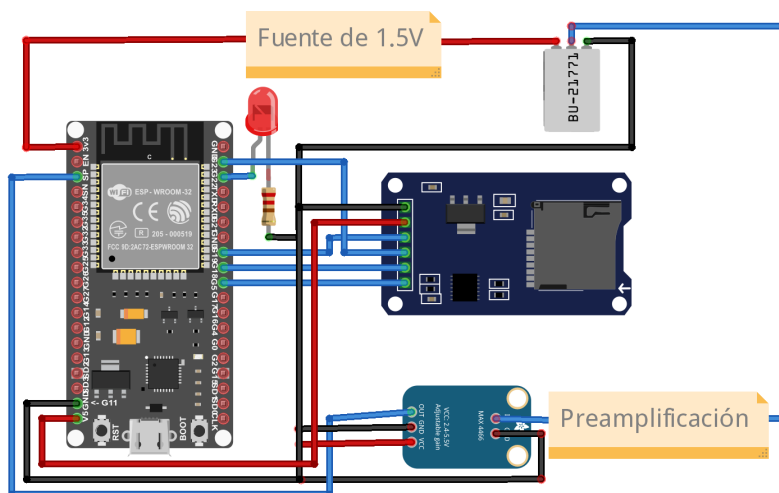


Figura 3.8: Circuito final. Autoría propia.

3.3.4. Diseño PCB del circuito final

El circuito esquemático mostrado anteriormente fue implementado en Proteus con el objetivo de aprovechar la herramienta de diseño Printed Circuit Boards (PCBs). Para el ESP32, lector micro SD y MAX4466 se utilizaron zócalos en lugar del componente como tal, de manera, que se los pueda colocar y retirar cuando se requiera. Se creó de igual manera una bornera de 3 pines para colocar el acelerómetro. Esto se puede visualizar de mejor manera en la figura 3.9, que muestra el diseño del circuito.

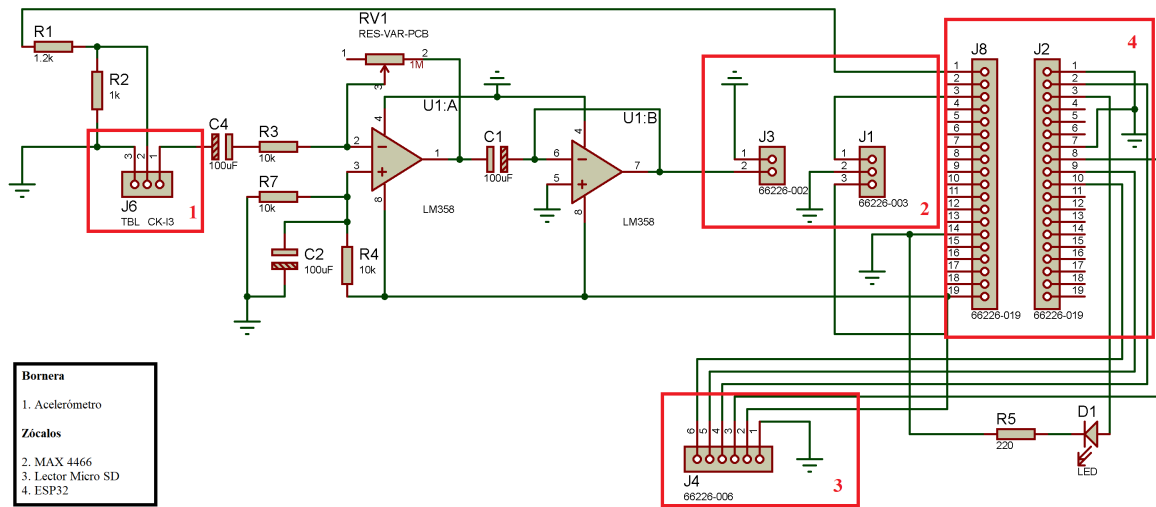


Figura 3.9: Circuito final. Autoría propia.

El diseño PCBs final se puede visualizar en la figura 3.10, como se observa, se ha colocado los elementos de tal manera que las pistas no tengan un amplio recorrido. El circuito PCBs está configurado en doble capa la frontal (*top*) y la trasera (*bottom*). Con la finalidad de evitar el ruido en la placa se ha creado también un plano de tierra.

En la figura 3.11, se observa el diseño en 3D de la placa, este diseño da como resultado una visión general de como quedaría la placa física real.

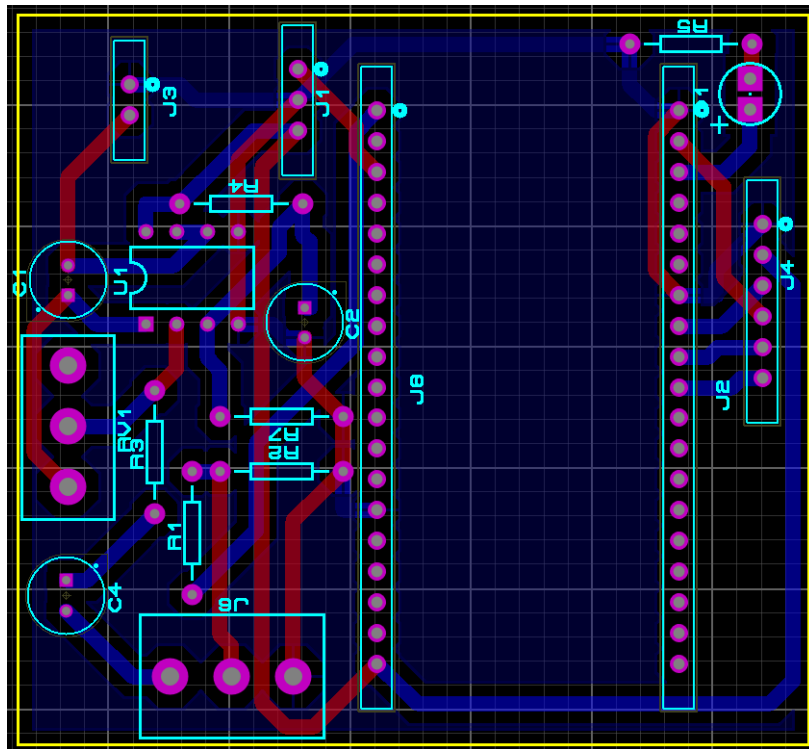


Figura 3.10: Circuito final en PCB. Autoría propia.

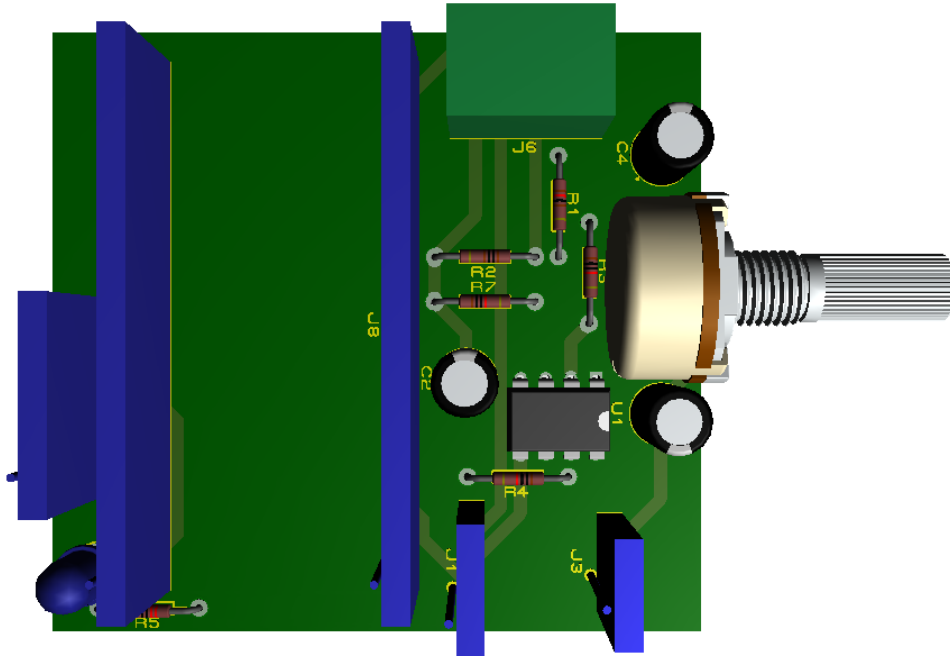


Figura 3.11: Circuito final en 3D. Autoría propia.

3.4. Script para ESP32

Se utilizó un *script* para la adquisición de datos a través del ESP32, el cual fue desarrollado en el entorno de desarrollo integrado de Arduino, en el anexo A se encuentra el código detallado. Se tomó como referencia los proyectos de *Github* [86] y [87]. El primero consta del montaje y código para la grabación de un audio utilizando el ESP32. Dentro de este proyecto se encuentran 5 archivos de programación, el *script* principal utiliza los 4 archivos restantes que son de código fuente (.cpp) y encabezado (.h). El *script* principal fue modificado de manera que se puedan realizar múltiples grabaciones de audio cada una con un nombre de archivo diferente. El segundo consta de un programa de gestión de archivos mediante servidor web. De esta forma, se tomaron ideas del primero y segundo *script* con el propósito de tener una interfaz de usuario con capacidad de descargar, grabar, borrar y visualizar la lista de archivos. En la figura 3.12, se puede visualizar la página principal de la interfaz de usuario.

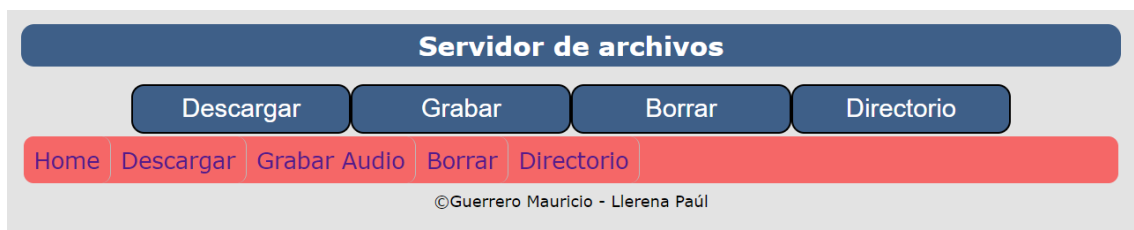


Figura 3.12: Interfaz del servidor web. Autoría propia.

3.5. Obtención de parámetros acústicos

3.5.1. Praat

Praat es un programa de libre distribución, de código abierto, multiplataforma y gratuito. Fue creado por Paul Boersma y David Weenink y, desarrollado en la Universidad de Amsterdam. Praat sirve principalmente para investigaciones en fonética y el análisis, síntesis y manipulación de señales del habla. Un aspecto relevante a considerar es que el programa es actualizado constantemente con mejoras desarrolladas por los autores, algunas de estas son sugeridas por los usuarios.

Al tratarse de un programa especializado en fonética, no cuenta con interfaces cómodas y simples para el usuario; se requiere un cierto adiestramiento para el manejo correcto del programa, además de un sustento teórico previo. Se recomienda, para la obtención de información más enfocada y especializada, revisar el menú de ayuda del mismo programa y participar de los foros dedicados a la que se puede acceder desde la página principal de Praat. En la misma página se encuentran todas las publicaciones relacionadas con los algoritmos implementados.

Entre las principales funciones de Praat se encuentra:

- **Análisis del habla:** El programa tiene la capacidad de procesar muestras de habla y adquirir información acústica como espectrogramas de banda ancha y estrecha, espectros, formantes, frecuencia fundamentales (F0s), intensidad, duración, entre otros.
- **Estadística:** Con el programa Praat, es posible aplicar métodos estadísticos (*análisis de escalamiento multidimensional, análisis de componentes principales y análisis discriminante*) a tablas con datos acústicos.
- **Scripts:** Con el lenguaje de programación de Praat, es posible automatizar tareas y evitar la repetición manual de diversos procedimientos, estas rutinas son creadas por el usuario para diferentes propósitos y fines.
- **Figuras:** Praat, ofrece una ventana de imágenes o *Praat Picture* que sirve para la rotulación y decoración de figuras para incluir en tesis, artículos, libros o presentaciones. Dicha herramienta cuenta con mucha flexibilidad ya que varía dimensiones, tipografía y colores de una manera fácil.

3.5.2. Obtención de parámetros a través de Praat

Se emplearon *scripts* con el fin de automatizar el análisis de la voz. Por otro lado, dicho análisis puede ser complementado con un análisis acústico gráfico que más adelante se expondrá. A continuación, se detallan las configuraciones utilizadas dentro de los *scripts* en Praat para la obtención de los parámetros.

3.5.2.1. Frecuencia fundamental

En primera instancia, es necesario la utilización del comando **Select Sound** [*File-name.ext*] para seleccionar el archivo de audio a analizar, dentro del comando se debe colocar el nombre del archivo de audio seguido de su extensión, la extensión por defecto es (*.wav*).

Para la obtención de la frecuencia fundamental se utiliza el comando **To Pitch...** [*Time step(s) Pitch floor(Hz) Pitch ceiling(Hz)*], donde:

- **Time step (s):** tiene un valor estándar de 0.0 y es el intervalo de medición (frame duration), en segundos (s).
- **Pitch floor (Hz):** es el límite inferior de la frecuencia un valor menor a 75Hz o valor estándar, no será reconocido.
- **Pitch ceiling (Hz):** es el límite superior de la frecuencia un valor mayor a 600Hz o valor estándar, no será reconocido.

Con la finalidad de obtener la media de la frecuencia fundamental en Hertz (Hz), se emplea el comando **Get mean...** [*t_i t_f unidad*], donde *t_i* y *t_f* es el tiempo inicial y final del audio correspondiente, es decir, es el intervalo de tiempo del audio a analizar, cabe señalar que si se coloca un cero en ambos parámetros la función analizará todo el rango del tiempo del audio, en *unidad* debe ser colocada la unidad de medida del parámetro solicitado, es decir, en este caso de la frecuencia fundamental [*Hertz*].

En base al comando anterior y sus parámetros, ahora para obtener la desviación estándar de la frecuencia fundamental se emplea el comando **Get standard deviation...** [*t_i t_f unidad*].

Posteriormente, para obtener el valor mínimo de la frecuencia fundamental en un rango específico de tiempo se utiliza el comando **Get minimum...** [*t_i t_f unidad interpolation*], donde el parámetro *interpolation* es el método de interpolación aplicado en el comando. El método estándar es el parabólico (*Parabolic*).

De la misma manera es posible obtener el valor máximo de la frecuencia fundamental en un rango específico de tiempo, para ello, se utiliza el comando ***Get maximum...*** [*t_i t_f unidad interpolation*].

A continuación, se muestra un ejemplo de la implementación de los comandos en el software Praat.

```
Select Sound [Filename.wav]
To Pitch... 0.0 75 600
Get mean... 0 0 Hertz
Get standard deviation... 0 0 Hertz
Get minimum... 0 0 Hertz Parabolic
Get maximum... 0 0 Hertz Parabolic
```

3.5.2.2. Intensidad

De igual manera, es necesario la utilización del comando ***Select Sound*** [*Filename.ext*] para seleccionar el archivo de audio a analizar. Para obtener la intensidad se utiliza el comando ***To Intensity...*** [*Minimum pitch(Hz) Time step(s)*], donde:

- **Minimum pitch (Hz):** es la frecuencia de periodicidad mínima de la señal.
- **Time step (s):** es el paso de tiempo del contorno de intensidad resultante. Si este parámetro es establecido en cero, el paso de tiempo se calcula como un cuarto de la longitud efectiva de la ventana, es decir, como $0.8/(Minimum\ pitch)$.

A continuación, se presenta un ejemplo de la aplicación de los comandos en el software Praat.

```
Select Sound [Filename.wav]
To Intensity... 75 0.001
```

3.5.2.3. Tiempo de fonación

Para obtener de este parámetro, se debe tener en cuenta que es necesario descartar el ataque y la filatura además de los silencios en la fonación, para ello se ha utilizado los dos parámetros anteriores (Frecuencia fundamental e Intensidad), posteriormente se analizan los dos parámetros simultáneamente en el tiempo. Para la obtención correcta del tiempo de fonación, es necesario descartar el tiempo en el que uno de los parámetros (Frecuencia fundamental o Intensidad) no esté definido o no tenga valor, para ello, se ha creado una lista en la que consta el tiempo, la frecuencia y la intensidad, luego se

analiza y se descartan los intervalos de tiempo en donde no existen valores de dichas magnitudes.

Para tal objetivo se ha empleado el comando *Select Sound [Filename.ext]* para seleccionar el archivo de audio a analizar. Posteriormente, se extrae el tiempo de inicio y final con los comandos *Get start time* y *Get end time* respectivamente. Finalmente, se separa el tiempo total en partes de 0.01s esto con el objetivo de extraer los intervalos en los cuales la frecuencia o la intensidad están indefinidos (*undefined*).

A continuación, se muestra un ejemplo para la obtención del tiempo de fonación en el software Praat.

```
select Sound 'name1$'  
sound = selected ("Sound")  
tmin = Get start time  
tmax = Get end time  
To Pitch: 0.001, 75, 300  
Rename: "pitch"  
selectObject: sound  
To Intensity: 75, 0.001  
Rename: "intensity"  
writeInfoLine: "Here_are_the_results:"  
sum = 0  
num = 0  
tiempo = 0  
for i to (tmax-tmin)/0.01  
    time = tmin + i * 0.01  
    selectObject: "Pitch_pitch"  
    pitch = Get value at time: time, "Hertz", "linear"  
    selectObject: "Intensity_intensity"  
    intensity = Get value at time: time, "cubic"  
    appendInfoLine: fixed$ (time, 2), "_", fixed$ (pitch, 3),  
        "_", fixed$ (intensity, 3)  
    if intensity = undefined  
        num=num+1  
    else  
        sum=sum+intensity  
    endif  
  
    if intensity = undefined or pitch = undefined  
        tiempo=tiempo+1
```

```
        endif
    endfor

    int_prom1 = sum/(((tmax-tmin)/0.01)-num)
    fon1 = time - (tiempo*0.01)
```

3.5.2.4. Jitter

Nuevamente, es necesario la utilización del comando **Select Sound** [*Filename.ext*] para seleccionar el archivo de audio a analizar.

También, es necesario emplear el comando **To PointProcess (periodic, cc)** [F_i F_f], donde F_i , es la frecuencia inicial y F_f la frecuencia final a analizar, este comando da como resultado las secciones periódicas que el programa interpreta como la secuencia de cierres glóticos en la vibración de las cuerdas vocales.

Para obtener el *Jitter local*, es necesario emplear el comando **Get jitter (local)...** [*Time range (min) (s) Time range (max) (s) Period floor(s) Period ceiling(s) Maximum period factor*], donde:

- **Time range (min) (s)**: el tiempo inicial, en segundos (s).
- **Time range (max) (s)**: el tiempo final, en segundos (s).
- **Period floor (s)**: el intervalo más corto posible que será utilizado en el cálculo de la fluctuación de fase, en segundos. Esta configuración por lo general es muy pequeña, normalmente de 0.1 ms.
- **Period ceiling (s)**: el intervalo máximo posible que será utilizado en el cálculo de la fluctuación de fase, en segundos. Esta configuración, por lo general, trabaja con una frecuencia mínima de periodicidad de 50Hz, por lo que, se establece intervalos cada 0.02s. En este contexto, resulta importante mencionar que intervalos más grandes podrían considerarse como tramos sin voz y por ende serán ignorados en el cálculo.
- **Maximum period factor**: la mayor diferencia posible entre intervalos consecutivos que será utilizado en el cálculo de la fluctuación de fase.

A continuación, se muestra un ejemplo de la aplicación de los comandos en el software Praat.

```
Select Sound [Filename.wav1]
To PointProcess (periodic, cc)... 75 500
Get jitter (local)... 0 0 0.0001 0.02 1.3
```


3.5.2.5. Shimmer

Siguiendo la misma estructura y sintaxis que el parámetro anterior (*Jitter*), Es necesario emplear el comando **plus PointProcess [Filename]** . Luego, se utiliza el comando **Get shimmer (local)... [Time range(min)(s) Time range(max)(s) Period floor(s) Period ceiling(s) Maximum period factor]**, donde los parámetros son similares al del *Jitter*, a continuación se presenta un ejemplo de aplicación del comando.

```
Select Sound [Filename.wav]
To PointProcess (periodic, cc)... 75 500
Get shimmer (local)... 0 0 0.0001 0.02 1.3
```

3.5.2.6. HNR (Harmonic-to-Noise Ratio)

Como en los anteriores parámetros primero es necesario emplear el comando **Select Sound [Filename.ext]** para seleccionar el archivo de audio a analizar. Para obtener el HNRs, es necesario emplear el comando **To Harmonicity (cc)... [Time step (s) Minimum pitch (Hz) Silence threshold Number of periods per window]**, donde:

- **Time step (s):** tiene un valor estándar de 0.01s y es el intervalo de medición (*frame duration*), en segundos (s).
- **Minimum pitch (Hz):** tiene un valor estándar de 75 Hz y determina la longitud de la ventana de análisis.
- **Silence threshold:** aquellas tramas que no contienen amplitudes por encima de este umbral (en relación con la amplitud máxima global) se consideran silenciosas.
- **Number of periods per window:** tiene un valor estándar de 4.5, este valor es el óptimo para el habla, ya que garantiza que los valores HNRs de hasta 37 dB se detecten de manera confiable.

Este comando detecta la periodicidad acústica sobre la base de un análisis de correlación cruzada directa. También, se han aplicado los comandos **Get mean... [t_i t_f unidad]** y **Get standard deviation... [t_i t_f unidad]** Para obtener el promedio y la desviación estándar del HNRs respectivamente. A continuación, se presenta un ejemplo de la aplicación de los comandos anteriores en el software Praat.

```
Select Sound [Filename.wav]
To Harmonicity (cc)... 0.01 75 0.1 1
Get mean... 0 0
Get standard deviation... 0 0
```

En la Tabla 3.3, se presenta un resumen de cada uno de los parámetros obtenidos con el software Praat junto con la unidad de medida correspondiente y una breve descripción del parámetro.

Tabla 3.3: Parámetros obtenidos en el software Praat.

Parámetro	Unidad	Descripción
Frecuencia fundamental (F_0)	Hz	Cantidad de veces que vibran las cuerdas vocales en un segundo.
Intensidad (I)	dB_{SPL}	Asociada directamente con la amplitud de la frecuencia, y depende primordialmente de la presión sonora.
Tiempo de fonación (T)	s	Período de tiempo durante el cual las cuerdas vocales ejercen esfuerzo o están vibrando.
Variabilidad de la F_0 (Jitter local)	%	Diferencia absoluta promedio entre períodos consecutivos, dividida por el período promedio.
Perturbación en amplitud (Shimmer local)	%	Diferencia absoluta promedio entre las amplitudes de períodos consecutivos, dividida por la amplitud promedio.
Relación armónico-ruido (HNRs)	dB	Medición de la relación entre la energía acústica de los armónicos y la energía del ruido agregado del tracto vocal durante la fonación.

3.5.3. Obtención de imágenes acústicas a través de Praat

Además, de la variedad de parámetros que se pueden obtener en el *software* Praat a través de los scripts, también es posible observar características acústicas de manera gráfica ya que el *software* permite obtener el oscilograma y espectrograma de un archivo de audio.

En la figura 3.13, se puede observar la obtención del oscilograma en la parte superior y el espectrograma en la parte inferior, además, se adquiere información de alguno de los parámetros como son la intensidad y la frecuencia fundamental estos se pueden observar sobrepuestos en el espectrograma la frecuencia fundamental de color azul y la intensidad de color amarillo, las escalas de dichos parámetros se encuentran en el margen derecho del espectrograma.

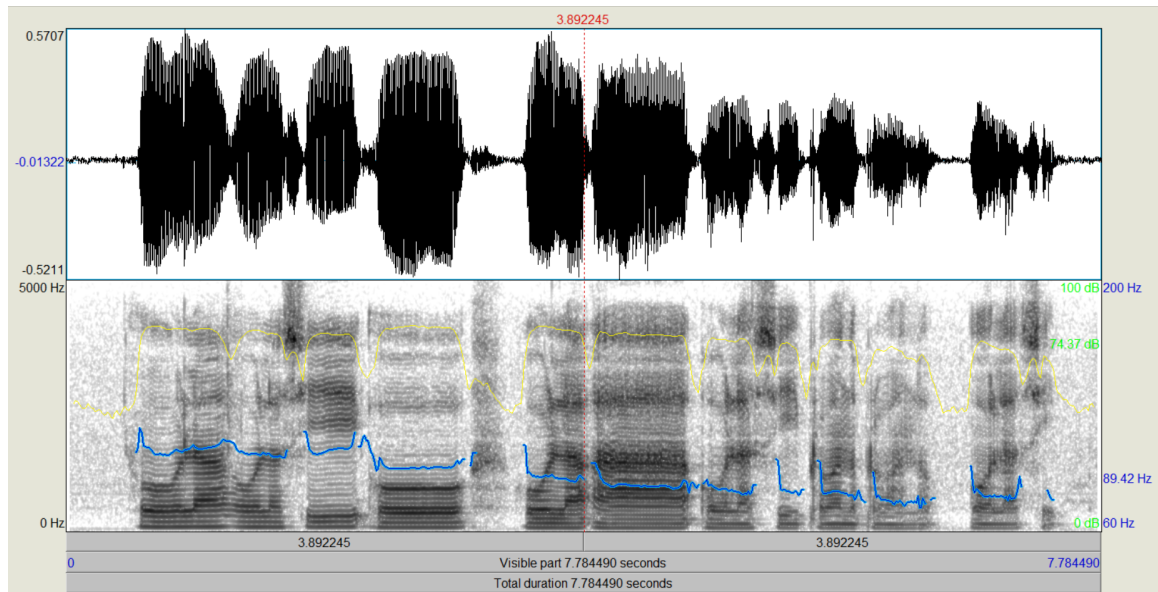


Figura 3.13: Oscilograma y espectrograma en Praat. Autoría propia.

Es importante destacar, que el programa Praat obtiene las imágenes acústicas en rangos de período de diez segundos, es decir, si el archivo de audio dura más de diez segundos se deberá evaluar por tramos, tomando en consideración la restricción mencionada.

4. Resultados

En este capítulo, se presentan los resultados obtenidos luego de haber realizado las pruebas respectivas con pacientes sanos y patológicos. Todo el proceso fue realizado conforme a la bibliografía encontrada.

4.1. Proceso de adquisición de la muestra

La población de la cual se recogen los datos, consta de 21 individuos post COVID-19 de los cuales se evaluarán cada uno de los parámetros y se obtendrán resultados contrastando el estudio con el informe de un fonoaudiólogo. Según [88] y [89], dentro del protocolo básico de valoración de la voz, se inicia con una grabación. Una grabación estándar consta de una vocal sostenida (típicamente la vocal /a/) y una frase sencilla.

4.1.1. Metodología de toma de muestras por parte de especialistas de la voz

Cada uno de los individuos fueron diagnosticados previamente en un aula insonorizada con equipos especializados y con supervisión de profesionales en el área de la voz. A continuación se describe de manera general las recomendaciones metodológicas para la grabación de la voz.

- Aula o habitación insonorizada (con ruido ambiente inferior a 40dB [90, 91]).
- Para una correcta y precisa medición de intensidad vocal es relevante la distancia del micrófono a la boca dicha distancia fue de 10cm.
- Los requisitos mínimos que el sistema de grabación debe tener, es una frecuencia de muestreo de al menos 20kHz, para evitar errores en el procesamiento de audio.
- La sensibilidad del micrófono debe ser de al menos 60dB, también considerando un ángulo de 45° a 90° para evitar ruidos provocados por la respiración.

4.1.2. Reactivos utilizados

Como parte de la metodología de la toma de muestras por parte de los especialistas de la voz, se encuentran los reactivos más importantes los cuales se describen en la tabla 4.1.

Tabla 4.1: Reactivos utilizados por parte de especialistas de la voz.

Reactivo	Uso clínico
Historia clínica	Recolección de información personal, antecedentes personales de salud y hábitos que permiten hacer predicciones del estado de salud del paciente.
GRABS	Parámetros para medir la cualidad vocal percibida. Propuesto por Hirano, 1981 desarrollado en la sociedad japonesa de logopedia y foniatría. Cuantifica del 0 al 4, según sus siglas: Grado de disfonía, Rugosidad, Astenia, Soplosidad y Tensión.
Cape - V	Parámetros para medir la cualidad vocal percibida. Propuesta por la Asociación Americana de Habla y Lenguaje. Mide parámetros de: severidad global, ronquera, escape aéreo, tensión, tono e intensidad.
VHI- 10	Descrito por Jacobson et al, es una versión simplificada del VHI-30, que mide parámetros físicos, emocionales y funcionales de la autopercepción de la disfonía y la calidad de vida del paciente que la padece.
Praat	Software para realizar análisis acústico de la voz, mide parámetros de: frecuencia fundamental, intensidad, medidas de perturbación e índice de armonía y ruido. Su utilidad radica en el monitoreo de la eficacia antes y después del tratamiento; además complementa al diagnóstico de imagen realizado por el médico laringólogo especialista.

Las muestras de voz obtenidas para el análisis fueron tomadas en un estudio de grabación con cámara insonorizada, con un micrófono con condensador de marca Beta4-080. Realizando las siguientes tareas fonatorias.

- Vocales sostenidas /a/ , /i/, después de emitir la frase “<esta, es una cabina>” y “<esta es una cabi>”.
- Frases del formador Cape-V, diseñadas para que aparezcan todas las vocales, con ataque vocal suave y duro, sonidos nasales y consonantes plosivas sordas.

- Habla espontánea.

4.2. Prototipo del dispositivo final

En la figura 4.1, 4.2 y 4.3, se muestra el prototipo final completo con el montaje de cada uno de sus elementos.



Figura 4.1: Prototipo del dispositivo final - vista superior. Autoría propia.



Figura 4.2: Prototipo del dispositivo final - vista frontal. Autoría propia.



Figura 4.3: Prototipo del dispositivo final - vista trasera. Autoría propia.

Por otra parte, en la figura 4.4, se muestra el PCBs sin los componentes soldados. La parte izquierda corresponde a la parte superior, mientras que la derecha corresponde a la parte inferior.

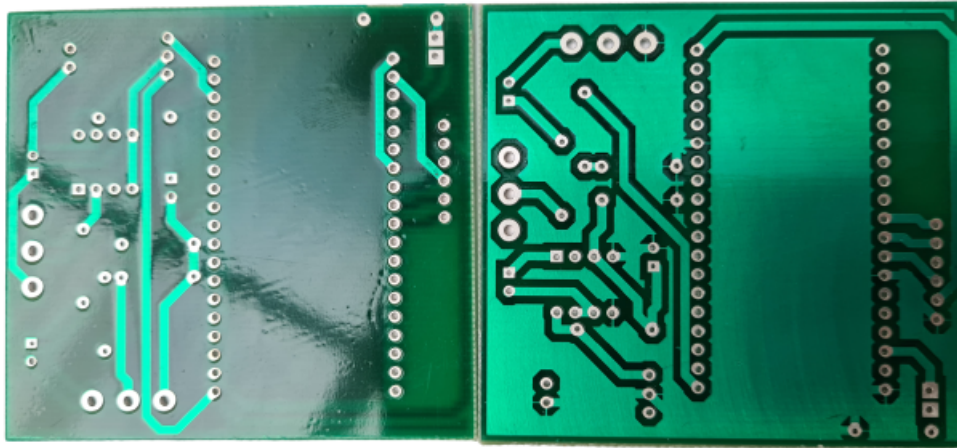


Figura 4.4: Placas impresas sin componentes soldados. Autoría propia.

4.3. Consumo energético del circuito

Conocer el consumo energético del circuito es de gran importancia si lo que se quiere es reducirlo. Esto es un punto clave para conseguir que el dispositivo sea portátil. Según los *datasheet* de cada elemento que se utilizó en el circuito de la figura 4.1, se realizó un estimado del consumo energético.

Tabla 4.2: Consumo energético de los elementos del circuito.

	Consumo energético typ. (mA)	Consumo energético max. (mA)
ESP32	160	260
BU-21771	0.005	0.005
LM358	1	2
Lector SD	80	200
LED	20	20
MAX4466	0.0024	0.0048
Total	261.05	482

Como se puede observar en la tabla 4.2, el consumo típico del dispositivo sería de **261.05mA**, mientras que el consumo máximo sería de **482mA**. En base a esto, si se tiene una batería de litio de 1700mAh, la duración de la batería para el consumo típico de energía, sería de 6 horas y media. Mientras que para el consumo de energía máximo,

la duración de la batería sería de 3 horas y media.

En este caso, para posibilitar la monitorización de los parámetros de la voz durante un tiempo prudencial de forma ininterrumpida, se alimentó el mismo mediante un cable micro USB que va conectado al ESP32.

4.4. Metodología de toma de muestras con el dispositivo electrónico

4.4.1. Posicionamiento del acelerómetro

Para una correcta detección y registro de las vibraciones de las cuerdas vocales, es necesario un correcto posicionamiento del acelerómetro, para ello se toma como referencia lo descrito en [16]. En la figura 4.5, con base en [16], se muestra cada uno de los puntos en donde es posible colocar el acelerómetro con la finalidad de detectar la mayor amplitud de las vibraciones de las cuerdas vocales.

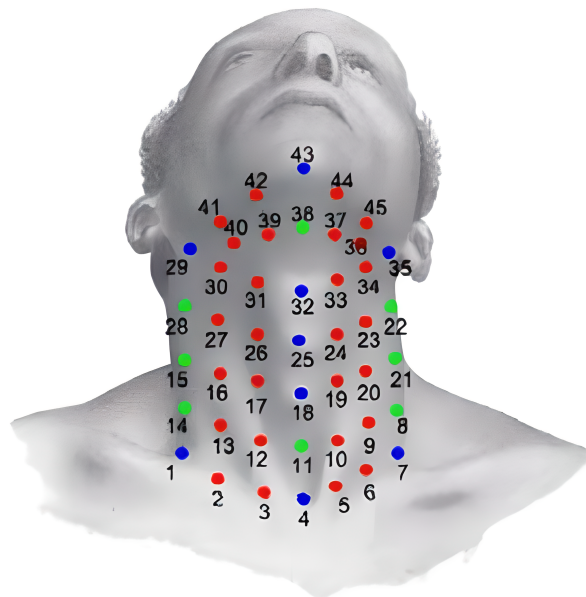


Figura 4.5: Conjunto de puntos donde es posible colocar el acelerómetro. Tomado de [16].

Ahora, en la figura 4.6, se detalla en una escala de colores los puntos con mayor o menor amplitud, siendo de importancia para el estudio el punto con mayor amplitud. En conclusión el acelerómetro será colocado en la parte inferior izquierda de la superficie del cuello, es decir, en el punto número 24 de la figura 4.5.

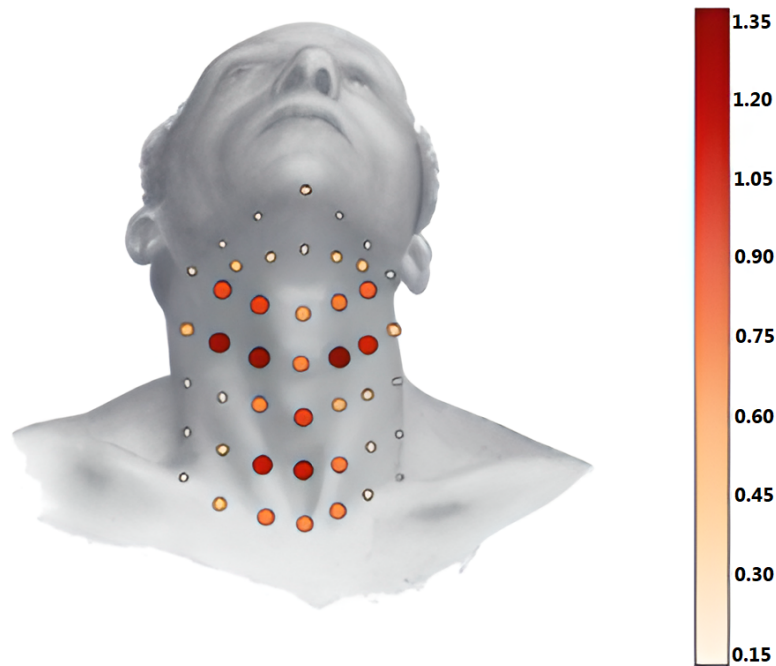


Figura 4.6: Amplitud de cada punto. Tomado de [16].

4.4.2. Protocolo de grabación

Una vez colocado de manera correcta el acelerómetro como se mencionó anteriormente, se seguirá una metodología sencilla de grabación. En primer lugar el individuo se debe colocar en una postura relajada, en este caso el individuo se sienta y mantiene la posición del cuello erguida, después de esto se captura una vocal /a/ sostenida en tono comfortable.

Es importante mencionar que, el tiempo de fonación se tomará como un parámetro aislado, debido a que si se evalúa en conjunto con los demás parámetros no se obtendrán resultados precisos, debido a la naturaleza misma del parámetro. Entonces, se capturan 3 muestras con una vocal /a/ sostenida durante el tiempo que el individuo pueda realizar la fonación correctamente. Posteriormente, se realiza un promedio de las 3 muestras con la finalidad de obtener resultados de mayor precisión y veracidad.

Por otro lado, para obtener los demás parámetros, de igual forma se ejecutarán 3 grabaciones con una vocal /a/ sostenida, con la diferencia de capturar solamente 10 segundos. En dichas muestras se deben seleccionar las regiones estables o continuas, despreciando el ataque y la filatura de la fonación.

4.5. Proceso de grabación y almacenamiento de los archivos de audio a través de la página web

En la figura 3.12, se tiene el menú de inicio de la página web. Para conectarse a dicha página es necesario estar en la misma red a la cual está conectado el ESP32. Al iniciarlo, este busca alguna red de la cual tenga las credenciales, de no ser el caso, el mismo actúa como Access Point (AP) para realizar una configuración WiFi. Por ejemplo en la figura 4.7 se muestran las redes disponibles, incluida la del ESP32 con el nombre de *ESP32-Ucuenca-WiFi*.

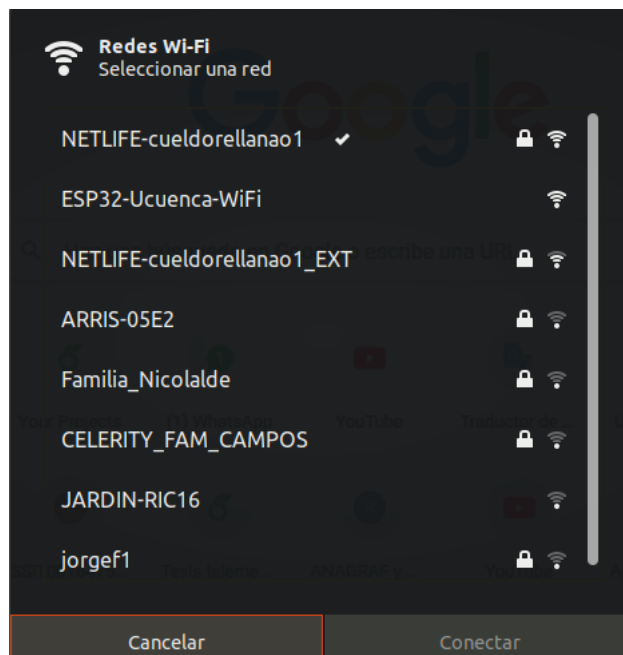
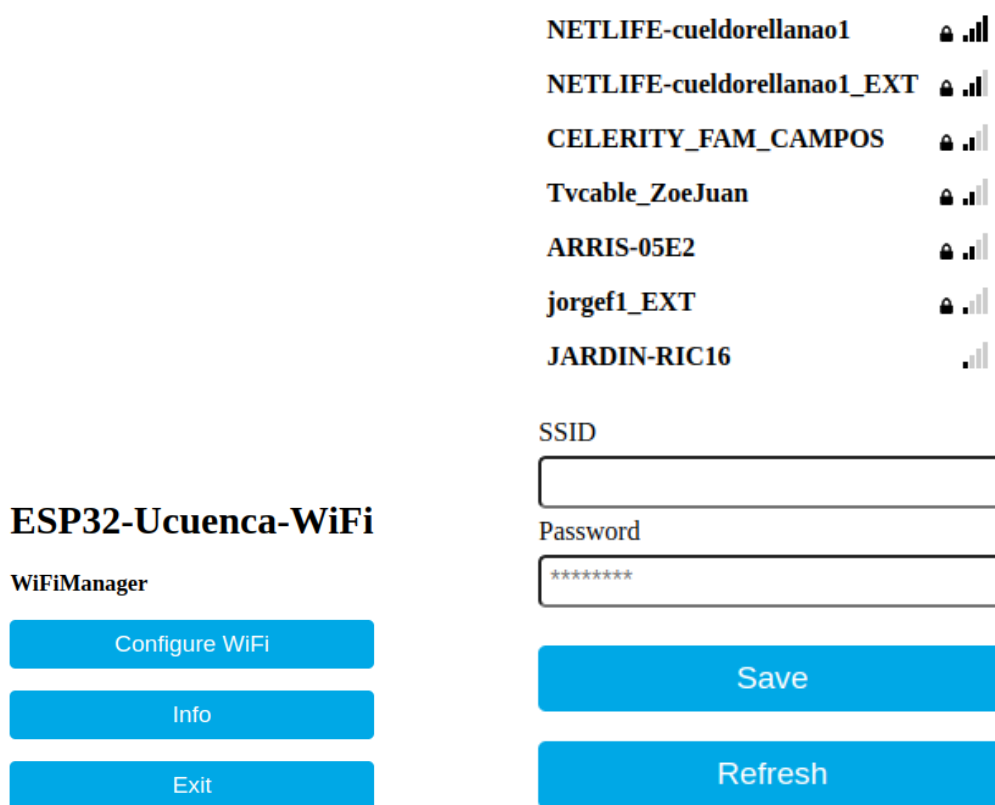


Figura 4.7: Redes WiFi disponibles, incluida la del ESP32. Autoría propia.

Para realizar la configuración se debe conectar a la red *ESP32-Ucuenca-WiFi* y una vez hecho esto, en un *browser* se ingresa la IP *192.168.4.1*. Aparecerá una página web, véase la figura 4.8a. Se selecciona la opción *Configure WiFi* y aparecerá lo que se muestra en la figura 4.8b, en esta página se tienen las redes WiFi detectadas por el ESP32 y recuadros para añadir el nombre de la red (Service Set Identifier (SSID)) y la contraseña (*Password*). Al colocar los datos respectivos de la red, se debe presionar en el botón *Save* y el módulo se reiniciará y tratará de conectarse a la red configurada. De no ser el caso, el ESP32 volverá al modo AP y el proceso se debe repetir.

Una vez conectado a la red, dentro de un navegador web se ingresa la dirección IP configurada en el código fuente la cual es *192.168.1.150*. El programa tiene las siguientes opciones:



(a) Página principal para configuración de la red. (b) Interfaz de ingreso de nombre y contraseña de la red.

Figura 4.8: Conexión a una red WiFi mediante el ESP32. Autoría propia.

4.5.1. Grabar

En esta opción se puede realizar la grabación de un archivo de audio en formato *.wav*. El nombre del archivo por defecto va a empezar con la palabra *sound*, seguida del número o letra que se ingresa en el cuadro de texto. En la figura 4.9, se tiene un número 1, por lo tanto, el archivo resultante tendrá el nombre de *sound1* con la extensión *.wav*. Seguido del número 1 se tiene una coma y el número 10, esto indica el tiempo en segundos de duración del archivo de audio. Para diferenciar entre el nombre y tiempo de grabación se utiliza el delimitador ”,”. Es importante que para empezar la grabación se debe presionar la tecla *ENTER* en el cuadro de texto. Si se utiliza un nombre de archivo que ya existe en el almacenamiento, será reemplazado por el nuevo archivo. En la figura 4.10, se muestra el mensaje de confirmación que indica que el archivo de audio se ha guardado correctamente en la tarjeta micro SD.

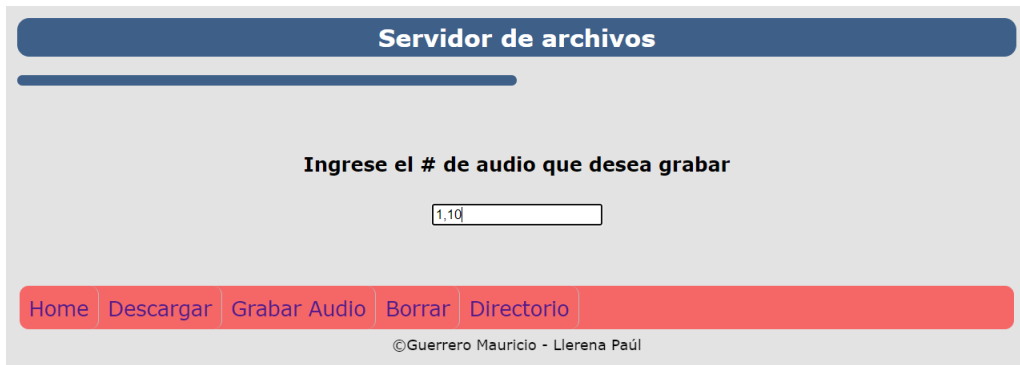


Figura 4.9: Menú para grabar el archivo. Autoría propia.



Figura 4.10: Mensaje que indica que el archivo se ha grabado correctamente. Autoría propia.

4.5.2. Descargar

Para descargar un archivo en específico, se tiene el cuadro de texto que se muestra en la figura 4.11. En el mismo se coloca el nombre del archivo con su extensión, en este caso el archivo *sound1.wav*. De igual manera para comenzar la descarga se presiona la tecla *ENTER* en el cuadro de texto. En la figura 4.12, se muestra la descarga en curso del archivo.



Figura 4.11: Menú para descargar el archivo. Autoría propia. Autoría propia.

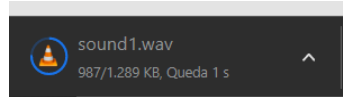


Figura 4.12: Descarga del archivo. Autoría propia.

4.5.3. Borrar

En esta opción se puede borrar un archivo específico, para ello en el cuadro de texto de la figura 4.13, se coloca el nombre del archivo junto con su extensión, en este caso el archivo *sound5.wav*. Para accionar el borrado del archivo de igual manera se presiona la tecla *ENTER* en el cuadro de texto. En la figura 4.14, se muestra un mensaje que indica que el archivo se ha borrado correctamente.



Figura 4.13: Menú para borrar el archivo. Autoría propia.



Figura 4.14: Mensaje que indica que el archivo se ha borrado correctamente. Autoría propia.

4.5.4. Directorio

Si se requiere verificar los archivos que se encuentran en la tarjeta micro SD se puede utilizar la opción de directorio. En la misma se despliega una lista de todos los archivos disponibles, junto con un indicador de archivo o directorio y por último el

tamaño del archivo. En este caso en la figura 4.15, se tiene una parte de la lista de archivos.

Servidor de archivos		
Contenido de la tarjeta SD		
Name/Type	Type File/Dir	File Size
Dir	/System Volume Information	
/System Volume Information/IndexerVolumeGuid	File	76 B
/System Volume Information/WPSettings.dat	File	12 B
/AUTORUN.INF	File	16 B
/test.txt	File	1.000 MB
/foo.txt	File	13 B
/sound3.wav	File	859.418 KB
/sound4.wav	File	859.418 KB
/Hola.txt	File	15 B
/sound6.wav	File	859.418 KB
/sound7.wav	File	859.418 KB
/sound8.wav	File	859.418 KB

Figura 4.15: Lista de archivos disponibles en la tarjeta micro SD. Autoría propia.

4.5.5. Posibles errores

Si la tarjeta micro SD no se encuentra insertada, al momento de realizar alguna acción, se mostrará el mensaje que se muestra en al figura 4.16. Basta con insertar la tarjeta micro SD y reiniciar el dispositivo para solucionar este inconveniente.

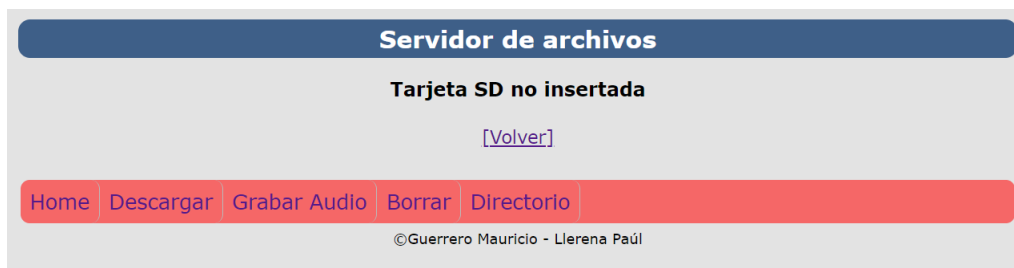


Figura 4.16: Error al no insertar la tarjeta micro SD. Autoría propia.

4.6. Adquisición de los parámetros a través de Praat

El proceso de grabación, como se ha mencionado anteriormente, se realiza con una vocal sostenida /a/, este proceso se lo realiza tres veces con una duración de 20 segundos para cada individuo. Como se muestra en la figura 4.17, en la parte inferior se indica la duración de la grabación total que es de alrededor de 20 segundos. Debido a que de manera gráfica Praat solo puede analizar ventanas de un máximo de 10 segundos se muestra dicha sección.

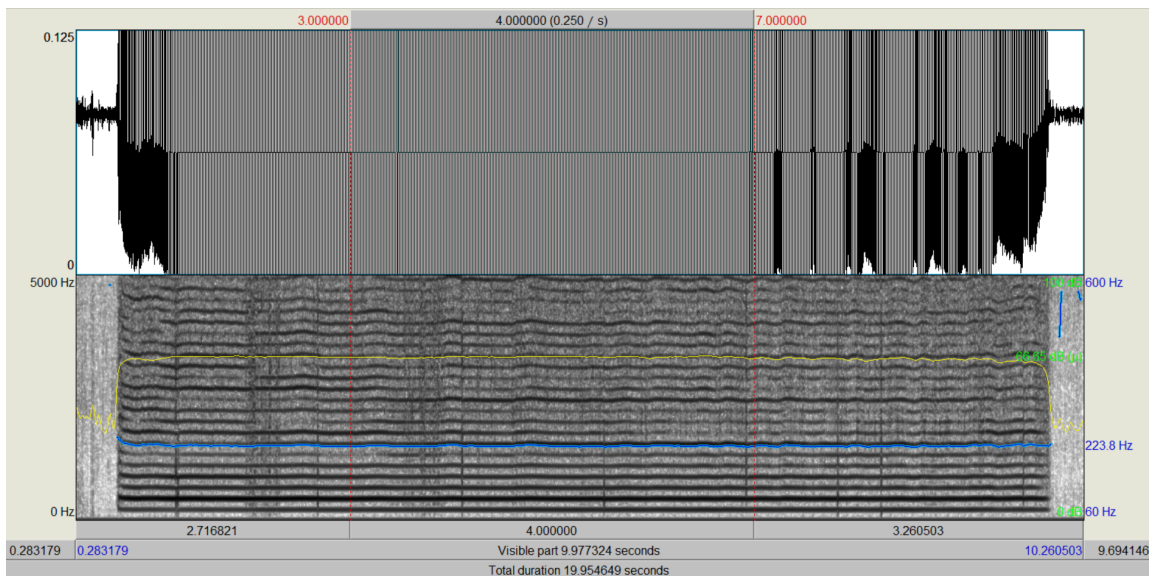


Figura 4.17: Ventana de 10 segundos y sub-ventana de 4 segundos en Praat. Autoría propia.

Se puede notar claramente que el rango dinámico de los parámetros a evaluar dura pocos segundos por lo que se ha decidido crear una sub-ventana para evaluar correctamente los parámetros, es decir, el parámetro de **tiempo de fonación** se lo realizará de manera independiente, evaluando toda la duración de la grabación sin tomar en cuenta el ataque, la filatura y los instantes en que no exista ni frecuencia ni intensidad constante, como se explicó de manera detallada anteriormente. Por otro lado, para los demás parámetros será necesario acortar el audio en un rango en la que los parámetros sean estables para que los resultados sean válidos, en este caso se ha escogido el intervalo de duración de 3 a 7 segundos por lo que esta nueva ventana tendrá una duración de 4 segundos, en la figura 4.17, se muestra dicha sección de tiempo entre líneas paralelas rojas punteadas. Ahora, para automatizar esto, es necesario añadir una sección dentro del script de Praat, dicho script se muestra a continuación.


```
sound = Read from file: file_name$
selectObject: sound
Edit
editor: sound
    Select: 3, 7
    cutsound1 = Extract selected sound (preserve times)
Close
Rename: "cutsound1"
```

Donde, al inicio se lee el archivo de audio, en este caso el archivo de audio completo de 20 segundos, luego con el comando **Edit** se selecciona el audio a editar y con el comando **Select: $[T_i],[T_f]$** , se limita el intervalo de la sub-ventana en este caso, se escogerá una ventana de 4 segundos, desde $T_i = 3$, hasta $T_f = 7$. Después se crea un nuevo objeto con el audio recortado con la condición de preservar los tiempos, se cierra el editor y se renombra el nuevo archivo de audio recortado en este caso como **cutsound1**. Esto se lo realiza con cada uno de los 3 archivos de audio grabados para promediar cada uno de los parámetros adquiridos para un resultado mucho más preciso y confiable.

Una vez ejecutado el *script* que se ha desarrollado, ver anexo B, los resultados se mostrarán en una ventana emergente del mismo programa Praat, en la que constará toda la información detallada acerca de los parámetros, como se puede observar en el recuadro inferior.

```
##### Parámetros #####
La frecuencia fundamental es: 224.46225450207024 Hz
La intensidad es: 63.46852243649075 dB
El tiempo de fonación es: 8.479999999999999 s
El jitter local es: 0.712732680818963 %
El shimmer local es: 1.4523527201821202 %
El HNR es: 20.69853998954751 dB

##### Estadísticas #####
La desviación estándar de la frecuencia fundamental es: 1.7778273020799962
La desviación estándar de la intensidad es: 3.767652005391048
El valor mínimo de la frecuencia fundamental es: 220.76558666168003 Hz
El valor máximo de la frecuencia fundamental es: 230.5282908820983 Hz
El valor mínimo de la intensidad es: 56.468542794616745 dB
El valor máximo de la intensidad es: 67.38079505321052 dB
La desviación estándar del HNR es: 4.094908925161924
```

4.7. Obtención de los parámetros acústicos

En primer lugar, para comprobar la eficiencia del dispositivo, se ha realizado pruebas en las que cada individuo realiza las tareas fonatorias dentro de la cabina de grabación, cuya característica primordial es que tiene bajo ruido ambiente y un micrófono profesional que tiene una sensibilidad superior a 60 dB. Esto lo convierte en un ambiente idóneo para la adquisición de los parámetros. De manera simultánea, se ha colocado el dispositivo a cada individuo, para obtener una tabla comparativa de cada parámetro con el dispositivo electrónico y con el micrófono dentro de la cabina de grabación, tal como se observa en la figura 4.18.



Figura 4.18: Prueba en cabina insonorizada simultánea con micrófono profesional y dispositivo. Autoría propia.

Los resultados de dichas pruebas se muestran en las tablas 4.3 y 4.4, que corresponden a los datos del micrófono y el dispositivo respectivamente.

Tabla 4.3: Resultados de parámetros obtenidos mediante el micrófono (Beta4-080) en prueba simultánea.

Ind	Sexo	F_0 (Hz)	Intensidad (dB)	Tiempo de fonación (s)	Jitter local (%)	Shimmer local (%)	HNR (dB)
1	M	143.725	58.433	7.387	0.374	3.857	16.935
2	M	108.405	60.658	11.5125	0.408	3.172	14.92
3	M	95.226	56.389	10.457	0.303	3.534	14.472
4	F	247.589	60.185	8.73	0.23	1.992	22.391
5	M	124.306	69.269	15.245	0.428	2.006	15.399
6	M	117.255	54.073	8.058	0.322	3.096	18.333
7	M	123.522	63.351	6.236	0.285	2.733	18.791
8	F	210.068	63.246	12.567	0.172	1.233	26.352

Tabla 4.4: Resultados de parámetros obtenidos mediante el dispositivo (acelerómetro) en prueba simultánea.

Ind	Sexo	F_0 (Hz)	Intensidad (dB)	Tiempo de fonación (s)	Jitter local (%)	Shimmer local (%)	HNR (dB)
1	M	135.942	50.0863	7.387	0.526	3.135	13.976
2	M	103.135	59.1293	11.5125	0.729	2.289	11.996
3	M	89.979	54.021	10.457	0.675	3.125	12.102
4	F	235.559	61.458	8.73	0.533	2.31	16.079
5	M	113.65	65.3998	15.245	0.635	1.969	14.057
6	M	106.419	53.108	8.058	0.785	3.678	15.698
7	M	113.888	58.86	6.236	0.533	2.587	13.322
8	F	213.465	59.201	12.567	0.505	1.654	21.985

4.7.1. Análisis estadístico

Con los datos de las tablas 4.3 y 4.4, se realizó un análisis estadístico a fin de comparar los resultados obtenidos mediante el micrófono y el dispositivo de manera simultánea. Para ello se utilizaron parámetros estadísticos aplicando el intervalo de confianza. En las tablas 4.5, 4.6, 4.7 y 4.8, se tienen los resultados utilizando un intervalo de confianza del 95 %.

Tabla 4.5: Estadísticas del dispositivo para hombres.

	Alfa	Media	Varianza	Intervalo	L. Superior	L. Inferior
F_0 (Hz)	0.05	110.502	15.236	12.191	122.693	98.311
Intensidad (dB)		56.767	5.478	4.383	61.151	52.384
Tiempo de fonación (s)		9.816	3.304	2.643	12.459	7.173
Jitter local (%)		0.647	0.104	0.083	0.731	0.564
Shimmer local (%)		2.797	0.630	0.504	3.301	2.293
HNR (dB)		13.525	1.387	1.109	14.635	12.416

Tabla 4.6: Estadísticas del dispositivo para mujeres.

	Alfa	Media	Varianza	Intervalo	L. Superior	L. Inferior
F_0 (Hz)	0.05	224.512	15.623	21.652	246.164	202.860
Intensidad (dB)		60.330	1.596	2.212	62.541	58.118
Tiempo de fonación (s)		10.649	2.713	3.760	14.409	6.888
Jitter local (%)		0.519	0.020	0.027	0.546	0.492
Shimmer local (%)		1.982	0.464	0.643	2.625	1.339
HNR (dB)		19.032	4.176	5.788	24.820	13.244

Tabla 4.7: Estadísticas del micrófono para hombres.

	Alfa	Media	Varianza	Intervalo	L. Superior	L. Inferior
F_0 (Hz)	0.05	118.740	16.371	13.099	131.839	105.641
Intensidad (dB)		60.363	5.430	4.345	64.707	56.018
Tiempo de fonación (s)		9.816	3.304	2.643	12.459	7.173
Jitter local (%)		0.353	0.059	0.047	0.400	0.306
Shimmer local (%)		3.066	0.647	0.517	3.584	2.549
HNR (dB)		16.475	1.823	1.459	17.934	15.016

Tabla 4.8: Estadísticas del micrófono para mujeres.

	Alfa	Media	Varianza	Intervalo	L. Superior	L. Inferior
F_0 (Hz)	0.05	228.829	26.531	36.770	265.598	192.059
Intensidad (dB)		61.716	2.164	3.000	64.716	58.716
Tiempo de fonación (s)		10.649	2.713	3.760	14.409	6.888
Jitter local (%)		0.201	0.041	0.057	0.258	0.144
Shimmer local (%)		1.613	0.537	0.744	2.356	0.869
HNR (dB)		24.372	2.801	3.882	28.253	20.490

En la figura 4.19 se muestra el análisis estadístico comparando la frecuencia fundamental obtenida mediante el dispositivo y el micrófono, en hombres y mujeres. Se evidencia que en el caso de los hombres la diferencia de frecuencia en promedio es de 8,238 Hz, mientras que en el caso de mujeres es un valor menor (4.317Hz). Esto demuestra que el dispositivo tiene una aproximación sobresaliente en el apartado de frecuencia fundamental.

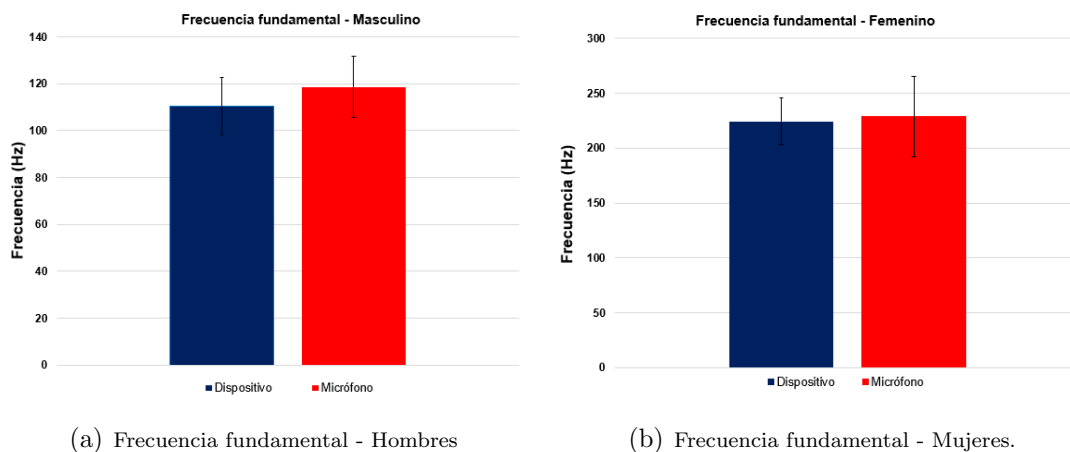


Figura 4.19: Análisis estadístico de la frecuencia fundamental. Autoría propia.

En la figura 4.20, se muestra el análisis estadístico comparando la intensidad obte-

nida mediante el dispositivo y el micrófono, en hombres y mujeres. Al igual que en el punto anterior, en el caso de los hombres la diferencia de intensidad en promedio es de 3.597dB, mientras que en el caso de mujeres es un valor menor (1.386). Esto demuestra que el dispositivo tiene una aproximación sobresaliente en el apartado de intensidad.

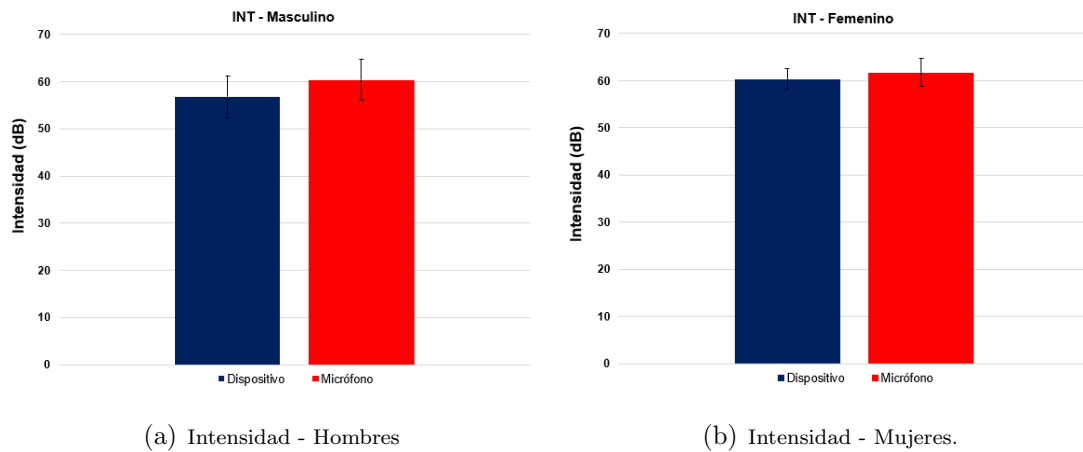


Figura 4.20: Análisis estadístico de la intensidad. Autoría propia.

En la figura 4.21, se muestra el análisis estadístico comparando el tiempo de fonación obtenido mediante el dispositivo y el micrófono, en hombres y mujeres. Debido a que las pruebas se hicieron de manera simultánea, el valor de tiempo de fonación en ambos casos será el mismo.

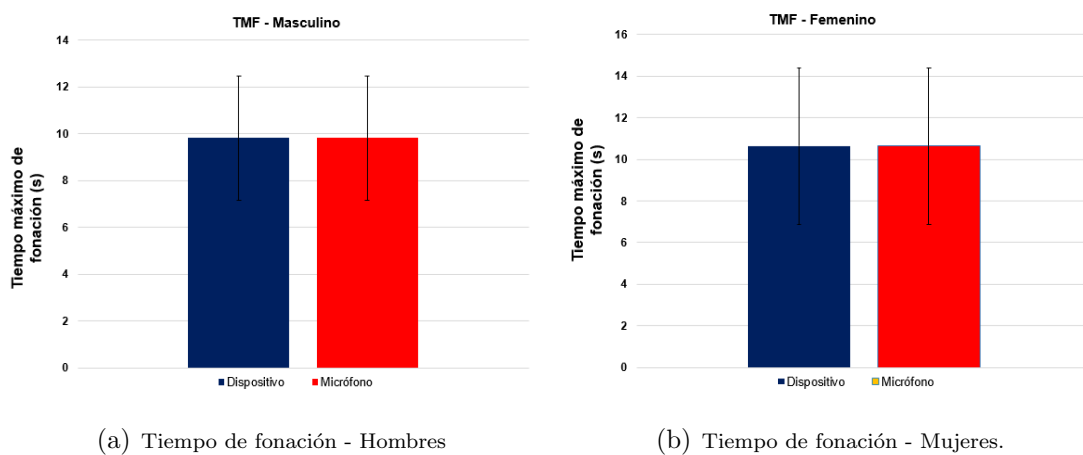


Figura 4.21: Análisis estadístico del tiempo de fonación. Autoría propia.

En la figura 4.22, se muestra el análisis estadístico comparando el *jitter* obtenido mediante el dispositivo y el micrófono, en hombres y mujeres. En este caso, tanto para los hombres como para las mujeres, existe una diferencia notable con respecto a los demás parámetros. Es el parámetro que más error tiene al realizar la comparativa,

puesto que la diferencia en promedio es de 0.294 % para los hombres y de 0.318 % para las mujeres, esta diferencia es grande ya que el umbral máximo para la comparativa entre un individuo sano o patológico es del 1 %, por lo que una mínima diferencia en este parámetro es significativa.

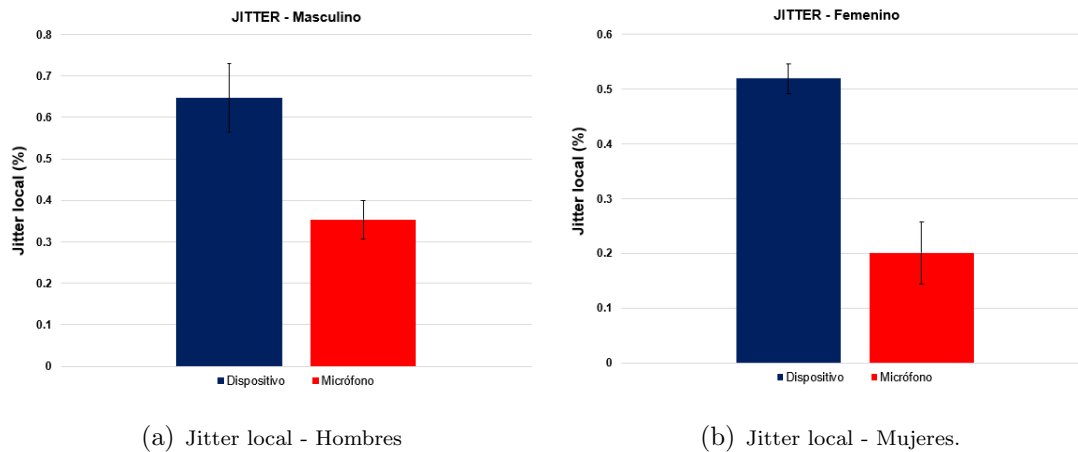


Figura 4.22: Análisis estadístico del jitter local. Autoría propia.

En la figura 4.23, se muestra el análisis estadístico comparando el *shimmer* obtenido mediante el dispositivo y el micrófono, en hombres y mujeres. Existe una diferencia en promedio de 0.269 % para los hombres y 0.369 % para las mujeres, esta diferencia es relativamente menor comparada con el *jitter*, ya que el umbral máximo de consideración entre individuos sanos y patológicos es del 3 %, por lo que la aproximación de este parámetro es buena.

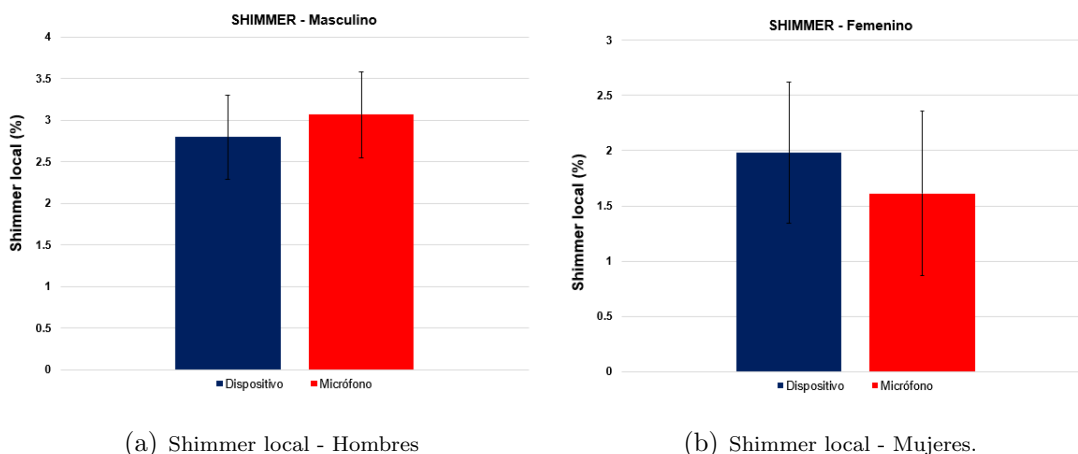


Figura 4.23: Análisis estadístico del shimmer local. Autoría propia.

En la figura 4.24, se muestra el análisis estadístico comparando el HNR obtenido mediante el dispositivo y el micrófono, en hombres y mujeres. Por último analizando

el HNR se observó una diferencia en promedio de 2.95dB para hombres y 5.34dB en mujeres, considerando la naturaleza del parámetro el resultado de la comparativa es bueno.

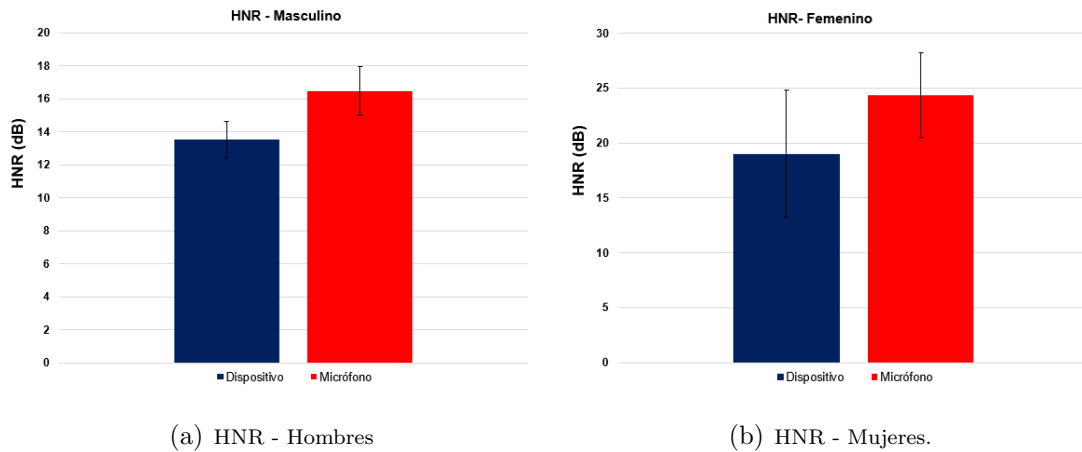


Figura 4.24: Análisis estadístico del HNR. Autoría propia.

Para comprobar la robustez del dispositivo electrónico frente al ruido ambiente, se han realizado las pruebas fuera de la cabina de grabación, como se observa en la figura 4.25. Los resultados se plasman en la tabla 4.9, donde es evidente que todos los parámetros tienen una variación mínima con respecto a las pruebas que se realizaron simultáneamente dentro de la cabina de grabación. Dichas pruebas forman parte de los resultados finales que se mostrarán posteriormente.



Figura 4.25: Prueba fuera de la cabina insonorizada del dispositivo. Autoría propia.

Tabla 4.9: Resultados de parámetros obtenidos mediante el dispositivo en prueba no simultánea.

Ind	Sexo	F_0 (Hz)	Intensidad (dB)	Tiempo de fonación (s)	Jitter local (%)	Shimmer local (%)	HNR (dB)
1	M	132.969	56.4356	9.3133	0.597	2.961	12.059
2	M	102.765	65.7	12.632	0.476	1.139	20.02
3	M	98.68	52.3	16.42	0.361	2.989	12.068
4	F	231.56	63.965	13.25	0.461	3.141	20.12
5	M	112.965	61.896	15.323	1.101	2.056	21.052
6	M	104.673	53.281	8.456	0.336	4.065	12.675
7	M	100.869	60.421	14.63	0.624	5.032	12.768
8	F	211.068	62.2465	12.567	0.972	1.933	21.352

En la tabla 4.13, se muestra el resumen de la obtención de los parámetros con cada uno de los individuos que fueron evaluados, cabe mencionar que las pruebas finales, se desarrollaron en un ambiente normal, es decir, en un ambiente con ruido y se comparará con el informe diagnóstico de parte de los fonoaudiólogos. Las pruebas se realizaron a un total de 21 individuos, de los cuales 12 son hombres y 9 mujeres, que previamente aceptaron ser partícipes del estudio tanto de parte de los fonoaudiólogos como de las pruebas con el dispositivo. Cabe mencionar que todos los individuos evaluados son pacientes post COVID-19. En la sección de anexos D, se muestra el modelo del consentimiento informado que cada individuo aceptó firmar de manera voluntaria, también, se anexa en E, el modelo de la ficha de historial fonoaudiológico que los especialistas de la voz realizan y por último se obtiene el informe de parte de los fonoaudiólogos, se extraen los primeros 10 individuos para focalizar la evaluación y comparación del informe fonoaudiológico comparado con los resultados de los parámetros obtenidos con el dispositivo electrónico, refiérase al anexo F.

Tabla 4.10: Parámetros del análisis vocal de individuos

Ind	Sexo	F_0 (Hz)	Intensidad (dB)	Tiempo de fonación (s)	Jitter local (%)	Shimmer local (%)	HNR (dB)
1	M	132.969	56.4356	9.3133	0.597	2.961	12.059
2	M	102.765	65.7	12.632	0.476	1.139	20.02
3	M	98.68	52.3	16.42	0.361	2.989	12.068
4	F	231.56	63.965	13.25	0.461	3.141	20.12
5	M	112.965	61.896	15.323	1.101	2.056	21.052
6	M	104.673	53.281	8.456	0.336	4.065	12.675
7	M	100.869	60.421	14.63	0.624	5.032	12.768
8	F	211.068	62.2465	12.567	0.972	1.933	21.352
9	M	108.698	53.25	12.334	0.561	3.937	12.865
10	F	210.965	57.063	9.256	0.903	4.577	14.768
11	F	224.468	63.469	8.479	0.729	1.453	20.703
12	M	119.464	58.234	17.54	0.563	2.748	12.397
13	M	97.8	60.443	10.897	0.555	2.757	15.967
14	F	206.645	59.392	9.35	0.856	2.545	16.842
15	F	198.544	62.252	8.86	1.068	3.056	16.534
16	F	234.096	63.136	13.647	0.466	1.699	18.05
17	M	94.855	58.804	9.607	0.5232	2.259	15.707
18	M	126.4	52.436	8.265	0.759	2.644	14.06
19	M	95.241	54.967	14.89	0.796	2.795	12.506
20	F	193.295	61.905	11.237	1.05	2.11	14.759
21	F	222.049	62.504	15.686	0.801	1.636	19.058

4.8. Análisis y evaluación de los parámetros acústicos

Existen puntos claves y aspectos a considerar al momento que se realizan las pruebas. En primer lugar las pruebas se deben ejecutar de manera simultánea, es decir, las pruebas elaboradas por parte de los fonoaudiólogos y las pruebas elaboradas por el dispositivo deben ser realizadas al mismo tiempo para descartar cualquier variación que el individuo pueda tener después en su voz, ya que, si las pruebas se ejecutan de forma independiente el individuo puede tener variaciones en su voz, por lo que las pruebas no tendrían validez. También existe la posibilidad de que se produzcan variaciones por los cambios en la metodología del registro de la voz, las diferentes gamas de equipos existentes al igual que los algoritmos empleados por los distintos programas. Por esta razón es necesario que las pruebas se lo hagan con presencia de los especialistas, ya que el dispositivo va a formar parte de una ayuda al diagnóstico y optimizar en alguna medida una parte del procedimiento llevado a cabo por los especialistas de la voz.

Para la evaluación de cada uno de los parámetros se ha analizado diferentes estudios en los existan rangos de normalidad o establezcan un promedio normal de cada parámetro. Para la frecuencia fundamental, *jitter* local, *shimmer* local y HNR, se ha tomado en cuenta la siguiente referencia [92]. Para los valores normales del parámetro de tiempo de fonación, la referencia [93], y para los valores normales de intensidad, la referencia [94]. Con el fin de automatizar el análisis de los parámetros dentro del rango de normalidad, se desarrolló un *script* en Python, en la figura 4.26, se muestra un ejemplo del análisis para un individuo.

```
El análisis vocal del individuo indica que:
La frecuencia fundamental está dentro del rango normal
La intensidad está dentro del rango normal
El tiempo máximo de fonación está debajo del promedio
El jitter está dentro del rango normal
El shimmer está dentro del rango normal
El HNR está dentro del rango normal
```

Figura 4.26: Resultado de análisis vocal de un individuo. Autoría propia.

4.8.1. Frecuencia Fundamental (F_0)

El parámetro, cambia según el género y la edad. Rangos normales en el hombre están alrededor de (87-182 HZ) y en la mujer alrededor de (165-264 Hz). Dichos valores pueden cambiar en presencia de alguna patología.

En este estudio, como se observa en la tabla 4.13, los valores de la frecuencia funda-

mental en hombres (M) se encuentran dentro de los rangos normales, de igual forma los valores de frecuencia fundamental en mujeres (F) se encuentran dentro de los rangos normales.

4.8.2. Intensidad

Este parámetro puede variar dependiendo de la sensibilidad, calidad del micrófono y la distancia a la que se coloca el individuo del micrófono, por lo que es indispensable una correcta colocación del individuo para que no exista variaciones, en el caso del dispositivo electrónico es importante el ajuste correcto del sensor acelerómetro para que no existan variaciones o cambios en este parámetro. Valores normales de este parámetro son valores mayores a 40dB ($> 40dB$) [94]. En este estudio, se puede observar que los valores de la intensidad se encuentran en rangos normales tanto para todos los individuos, tal como se observa en la tabla 4.13.

4.8.3. Tiempo de fonación

En este caso, existen diversos estudios que aportan con datos normativos sobre el tiempo máximo de fonación, dichos valores mencionados presentan una alta variabilidad entre ellos y depende de varios factores como el entorno de grabación, la condición física de cada individuo, el escenario de las pruebas, etc. El estudio que ha sido escogido para tomar de referencia ha sido [93], en el que manifiestan que la medida del tiempo máximo de fonación en promedio para un hombre es de alrededor de 14 segundos y para la mujer es de alrededor de 11 segundos. En este estudio, hay varios individuos que están por debajo del promedio del tiempo de fonación, como se muestra en la tabla 4.11.

Tabla 4.11: Evaluación del parámetro tiempo de fonación en cada individuo.

Ind	Sexo	Tiempo de fonación (s)	Valoración
1	M	9.3133	Por debajo del promedio
2	M	12.632	Por debajo del promedio
3	M	16.42	Por encima del promedio
4	F	13.25	Por encima del promedio
5	M	15.323	Por encima del promedio
6	M	8.456	Por debajo del promedio
7	M	14.63	Por encima del promedio
8	F	12.567	Por encima del promedio
9	M	12.334	Por debajo del promedio
10	F	9.256	Por debajo del promedio
11	F	8.479	Por debajo del promedio
12	M	17.54	Por encima del promedio
13	M	10.897	Por debajo del promedio
14	F	9.35	Por debajo del promedio
15	F	8.86	Por debajo del promedio
16	F	13.647	Por encima del promedio
17	M	9.607	Por debajo del promedio
18	M	8.265	Por debajo del promedio
19	M	14.89	Por encima del promedio
20	F	11.237	Por encima del promedio
21	F	15.686	Por encima del promedio

4.8.4. Jitter local

Este parámetro relacionado directamente con la perturbación de la frecuencia, mide el grado de variación entre la duración de ciclos vocales consecutivos, es asociado al grado de ronquera, rozamiento de las cuerdas vocales o dureza de la voz. Según [92], el rango normal del parámetro en hombres, se encuentra en (0.16 % a 0.95 %) y en mujeres, se encuentra en (0.16 % a 1 %). En este estudio, existen tres individuos (5, 15

y 20) que están brevemente fuera de rango, como se muestra en la tabla 4.12.

Tabla 4.12: Evaluación del parámetro jitter local en cada individuo.

Ind	Sexo	Jitter local (%)	Valoración
1	M	0.597	Dentro del rango normal
2	M	0.476	Dentro del rango normal
3	M	0.361	Dentro del rango normal
4	F	0.461	Dentro del rango normal
5	M	1.101	Fuera del rango normal
6	M	0.336	Dentro del rango normal
7	M	0.624	Dentro del rango normal
8	F	0.972	Dentro del rango normal
9	M	0.561	Dentro del rango normal
10	F	0.903	Dentro del rango normal
11	F	0.729	Dentro del rango normal
12	M	0.563	Dentro del rango normal
13	M	0.555	Dentro del rango normal
14	F	0.856	Dentro del rango normal
15	F	1.068	Fuera del rango normal
16	F	0.466	Dentro del rango normal
17	M	0.5232	Dentro del rango normal
18	M	0.759	Dentro del rango normal
19	M	0.796	Dentro del rango normal
20	F	1.05	Fuera del rango normal
21	F	0.801	Dentro del rango normal

4.8.5. Shimmer local

Este parámetro relacionado directamente con la perturbación de la intensidad, mide el grado de variación que existe en la amplitud entre ciclos vocales consecutivos. Según [92], el rango normal del parámetro, en hombres, se encuentra en (0.88 % a 8.76 %) y

en mujeres, se encuentra en (1.02 % a 3.32 %). En este estudio, existe un individuo (10) que se encuentra fuera del rango normal de este parámetro, tal como se observa en la tabla 4.13.

Tabla 4.13: Evaluación del parámetro shimmer local en cada individuo.

Ind	Sexo	Shimmer local (%)	Valoración
1	M	2.961	Dentro del rango normal
2	M	1.139	Dentro del rango normal
3	M	2.989	Dentro del rango normal
4	F	3.141	Dentro del rango normal
5	M	2.056	Dentro del rango normal
6	M	4.065	Dentro del rango normal
7	M	5.032	Dentro del rango normal
8	F	1.933	Dentro del rango normal
9	M	3.937	Dentro del rango normal
10	F	4.577	Fuera del rango normal
11	F	1.453	Dentro del rango normal
12	M	2.748	Dentro del rango normal
13	M	2.757	Dentro del rango normal
14	F	2.545	Dentro del rango normal
15	F	3.056	Dentro del rango normal
16	F	1.699	Dentro del rango normal
17	M	2.259	Dentro del rango normal
18	M	2.644	Dentro del rango normal
19	M	2.795	Dentro del rango normal
20	F	2.11	Dentro del rango normal
21	F	1.636	Dentro del rango normal

4.8.6. HNR

Este parámetro mide la relación armónico-ruido, comparando el sonido de la emisión de la vocal sostenida, conjuntamente con el componente armónico del componente de ruido acompañante. Según [92], el rango normal del parámetro en hombres, se encuentra en (12.63 a 29.86) y en mujeres, se encuentra en (15.12 a 29.04). En este estudio, existen cinco individuos (1, 3, 10, 12 y 19) que están ligeramente fuera de rango, como se muestra en la tabla 4.14.

Tabla 4.14: Evaluación del parámetro jitter local en cada individuo.

Ind	Sexo	HNR (dB)	Valoración
1	M	12.059	Fuera del rango normal
2	M	20.02	Dentro del rango normal
3	M	12.068	Fuera del rango normal
4	F	20.12	Dentro del rango normal
5	M	21.052	Dentro del rango normal
6	M	12.675	Dentro del rango normal
7	M	12.768	Dentro del rango normal
8	F	21.352	Dentro del rango normal
9	M	12.865	Dentro del rango normal
10	F	14.768	Fuera del rango normal
11	F	20.703	Dentro del rango normal
12	M	12.397	Fuera del rango normal
13	M	15.967	Dentro del rango normal
14	F	16.842	Dentro del rango normal
15	F	16.534	Dentro del rango normal
16	F	18.05	Dentro del rango normal
17	M	15.707	Dentro del rango normal
18	M	14.06	Dentro del rango normal
19	M	12.506	Fuera del rango normal
20	F	14.759	Dentro del rango normal
21	F	19.058	Dentro del rango normal

En la siguiente tabla 4.15, se presenta un resumen de cada individuo conjuntamente con el parámetro que se ha medido, marcando si existe una variación de los rangos normales del parámetro, o en el caso del parámetro de tiempo de fonación está por debajo o encima del promedio. El visto (✓) en la tabla, indica que el parámetro se encuentra dentro de los rangos normales y en el caso del parámetro del tiempo de fonación, que el resultado está por encima de la media. El asterisco (*) en la tabla, indica que se encuentra fuera del rango normal o en el caso del parámetro de tiempo de fonación, que está por debajo del promedio. En general, se puede concluir que en algunos individuos existe una desviación del rango normal de algunos parámetros, pero dicha desviación es mínima, en el caso del parámetro del tiempo de fonación, como ya se había mencionado el parámetro tiene una variabilidad alta en cada individuo.

Tabla 4.15: Resumen de cada individuo y el parámetro medido.

ind	Sexo	F_0 (Hz)	I (dB)	T_f (s)	Jitt (%)	Shimm (%)	HNR (dB)
Rangos N. (M)		(87-182)	(>40)	14	(0.16-0.95)	(0.88-8.76)	(12.63-29.86)
Rangos N. (F)		(165-264)	(>40)	11	(0.16-1)	(1.02-3.32)	(15.12-29.04)
1	M	✓	✓	*	✓	✓	*
2	M	✓	✓	*	✓	✓	✓
3	M	✓	✓	✓	✓	✓	*
4	F	✓	✓	✓	✓	✓	✓
5	M	✓	✓	✓	*	✓	✓
6	M	✓	✓	*	✓	✓	✓
7	M	✓	✓	✓	✓	✓	✓
8	F	✓	✓	✓	✓	✓	✓
9	M	✓	✓	*	✓	✓	✓
10	F	✓	✓	*	✓	*	*
11	F	✓	✓	*	✓	✓	✓
12	M	✓	✓	✓	✓	✓	*
13	M	✓	✓	*	✓	✓	✓
14	F	✓	✓	*	✓	✓	✓
15	F	✓	✓	*	*	✓	✓
16	F	✓	✓	✓	✓	✓	✓
17	M	✓	✓	*	✓	✓	✓
18	M	✓	✓	*	✓	✓	*
19	M	✓	✓	✓	✓	✓	*
20	F	✓	✓	✓	*	✓	✓
21	F	✓	✓	✓	✓	✓	✓

4.9. Comparación de los parámetros adquiridos con el dispositivo y el informe fonoaudiológico

Para sintetizar la comparación, se revisa a detalle cada uno de los comentarios y el diagnóstico presuntivo en el informe fonoaudiológico (F), de los 10 primeros individuos que tengan disfonía de base funcional, con derivación al Centro de Otorrinolaringología (ORL) para estudio estroboscópico y se compara con la tabla 4.15 en donde se expone el resumen de rangos de normalidad de cada individuo.

Los individuos con dicho diagnóstico presuntivo son: (1, 3, 5, 6, 8 y 9).

- **Individuo 1:** El informe fonoaudiológico señala que el individuo está fuera de rangos normales de los parámetros: tiempo de fonación y HNR, lo cual de igual manera se ve reflejado en la tabla 4.15, por lo que existe una congruencia positiva en las pruebas.
- **Individuo 3:** El informe fonoaudiológico señala que el individuo está dentro de rangos normales de los parámetros: pero al final de las tareas fonatorias existe una voz más grave (*vocal fry*). De alguna manera eso se ve reflejado en el HNR de la tabla 4.15. Cabe mencionar, que el individuo usa su voz con frecuencia.
- **Individuo 5:** El informe fonoaudiológico señala que el individuo está dentro de rangos normales de los parámetros y según las cualidades de su voz mantiene un grado de alteración vocal moderada, tiene una autopercepción de irregular de su voz y se considera como una persona normalmente habladora. La tabla 4.15 muestra una pequeña desviación del rango normal del parámetro *jitter*.
- **Individuo 6:** El informe fonoaudiológico señala que el individuo está dentro de rangos normales a excepción del parámetro Tiempo Máximo de Fonación (TMF) lo cual también se evidencia en la tabla 4.15. Cabe señalar que el parámetro HNR se encuentra cercano al límite del rango normal.
- **Individuo 8:** El informe fonoaudiológico señala que el individuo está dentro de rangos normales a excepción del parámetro TMF, en este caso, en la tabla 4.15 se puede observar que el parámetro TMF se encuentra en un rango normal, pero en la tabla 4.11 se observa que el parámetro se encuentra cercano al límite del rango de normalidad.
- **Individuo 9:** El informe fonoaudiológico señala que el individuo está dentro de rangos normales a excepción del parámetro TMF, lo cual también se evidencia en la tabla 4.15.

4.10. Adquisición del oscilograma y espectrograma a través de Praat

Para complementar los resultados de los parámetros, se puede optar por la visualización de las imágenes acústicas como son el oscilograma y espectrograma, para ello se debe ingresar a través del Graphical User Interface (GUI) del programa Praat y abrir el archivo de audio que se desee analizar. En este caso en la figura 4.27, se muestra el archivo de audio completo, no es posible visualizar el espectrograma ni los parámetros debido a que se debe seccionar y seleccionar una ventana de 10 segundos máximo. En la figura 4.28 se muestra dicha ventana; se puede observar el oscilograma en la parte superior, el espectrograma en la parte inferior conjuntamente con los parámetros de frecuencia fundamental e intensidad con sus respectivas escalas, esto será de utilidad para el especialista de la voz.

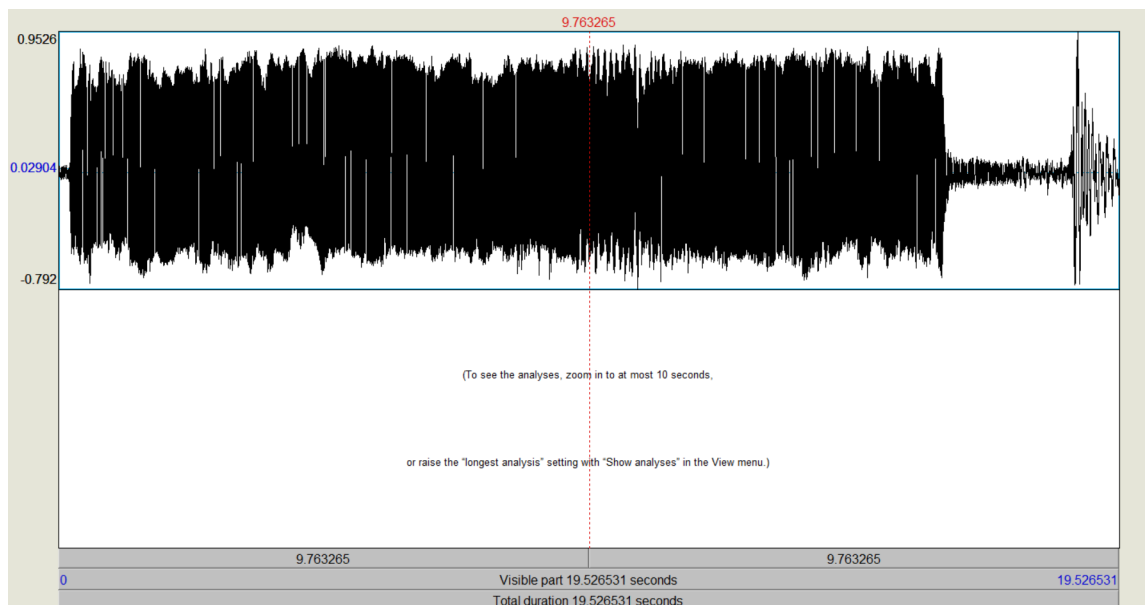


Figura 4.27: Adquisición del oscilograma completo. Autoría propia.

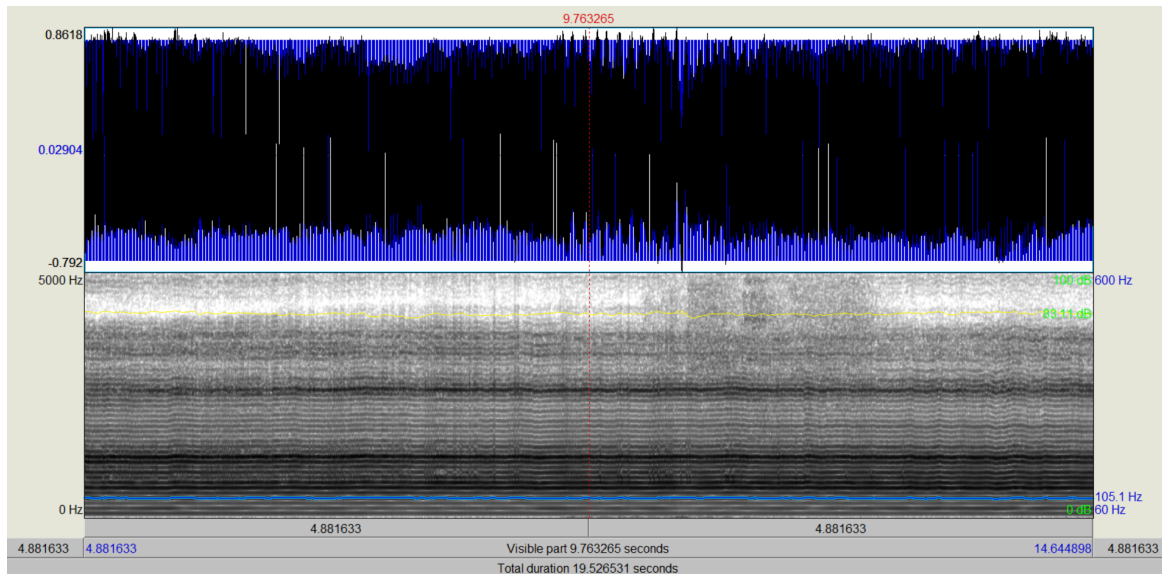


Figura 4.28: Adquisición del oscilograma y espectrograma en un rango de 10 segundos. Autoría propia.

4.11. Análisis y evaluación de oscilograma y espectrograma

El análisis de las imágenes acústicas es una herramienta complementaria dentro de la valoración patológica de la voz y forma parte del diagnóstico de un especialista de la voz. Sin embargo, existe el inconveniente del punto de vista del observador, es decir, que la interpretación depende de la valoración visual del que analiza [95].

Cada una de las líneas que se observa en el espectrograma, representa un armónico de la descomposición del sonido complejo. La frecuencia más baja es llamada la frecuencia fundamental, en algunos casos existen líneas debajo de la línea de frecuencia fundamental llamadas sub-armónicos y las líneas por encima de la frecuencia fundamental corresponden a los demás armónicos. La amplitud de cada una de las líneas es representada por el grosor de cada línea, mientras mayor es el grosor, mayor amplitud o intensidad del sonido [96].

4.11.1. Evaluación del oscilograma y espectrograma de un individuo con diagnóstico presuntivo (disfónico)

En la figura 4.29 se muestra un ejemplo de un paciente con diagnóstico presuntivo disfónico, el espectrograma está configurado con un filtro de banda ancha (300Hz). De manera clara se observa que en el espectrograma existe una alta variabilidad en

el parámetro de la intensidad (color amarillo) al igual que la frecuencia fundamental (color azul). Por lo que se podría estudiar más a fondo las imágenes acústicas para una valoración exhaustiva de ser el caso.

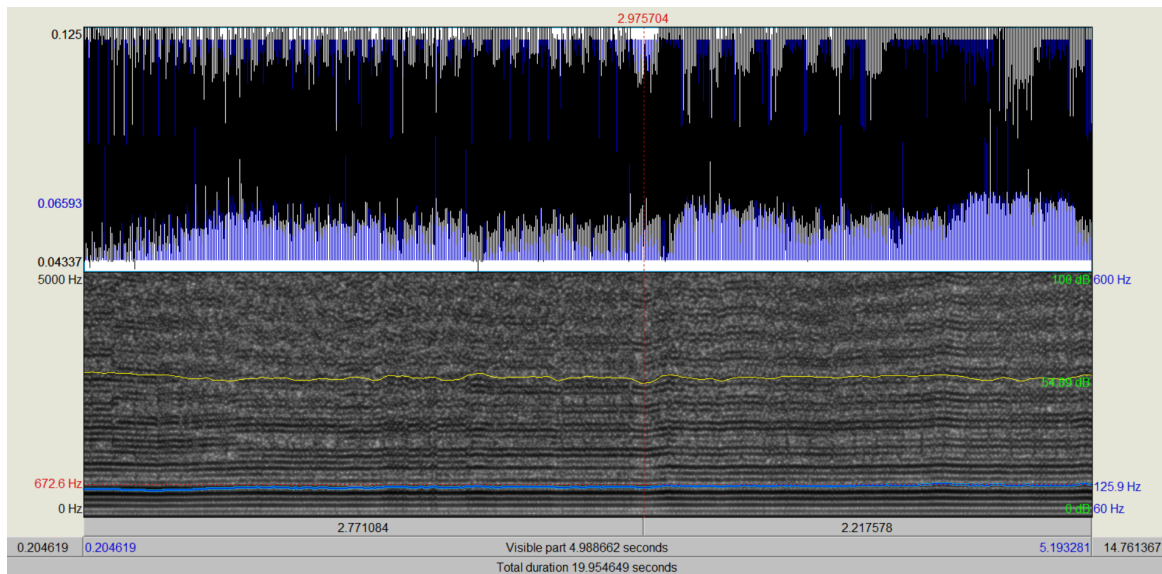


Figura 4.29: Adquisición del oscilograma y espectrograma en un paciente con diagnóstico presuntivo disfónico. Autoría propia.

4.11.2. Evaluación del oscilograma y espectrograma de un individuo con diagnóstico presuntivo sano (eufonía)

En la figura 4.30, se muestra un ejemplo de un paciente con diagnóstico presuntivo sano (eufonía); el espectrograma está configurado con un filtro de banda ancha (300Hz). En comparación con la figura 4.29, es evidente la diferencia en cuanto a la variabilidad de los parámetros, en este caso la intensidad y la frecuencia fundamental son mucho más constantes, casi lineales.

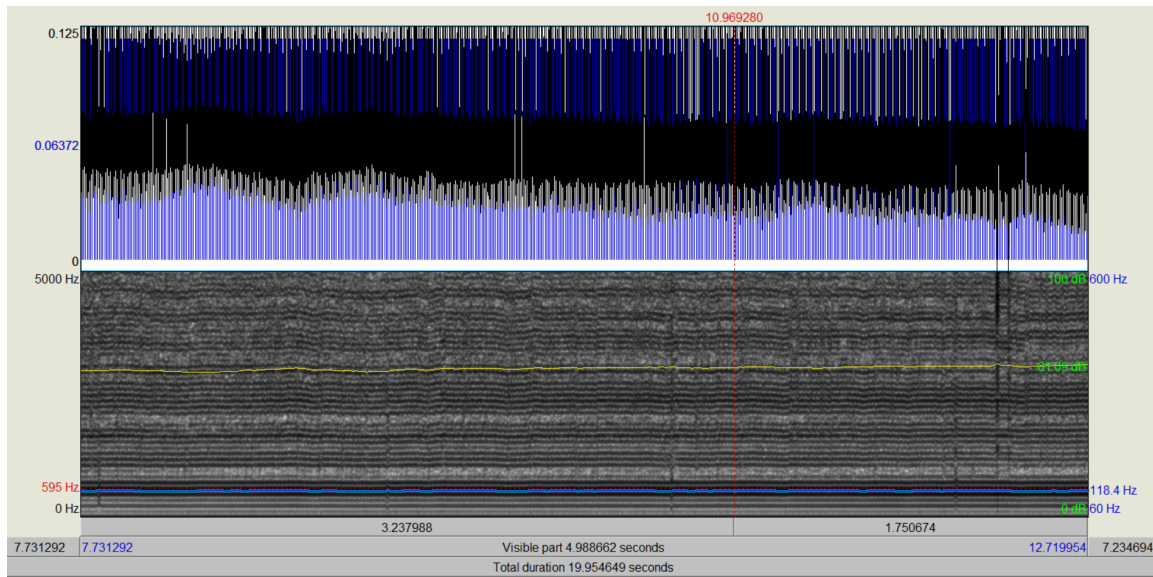


Figura 4.30: Adquisición del oscilograma y espectrograma en un paciente con diagnóstico presuntivo sano (eufonía). Autoría propia.

5. Conclusiones y recomendaciones

5.1. Conclusiones

- Se diseñó e implementó un dispositivo electrónico para el apoyo a un especialista en la voz, en este caso, un fonoaudiólogo. Dicho dispositivo facilita en gran medida la adquisición y valoración de los parámetros acústicos de la voz. El dispositivo en mención, tiene la característica esencial de no ser invasivo, debido a que trabaja directamente con las vibraciones de las cuerdas vocales, por lo que la adquisición de los parámetros se lo realiza de una forma mucho más sencilla y rápida.
- Se ha optado por la utilización de un sensor acelerómetro para la recolección de datos acústicos, gracias a la naturaleza del sensor el dispositivo no es invasivo, ya que dicho sensor es colocado en la superficie del cuello y trabaja directamente sobre las vibraciones de las cuerdas vocales. Se aprovecha también la característica de la robustez frente al ruido ambiente.
- El diagnóstico completo de un especialista de la voz, en este caso de un fonoaudiólogo, es un proceso exhaustivo en el que se realizan varios procedimientos tanto fisiológicos como de análisis acústico para detectar anomalías en la voz. Este proyecto focaliza una pequeña parte del diagnóstico de un fonoaudiólogo, que es el estudio acústico, es decir, identificar los parámetros fundamentales de la voz del individuo. Los parámetros como: frecuencia fundamental, intensidad, tiempo de fonación, *jitter*, *shimmer*, *HNR*, entre otros, tienen un rango definido para personas sanas y patológicas. De esta manera, el fonoaudiólogo tiene una herramienta para dar paso a un análisis más profundo y detallado en base a lo ya descrito anteriormente.
- Una característica primordial del dispositivo desarrollado es la robustez frente al ruido ambiente, por lo que no necesita de una cabina insonorizada para proceder con las pruebas a individuos. La fiabilidad del mismo con respecto a un micrófono profesional (que es el equipo que se utiliza a menudo para este trabajo) es sobresaliente en los parámetros principales (frecuencia fundamental, intensidad y tiempo de fonación). La frecuencia fundamental tiene una diferencia de promedio de 8.238Hz para los hombres y 4.317Hz para las mujeres, tomando en cuenta la naturaleza del parámetro la comparativa resulta sobresaliente. De igual manera, en el caso de la intensidad la diferencia en el promedio es de 3.597dB para los hombres y 1.386 dB para las mujeres, por lo que la comparativa es sobresaliente. En el caso del parámetro de tiempo de fonación, al realizarse las pruebas

simultáneamente este parámetro es el mismo tanto para el dispositivo, como para el micrófono.

- Existe un mayor margen de error en los tres siguientes parámetros: (*jitter*, *shimmer* y *HNR*). En el caso del *jitter*, es el parámetro que más error tiene al realizar la comparativa, puesto que la diferencia en promedio es de 0.294% para los hombres y de 0.318% para las mujeres, esta diferencia es grande ya que el umbral máximo para la comparativa entre un individuo sano o patológico es del 1%, por lo que una mínima diferencia en este parámetro es significativa. En cuanto al *shimmer*, existe una diferencia en promedio de 0.269% para los hombres y 0.369% para las mujeres, esta diferencia es relativamente menor comparada con el parámetro *jitter*, ya que el umbral máximo de consideración entre individuos sanos y patológicos en el *shimmer* es del 3%, por lo que la aproximación de este parámetro es buena. Por último analizando el *HNR* se observó una diferencia en promedio de 2.95dB para hombres y 5.34dB en mujeres, considerando la naturaleza del parámetro el resultado de la comparativa es bueno. Si bien los resultados no son idénticos, son bastante similares lo que permite corroborar la fiabilidad del equipo.
- El ESP32 como parte principal del diseño del circuito ocasiona un gran consumo energético. Para posibilitar la monitorización de los parámetros de forma ininterrumpida es necesario contar con una batería de una capacidad razonable que permita al menos la utilización del mismo durante un día regular de consulta. En caso de ser necesario también se puede alimentar al ESP32 con un cable micro USB lo que posibilita la utilización del dispositivo de forma indefinida sin preocupaciones de desgaste de batería.
- Los datos recolectados por el dispositivo fueron procesados en el software Praat que es una herramienta de software libre para el análisis científico del habla usado en lingüística. Se puede realizar un análisis manual con un archivo de audio *.wav* o también se puede programar un *script*. Con el propósito de que el análisis de los datos se automatice, se programó un *script*, el mismo recibe los archivos de audio para posteriormente procesarlos. Complementando esto, se desarrolló un *script* en Python para clasificar los parámetros que se obtienen en el *script* de Praat dentro o fuera del rango de normalidad. Así pues, en seguida de tomar los datos ya se puede tener análisis y resultados de la muestra.
- Se desarrolló el análisis de los parámetros acústicos a 21 individuos (12 hombres y 9 mujeres), posteriormente se realizó un banco de pruebas, con la finalidad de comparar los resultados del dispositivo con el informe fonaudiológico final. Los resultados obtenidos en dicha comparativa son en general positivos, ya que existen coincidencias evidentes en ambos procesos.

- La principal ventaja de utilizar el ESP32 es su módulo WiFi. Para la toma de datos se realizó un servidor web con interfaz de usuario de manera que el fonoaudiólogo pueda utilizar el dispositivo de la manera más amigable posible. Entre las funcionalidades de la página web se encuentran grabar un audio, descargar un audio, borrar un audio y revisar el directorio de datos almacenados en la tarjeta micro SD. De esta manera se evita que el especialista tenga que retirar la tarjeta de memoria para extraer los datos y lo pueda hacer mediante la red.
- El dispositivo en mención, puede ser utilizado en múltiples aplicaciones dentro del área de fonoaudiología, una de las principales aplicaciones son: la detección oportuna de afecciones de la voz, comparación de la calidad vocal antes y después de una intervención quirúrgica, monitoreo constante del avance o recuperación dentro del proceso de rehabilitación vocal.
- Finalmente, el dispositivo desarrollado tuvo una respuesta satisfactoria, por lo que podrá ser usada como herramienta de primer frente para el apoyo a los especialistas en el diagnóstico oportuno de anomalías en la producción de la voz

5.2. Recomendaciones

- Trabajar con archivos de audio es un problema puesto que el diseño de la placa debe ser preciso a fin de evitar cualquier ruido, por ello en nuestra experiencia recomendamos colocar un plano de tierra en el diseño de la placa para evitar la suma de ruido en modo común. De igual manera la colocación de los componentes electrónicos en el diseño PCB debe ser lo más cercano al esquemático, de manera que no exista ruido de recorrido.
- Realizar múltiples pruebas electrónicas y ajustar la sensibilidad del dispositivo electrónico en cada una de ellas, de manera que, se encuentre en un rango de buena calidad.
- Referirse a las hojas de especificaciones técnicas de cada sensor o elemento electrónico para resultados mucho más precisos y evitar errores comunes al realizar un circuito electrónico.
- Tener una buena base teórica en el tema de fonoaudiología agilizará el proceso de las pruebas con individuos y los resultados serán direccionados de manera correcta.
- Antes de tener el dispositivo armado y funcional, buscar una población la cual esté dispuesta a realizar pruebas de funcionamiento para evitar retrasos en el proyecto.

5.3. Trabajos futuros

- Implementar un sistema de redes neuronales que permita el procesamiento, análisis y resultados de los datos de forma inmediata.
- Modificar el código fuente para mejorar la interfaz de usuario de la página web, añadiendo más funcionalidades para el especialista y permitiendo que el coste de procesamiento que realiza el ESP32 disminuya.
- Optimizar el circuito electrónico para obtener los parámetros acústicos más cercanos a como se obtendría con una cabina insonorizada y un micrófono profesional, también, optimizar el consumo de corriente y voltaje del circuito electrónico, con la finalidad de ocuparlo de manera portátil, con banco de baterías.
- Implementar una base de datos donde se encuentre cada uno de los pacientes, anexando su expediente clínico y la obtención de los parámetros acústicos, para así permitir un seguimiento de cada individuo.

A. ANEXOS

A. Scripts de servidor web y adquisición de datos desde el ESP32

A.1. Script principal del programa .ino

https://github.com/paulllerena8/ESP32-analisisvocal/blob/main/ESP_File_Download_v01.ino

A.2. Script .h CSS para el diseño de la página web

<https://github.com/paulllerena8/ESP32-analisisvocal/blob/main/CSS.h>

A.3. Script .cpp para configuración del modo I2S

<https://github.com/paulllerena8/ESP32-analisisvocal/blob/main/I2S.cpp>

A.4. Script .h para configuración del modo I2S

<https://github.com/paulllerena8/ESP32-analisisvocal/blob/main/I2S.h>

A.5. Script .h para configuración de la red

<https://github.com/paulllerena8/ESP32-analisisvocal/blob/main/Network.h>

A.6. Script .h para configuración de pin CS del lector SD

https://github.com/paulllerena8/ESP32-analisisvocal/blob/main/Sys_Variables.h

A.7. Script .cpp para configuración del archivo .wav de audio

<https://github.com/paulllerena8/ESP32-analisisvocal/blob/main/Wav.cpp>

A.8. Script .h para configuración del archivo .wav de audio

<https://github.com/paulllerena8/ESP32-analisisvocal/blob/main/Wav.h>

B. Script de Praat para obtención de parámetros

```
#####  
###          Obtenci n de par metros de la voz en Praat  
### Autores:  
###  
### Mauricio Guerrero  
### Pa l Llerena  
###  
### 21/03/2022  
#####  
  
Erase all  
clearinfo  
  
#Carga el archivo de audio con nombre predeterminado a  
  analizar#  
name01$ = "sound1.wav"  
name02$ = "sound2.wav"  
name03$ = "sound3.wav"  
  
#Ingresar el sexo del individuo#  
sexo$ ="F"  
  
sound = Read from file: name01$  
selectObject: sound  
Edit  
editor: sound  
  Select: 3, 7  
  cutsound1 = Extract selected sound (preserve times)  
Close  
Rename: "cutsound1"
```

```
sound = Read from file: name02$
selectObject: sound
Edit
editor: sound
    Select: 3, 7
    cutsound2 = Extract selected sound (preserve times)
Close
Rename: "cutsound2"
```

```
sound = Read from file: name03$
selectObject: sound
Edit
editor: sound
    Select: 3, 7
    cutsound3 = Extract selected sound (preserve times)
Close
Rename: "cutsound3"
```

```
#Separamos el nombre de la extensión del archivo para
    utilizar el nombre#
```

```
Read from file: name01$
name01$ = name01$ - ".wav"
Read from file: name02$
name02$ = name02$ - ".wav"
Read from file: name03$
name03$ = name03$ - ".wav"
```

```
##### Obtención de Frecuencia Fundamental
```

```
#####
```

```
selectObject: cutsound1
To Pitch... 0.0 75 500
meanpitch1 = Get mean... 0 0 Hertz
meansd1 = Get standard deviation... 0 0 Hertz
minpitch1 = Get minimum... 0 0 Hertz Parabolic
maxpitch1 = Get maximum... 0 0 Hertz Parabolic
```

```
selectObject: cutsound2
To Pitch... 0.0 75 500
```

```
meanpitch2 = Get mean... 0 0 Hertz  
meansd2 = Get standard deviation... 0 0 Hertz  
minpitch2 = Get minimum... 0 0 Hertz Parabolic  
maxpitch2 = Get maximum... 0 0 Hertz Parabolic
```

```
selectObject: cutsound3  
To Pitch... 0.0 75 500  
meanpitch3 = Get mean... 0 0 Hertz  
meansd3 = Get standard deviation... 0 0 Hertz  
minpitch3 = Get minimum... 0 0 Hertz Parabolic  
maxpitch3 = Get maximum... 0 0 Hertz Parabolic
```

```
##### Obtenci n de Intensidad #####
```

```
selectObject: cutsound1  
To Intensity... 75 0.001 true  
meanint1 = Get mean... 0 0  
meansdi1 = Get standard deviation... 0 0  
minint1 = Get minimum... 0 0 Parabolic  
maxint1 = Get maximum... 0 0 Parabolic
```

```
selectObject: cutsound2  
To Intensity... 75 0.001 true  
meanint2 = Get mean... 0 0  
meansdi2 = Get standard deviation... 0 0  
minint2 = Get minimum... 0 0 Parabolic  
maxint2 = Get maximum... 0 0 Parabolic
```

```
selectObject: cutsound3  
To Intensity... 75 0.001 true  
meanint3 = Get mean... 0 0  
meansdi3 = Get standard deviation... 0 0  
minint3 = Get minimum... 0 0 Parabolic  
maxint3 = Get maximum... 0 0 Parabolic
```

```
##### Obtenci n del tiempo de fonaci n
```



```
#####

select Sound 'name01$'
sound = selected ("Sound")
tmin = Get start time
tmax = Get end time
To Pitch: 0.001, 75, 300
Rename: "pitch"
selectObject: sound
To Intensity: 75, 0.001
Rename: "intensity"
#writeInfoLine: "Here_are_the_results:"
sum = 0
num = 0
tiempo = 0
for i to (tmax-tmin)/0.01
    time = tmin + i * 0.01
    selectObject: "Pitch_pitch"
    pitch = Get value at time: time, "Hertz", "linear"
    selectObject: "Intensity_intensity"
    intensity = Get value at time: time, "cubic"
    #appendInfoLine: fixed$ (time, 2), "_", fixed$ (pitch, 3)
        , "_", fixed$ (intensity, 3)
    if intensity = undefined
        num=num+1
    else
        sum=sum+intensity
    endif

    if intensity = undefined or pitch = undefined
        tiempo=tiempo+1
    endif
endifor

int_prom1 = sum/(((tmax-tmin)/0.01)-num)
fon1 = time - (tiempo*0.01)

select Sound 'name02$'
```

UCUENCA

```
sound = selected ("Sound")
tmin = Get start time
tmax = Get end time
To Pitch: 0.001, 75, 300
Rename: "pitch"
selectObject: sound
To Intensity: 75, 0.001
Rename: "intensity"
#writeInfoLine: "Here_are_the_results:"
sum = 0
num = 0
tiempo = 0
for i to (tmax-tmin)/0.01
    time = tmin + i * 0.01
    selectObject: "Pitch_pitch"
    pitch = Get value at time: time, "Hertz", "linear"
    selectObject: "Intensity_intensity"
    intensity = Get value at time: time, "cubic"
    #appendInfoLine: fixed$ (time, 2), "_", fixed$ (pitch, 3)
        , "_", fixed$ (intensity, 3)
    if intensity = undefined
        num=num+1
    else
        sum=sum+intensity
    endif

    if intensity = undefined or pitch = undefined
        tiempo=tiempo+1
    endif
endfor

int_prom2 = sum/(((tmax-tmin)/0.01)-num)
fon2 = time - (tiempo*0.01)

select Sound 'name03$'
sound = selected ("Sound")
tmin = Get start time
tmax = Get end time
```

UCUENCA

```
To Pitch: 0.001, 75, 300
Rename: "pitch"
selectObject: sound
To Intensity: 75, 0.001
Rename: "intensity"
#writeInfoLine: "Here_are_the_results:"
sum = 0
num = 0
tiempo = 0
for i to (tmax-tmin)/0.01
    time = tmin + i * 0.01
    selectObject: "Pitch_pitch"
    pitch = Get value at time: time, "Hertz", "linear"
    selectObject: "Intensity_intensity"
    intensity = Get value at time: time, "cubic"
    #appendInfoLine: fixed$ (time, 2), "_", fixed$ (pitch, 3)
        , "_", fixed$ (intensity, 3)
    if intensity = undefined
        num=num+1
    else
        sum=sum+intensity
    endif

    if intensity = undefined or pitch = undefined
        tiempo=tiempo+1
    endif
endifor

int_prom3 = sum/(((tmax-tmin)/0.01)-num)
fon3 = time - (tiempo*0.01)

##### Obtenci n de Jitter
#####

selectObject: cutsound1
To PointProcess (periodic, cc)... 75 500
meanlocal1 = Get jitter (local)... 0 0 0.0001 0.02 1.3

selectObject: cutsound2
```

```
To PointProcess (periodic, cc)... 75 500
meanlocal2 = Get jitter (local)... 0 0 0.0001 0.02 1.3

selectObject: cutsound3
To PointProcess (periodic, cc)... 75 500
meanlocal3 = Get jitter (local)... 0 0 0.0001 0.02 1.3

##### Obtenci n de Shimmer
#####

selectObject: cutsound1
plus PointProcess cutsound1
meanlocals1 = Get shimmer (local)... 0 0 0.0001 0.02 1.3 1.6

selectObject: cutsound2
plus PointProcess cutsound2
meanlocals2 = Get shimmer (local)... 0 0 0.0001 0.02 1.3 1.6

selectObject: cutsound3
plus PointProcess cutsound3
meanlocals3 = Get shimmer (local)... 0 0 0.0001 0.02 1.3 1.6

#'name3$'

##### Mediciones de la relaci n arm nico -ruido
#####

selectObject: cutsound1
To Harmonicity (cc)... 0.01 75 0.1 1
meanHNR_1 = Get mean... 0 0
meansdHNR_1 = Get standard deviation... 0 0

selectObject: cutsound2
To Harmonicity (cc)... 0.01 75 0.1 1
meanHNR_2 = Get mean... 0 0
meansdHNR_2 = Get standard deviation... 0 0
```

```
selectObject: cutsound3
To Harmonicity (cc)... 0.01 75 0.1 1
meanHNR_3 = Get mean... 0 0
meansdHNR_3 = Get standard deviation... 0 0

##### Mostrar par metros
#####

meanpitch = (meanpitch1 + meanpitch2 + meanpitch3)/3
meansd = (meansd1 + meansd2 + meansd3)/3
meanint = (meanint1 + meanint2 + meanint3)/3
meansdi = (meansdi1 + meansdi2 + meansdi3)/3
minpitch = (minpitch1 + minpitch2 + minpitch3)/3
maxpitch = (maxpitch1 + maxpitch2 + maxpitch3)/3
minint = (minint1 + minint2 + minint3)/3
maxint = (maxint1 + maxint2 + maxint3)/3
fon = (fon1 + fon2 + fon3)/3
jitt = ((meanlocal1 + meanlocal2 + meanlocal3)/3) *100
shimm = ((meanlocals1 + meanlocals2 + meanlocals3)/3) *100
hnr = (meanHNR_1 + meanHNR_2 + meanHNR_3)/3
hnr_sd = (meansdHNR_1 + meansdHNR_2 + meansdHNR_3)/3

appendInfoLine: "#####_Par metros_
#####"
appendInfoLine: "_____"
appendInfoLine: "La_frecuencia_fundamental_es:_", 'meanpitch'
, "_Hz"
appendInfoLine: "La_intensidad_es:_", 'meanint', "_dB"
appendInfoLine: "El_tiempo_de_fonacion_es:_", 'fon', "_s"
appendInfoLine: "El_jitter_local_es:_", 'jitt', "_%"
appendInfoLine: "El_shimmer_local_es:_", 'shimm', "_%"
appendInfoLine: "El_HNR_es:_", 'hnr', "_dB"

appendInfoLine: "_____"
appendInfoLine: "#####_Estad sticas_
#####"
appendInfoLine: "_____"
```

```
appendInfoLine: "La_desviación_estandar_de_la_frecuencia_
    fundamental_es:_", 'meansd'
appendInfoLine: "La_desviación_estandar_de_la_intensidad_
    :_", 'meansdi'
appendInfoLine: "El_valor_mínimo_de_la_frecuencia_
    fundamental_es:_", 'minpitch', "_Hz"
appendInfoLine: "El_valor_máximo_de_la_frecuencia_
    fundamental_es:_", 'maxpitch', "_Hz"
appendInfoLine: "El_valor_mínimo_de_la_intensidad_es:_", '
    minint', "_dB"
appendInfoLine: "El_valor_máximo_de_la_intensidad_es:_", '
    maxint', "_dB"
appendInfoLine: "La_desviación_estandar_del_HNR_es:_", '
    hnr_sd'
```

```
#agregan las variables creadas para los valores obtenidos en
    el archivo .csv
```

```
#fileappend "'directory$'Parametros.txt" 'meanpitch:10' , '
    meanint:10' , 'fon:10' , 'jitt:10' , 'shimm:10' , 'hnr:10'
    'newline$'
```

```
deleteFile ("datos.txt")
```

```
fileappend "datos.txt" 'sexo$''newline$''meanpitch:10''
    newline$''meanint:10''newline$''fon:10''newline$''jitt:10'
    'newline$''shimm:10''newline$''hnr:10''newline$''meansd:10
    ''newline$''meansdi:10''newline$''minpitch:10''newline$''
    maxpitch:10''newline$''minint:10''newline$''maxint:10''
    newline$''hnr_sd:10''newline$'
```

```
select all
```

```
Remove
```

C. Script de Python para análisis de parámetros

```
with open("datos.txt", "r") as tf:
```

```
sexo = tf.readline()
freq = float(tf.readline())
inte = float(tf.readline())
TMF = float(tf.readline())
jitter = float(tf.readline())
shimmer = float(tf.readline())
HNR = float(tf.readline())

if sexo == "F\n":
    if freq > 165 and freq < 264:
        res2 = "dentro_del_rango_normal"
    else:
        res2 = "fuera_del_rango_normal"

    if inte < 40:
        res3 = "fuera_del_rango_normal"
    else:
        res3 = "dentro_del_rango_normal"

    if TMF < 11:
        res4 = "debajo_del_promedio"
    else:
        res4 = "sobre_el_promedio"

    if jitter > 0.16 and jitter < 1:
        res5 = "dentro_del_rango_normal"
    else:
        res5 = "fuera_del_rango_normal"

    if shimmer > 1.02 and shimmer < 3.32:
        res6 = "dentro_del_rango_normal"
    else:
        res6 = "fuera_del_rango_normal"

    if HNR > 15.12 and HNR < 29.04:
        res7 = "dentro_del_rango_normal"
    else:
        res7 = "fuera_del_rango_normal"
else:
```

```
if freq > 87 and freq < 182:
    res2 = "dentro_del_rango_normal"
else:
    res2 = "fuera_del_rango_normal"

if inte < 40 :
    res3 = "fuera_del_rango_normal"
else:
    res3 = "dentro_del_rango_normal"

if TMF < 14:
    res4 = "por_debajo_del_promedio"
else:
    res4 = "sobre_el_promedio"

if jitter > 0.16 and jitter < 0.95:
    res5 = "dentro_del_rango_normal"
else:
    res5 = "fuera_del_rango_normal"

if shimmer > 0.88 and shimmer < 8.76:
    res6 = "dentro_del_rango_normal"
else:
    res6 = "fuera_del_rango_normal"

if HNR > 12.63 and HNR < 29.86:
    res7 = "dentro_del_rango_normal"
else:
    res7 = "fuera_del_rango_normal"

print("El analisis_vocal_del_individuo_indica_que:")
print("La_frecuencia_fundamental_est ",res2)
print("La_intensidad_est ",res3)
print("El_tiempo_maximo_de_fonacion_est ",res4)
print("El_jitter_est ",res5)
print("El_shimmer_est ",res6)
print("El_HNR_est ",res7)
```


D. Modelo de consentimiento informado

Consentimiento Informado

Cuenca, _____ de 2022

Yo _____, con cédula de identidad _____, en calidad de sujeto de estudio, doy mi consentimiento para participar en la validación del estudio titulado "Diseño e implementación de un dispositivo para monitoreo y apoyo al diagnóstico de fonoaudiología en pacientes post COVID". El estudio de investigación incluirá: recolección de datos y muestras.

Yo he leído la información anterior previamente, de la cual tengo una copia. Estoy de acuerdo en conceder **información personal verídica**. He tenido la oportunidad de hacer preguntas sobre la información y cada pregunta que he hecho ha sido respondida satisfactoriamente. He tenido el tiempo suficiente para leer y comprender los riesgos y beneficios de mi participación.

Doy mi consentimiento para la toma de una muestra de mi voz para análisis acústico con el dispositivo creado por Mauricio Esteban Guerrero Sarmiento y Paúl Vladimir Llerena Orellana estudiantes en calidad de egresados de la facultad de Ingeniería, carrera en electrónica y telecomunicaciones de la Universidad de Cuenca, cuya finalidad **será única y exclusivamente para fines académicos**; en adición, comprendo la obligación de censurar la identidad para mantener la confidencialidad.

Firma del participante del Estudio

Firma del Investigador 1

Firma del Investigador 2

Nombres de los investigadores: (Mauricio Esteban Guerrero Sarmiento y Paúl Vladimir Llerena Orellana)

E. Modelo de historia fonoaudiológica

Historia fonoaudiológica N° ...	Historia clínica	Fecha de aplicación Cuenca,..... de Agosto de 2022
	Valoración subjetiva de la voz	
	Valoración objetiva de la voz	
	Valoración de la eficiencia vocal	

1. Datos de filiación

Nombre completo			
Diagnóstico médico			
Edad		Género	
Profesión/Ocupación			
Informante/representante			

2. Motivo de Consulta

--

3. Antecedentes personales

Enfermedades pulmonares	Si: No: Tiempo:	Enfermedades gastrointestinales	Si: No: Tiempo:	Enfermedades neurológicas	Si: No: Tiempo:
Enfermedades autoinmunitarias	Si: No: Tiempo:	Enfermedades Endocrinas	Si: No: Tiempo:	Radioterapia	Si: No: Describir:
Cirugías prvevias	Si: No: Tipo:	Consumo de medicamentos	Si: No: Describir:	Hospitalizaciones	Si: No: Tipo:

4. Problemas ORL

Reflujo gastroesofágico	Si: No: Tiempo:	Rinitis Alérgica	Si: No: Tiempo:	hipoacusia	Si: No: Tiempo:
Otros:					

5. Hábitos

Consumo de tabaco	Si: No: Cantidad:	Consumo de alcohol	Si: No: Cantidad:	Consumo de estupefacientes	Si: No: Cantidad:
Ingesta de líquidos diarios	Si: No: Cantidad:	Tiempo de uso de la voz hablada y cantada	Aprox. Cantidad:	Consumo de café, teína, chocolate	Si: No: Cantidad:
Otros:					

Elaborado por:

Sofía Carolina Cedillo Ochoa. Licenciada en fonoaudiología por la Universidad de Cuenca & Estefanía Geovanna Mejía Macas. Egresada de licenciatura en fonoaudiología por la Universidad de Cuenca, Diplomada en Ciencia Habilitación y Rehabilitación de la voz humana por Vocology Center, Colombia.

Formato de calificación GRBAS

GRBAS

GRBAS	Traducción	Calificación
Grade	Grado global de alteración vocal o disfonía	
Rough	Irregularidad vocal , en frecuencia , amplitud o ambas	
Breathy	Fuga de aire a través de la glotis	
Asthenic	Debilidad , fatiga o falta de poder de la voz	
Strain	Hiperfunción o esfuerzo vocal	
Total		

Formato de calificación Voice Handicap Index -10

Descripción del perfil del hablante

Uso mi voz principalmente en

Mi profesión (maestro, abogado, sacerdote, vendedor, etc.)	
Actividades fuera del trabajo (asociaciones, teatro, ONG, etc.)	
Ninguna de las anteriores	

Uso mi voz cantada principalmente en

Mi profesión (cantante, estudiante de canto, profesor de música, etc.)	
Actividades fuera del trabajo (coros , cantante aficionado)	
Ninguna de las anteriores. No canto	

	Muy poco habladora	Normalmente habladora	Muy habladora
Se considera una persona			

Elaborado por:

Sofía Carolina Cedillo Ochoa. Licenciada en fonoaudiología por la Universidad de Cuenca & Estefanía Geovanna Mejía Macas. Egresada de licenciatura en fonoaudiología por la Universidad de Cuenca, Diplomada en Ciencia Habilitación y Rehabilitación de la voz humana por Vocology Center, Colombia.

		Nunca	Casi nunca	A veces	Casi siempre	Siempre
F1	La gente me oye con dificultad debido a mi voz					
F2	La gente no me entiende en sitios ruidosos					
F3	Mis problemas con la voz alteran mi vida personal y social					
F4	Me siento desplazado de las conversaciones por mi voz					
F5	Mi problema con la voz afecta al rendimiento laboral					
F6	Siento que necesito tensar la garganta para producir la voz					
F7	La calidad de mi voz es impredecible					
F8	Mi voz me molesta					
F9	Mi voz me hace sentir discapacitado					
F10	La gente me pregunta : "Que te pasa en la voz "					

Formato de calificación CAPE V

Consenso sobre la evaluación auditivo – perceptual de la voz

Los siguientes parámetros de la calidad vocal serán calificados tras completar las siguientes tareas vocales:

Tarea 1: Vocal sostenida /a/ /e/ /i/ durante de 3 a 5 segundos

Tarea 2: Oraciones: (a) Nuria ojeo una pajarita y una blusa amarilla. (b) Marta multa mi moto más mágica. (c) La llave brilla en la mano. (d) Irene adora hacer huevos al horno. (e) Mamá me mimó una mano. (f) Ala tu zapato a tu pata

Severidad global	Puntuación
	_____/100

Elaborado por:

Sofía Carolina Cedillo Ochoa. Licenciada en fonoaudiología por la Universidad de Cuenca & Estefanía Geovanna Mejía Macas. Egresada de licenciatura en fonoaudiología por la Universidad de Cuenca, Diplomada en Ciencia Habilitación y Rehabilitación de la voz humana por Vocology Center, Colombia.

Aspereza	
LE _____ MO _____ SE	_____/100
Escape	
LE _____ MO _____ SE	_____/100
Tensión	
LE _____ MO _____ SE	_____/100
Tono : Indicar la naturaleza de la desviación _____	
LE _____ MO _____ SE	_____/100
Intensidad : Indicar la naturaleza de la desviación _____	
LE _____ MO _____ SE	_____/100
Característica adicional	
1: LE _____ MO _____ SE	_____/100
Característica adicional	
2: LE _____ MO _____ SE	_____/100
Comentarios sobre la nasalidad : normal otros:	
Describir :	
Características adicionales : Diplofonía , vocal fry , falsetto , astenia , afonía, inestabilidad del tono , temblor , fonación húmeda , otros términos relvantes :	

Elaborado por:

Sofía Carolina Cedillo Ochoa. Licenciada en fonoaudiología por la Universidad de Cuenca & Estefanía Geovanna Mejía Macas. Egresada de licenciatura en fonoaudiología por la Universidad de Cuenca, Diplomada en Ciencia Habilitación y Rehabilitación de la voz humana por Vocology Center, Colombia.

Valoración objetiva de la voz

Análisis Acústico de la Voz

Valoración objetiva

Frecuencia Fundamental Hz
Intensidad dB
Medidas de perturbación	Jitter local ... %
	Jitter ppq5 ... %
	Shimmer local ...%
	Shimmer appq3 ...%
Indices de armonía - ruido	Armónico ruido ... dB
	Ruido Armónico ... dB

Valoración de la eficiencia vocal

Tiempo máximo de fonación

1era Emisión:seg	2da Emisión:seg	3era Emisión:seg
-----------------------	----------------------	-----------------------

- Comentarios

- Impresiones y recomendaciones

- Evaluador: Nombre y firma

Lic. Sofia Cedillo
Licenciada en fonoaudiología

Estefanía Mejía
Egresada de licenciatura en
fonoaudiología

Elaborado por:

Sofia Carolina Cedillo Ochoa. Licenciada en fonoaudiología por la Universidad de Cuenca & Estefanía Geovanna Mejía Macas. Egresada de licenciatura en fonoaudiología por la Universidad de Cuenca, Diplomada en Ciencia Habilitación y Rehabilitación de la voz humana por Vocology Center, Colombia.

F. Informe fonoaudiológico general

Cuenca, 4 de Agosto de 2022

Informe Fonoaudiológico General

Mediante el presente documento, se informa el procedimiento fonoaudiológico realizado para el proyecto de investigación titulado “Diseño e implementación de un dispositivo electrónico para monitoreo y apoyo al diagnóstico de fonoaudiología en pacientes post COVID” previo a la obtención del título Ingenieros en electrónica y telecomunicaciones, de la Universidad de Cuenca.

Reactivos Utilizados

Reactivo	Uso clínico
Historia clínica	Recolección de información personal, antecedentes personales de salud y hábitos que permiten hacer predicciones del estado de salud del paciente.
GRABS	Parámetros para medir la cualidad vocal percibida. Propuesto por Hirano, 1981 desarrollado en la sociedad japonesa de logopedia y foniatría. Cuantifica del 0 al 4, según sus siglas: Grado de disfonía, Rugosidad, Astenia, Soplosidad y Tensión.
Cape - V	Parámetros para medir la cualidad vocal percibida. Propuesta por la Asociación Americana de Habla y Lenguaje. Mide parámetros de: severidad global, ronquera, escape aéreo, tensión, tono e intensidad.
VHI- 10	Descrito por Jacobson et al, es una versión simplificada del VHI-30, que mide parámetros físicos, emocionales y funcionales de la autopercepción de la disfonía y la calidad de vida del paciente que la padece.
Praat	Software para realizar análisis acústico de la voz, mide parámetros de: frecuencia fundamental, intensidad, medidas de perturbación e índice de armonía y ruido. Su utilidad radica en el monitoreo de la eficacia antes y después del tratamiento; a demás complementa al diagnóstico de imagen realizado por el médico laringólogo especialista.

Las muestras de voz obtenidas para el análisis fueron tomadas en un estudio de grabación con cámara insonorizada, con un micrófono con condensador de marca Beta4-080. Realizando las siguientes tareas fonatorias.

1. Vocales sostenidas /a/ , /i/, después de emitir la frase “<esta, es una cabina>” y “<esta es una cabi>”
2. Frases del formator Cape-V, diseñadas para que aparezcan todas las vocales, con ataque vocal suave y duro, sonidos nasales y consonantes plosivas sordas.
3. Habla espontánea

Comentarios, e impresión diagnóstica según los resultados obtenidos de los reactivos aplicados.

Ind	Comentarios	Diagnóstico presuntivo
1	<p>Sujeto de estudio de sexo masculino refiere rinitis alérgica ocasional y actualmente cursa con un episodio de laringitis viral aguda. En cantidad inespecificada ingiere irritantes para la mucosa laríngea como teína, cafeína, etc.</p> <p>Según las cualidades de su voz mantiene un grado de alteración vocal leve, tiene una autopercepción normal de su voz; el TMF se encuentra disminuido secundario a una disfonía leve; los parámetros de análisis acústico: frecuencia fundamental, intensidad, medidas de perturbación e índices de armonía se encuentran dentro de los parámetros normales, excepto el índice ruido- armónico que se encuentra elevado.</p>	<p>Disfonía de base funcional, se deriva a ORL para estudio estroboscópico</p>
2	<p>Sujeto de estudio de sexo masculino no refiere diagnóstico médicos anteriores, consumo de fármacos o cirugías recientes. En cantidad inespecificada ingiere irritantes para la mucosa laríngea como teína, cafeína, etc.</p> <p>Según las cualidades de su voz no tiene alteración, tiene una autopercepción normal de su voz; el TMF se encuentra disminuido; los parámetros de análisis acústico: frecuencia fundamental, intensidad, medidas</p>	<p>Eufonía</p>

	de perturbación e índices de armonía y ruido se encuentran dentro de los parámetros normales.	
3	<p>Sujeto de estudio de sexo masculino no refiere diagnóstico médicos anteriores, consumo de fármacos o cirugías recientes. En cantidad inespecificada ingiere irritantes para la mucosa laríngea como teína, cafeína, etc. Utiliza su voz cantada de manera aficionada.</p> <p>Según las cualidades de su voz mantiene un grado de alteración vocal leve, tiene una autopercepción normal de su voz y se considera como una persona muy habladora; el TMF se encuentra dentro de los parámetros normales sin embargo al final de las emisiones se percibe vocal fry; los parámetros de análisis acústico: frecuencia fundamental, intensidad, medidas de perturbación e índices de armonía y ruido se encuentran dentro de los parámetros normales.</p>	Disfonía de base funcional, se deriva a ORL para estudio estroboscópico
4	<p>Sujeto de estudio de sexo femenino refiere ovario poliquístico en tratamiento farmacológico, actualmente cursa con un episodio de laringitis viral aguda. No ingiere irritantes para la mucosa laríngea como teína, cafeína, etc.</p> <p>Según las cualidades de su voz no tiene alteración, tiene una autopercepción normal de su voz, se considera como una persona normalmente habladora; el TMF se encuentra dentro de los parámetros normales; los parámetros de análisis acústico: frecuencia fundamental, intensidad, medidas de perturbación e índices de armonía y ruido se encuentran dentro de los parámetros normales.</p>	Eufonía
5	Sujeto de estudio de sexo masculino no refiere diagnóstico médicos anteriores, consumo de fármacos o cirugías recientes. En cantidad inespecificada	Disfonía de base funcional, se deriva a ORL para estudio estroboscópico

	<p>ingiere irritantes para la mucosa laríngea como teína, cafeína, alcohol, etc.</p> <p>Según las cualidades de su voz mantiene un grado de alteración vocal moderada, tiene una autopercepción de irregular de su voz y se considera como una persona normalmente habladora; el TMF se encuentra dentro de los parámetros normales; los parámetros de análisis acústico: frecuencia fundamental, intensidad, medidas de perturbación e índices de armonía y ruido se encuentran dentro de los parámetros normales.</p>	
6	<p>Sujeto de estudio de sexo masculino refiere rinitis alérgica tratada con terapia farmacológica y actualmente cursa con un episodio de laringitis viral aguda. En cantidad inespecificada ingiere irritantes para la mucosa laríngea como teína, cafeína, etc.</p> <p>Según las cualidades de su voz mantiene un grado de alteración vocal leve, tiene una autopercepción normal de su voz; el TMF se encuentra disminuido secundario a una disfonía leve; los parámetros de análisis acústico: frecuencia fundamental, intensidad, medidas de perturbación e índices de armonía y ruido se encuentran dentro de los parámetros normales.</p>	<p>Disfonía de base funcional, se deriva a ORL para estudio estroboscópico</p>
7	<p>Sujeto de estudio de sexo masculino no refiere diagnóstico médicos anteriores, consumo de fármacos o cirugías recientes. En cantidad inespecificada ingiere irritantes para la mucosa laríngea como teína, cafeína, etc.</p> <p>Según las cualidades de su voz no tiene alteración, tiene una autopercepción normal de su voz y se considera una persona normalmente habladora; el TMF se encuentra disminuido; los parámetros de análisis acústico: frecuencia fundamental, intensidad,</p>	<p>Eufonía</p>

	medidas de perturbación e índices de armonía y ruido se encuentran dentro de los parámetros normales.	
8	<p>Sujeto de estudio de sexo femenino no refiere diagnóstico médicos anteriores, consumo de fármacos o cirugías recientes. En cantidad inespecificada ingiere irritantes para la mucosa laríngea como teína, cafeína, alcohol, etc. Utiliza su voz de manera ocupacional.</p> <p>Según las cualidades de su voz mantiene un grado de alteración vocal moderada, tiene una autopercepción de irregular de su voz y se considera como una persona muy habladora; el TMF no se encuentra dentro de los parámetros normales secundario a disfonía; los parámetros de análisis acústico: frecuencia fundamental, intensidad, medidas de perturbación e índices de armonía y ruido se encuentran dentro de los parámetros normales.</p>	Disfonía de base funcional, se deriva a ORL para estudio estroboscópico
9	<p>Sujeto de estudio de sexo femenino no refiere diagnóstico médicos anteriores, consumo de fármacos o cirugías recientes. En cantidad inespecificada ingiere irritantes para la mucosa laríngea como teína, cafeína, alcohol, etc. Utiliza su voz de manera ocupacional.</p> <p>Según las cualidades de su voz mantiene un grado de alteración vocal leve, tiene una autopercepción de irregular de su voz y se considera como una persona normalmente habladora; el TMF no se encuentra dentro de los parámetros normales secundario a disfonía; los parámetros de análisis acústico: frecuencia fundamental, intensidad, medidas de perturbación e índices de armonía y ruido se encuentran dentro de los parámetros normales.</p>	Disfonía de base funcional, se deriva a ORL para estudio estroboscópico

10	<p>Sujeto de estudio de sexo femenino refiere rinoplastía.</p> <p>En cantidad inespecificada ingiere irritantes para la mucosa laríngea como teína, cafeína, alcohol, etc.</p> <p>Según las cualidades de su voz mantiene un grado de alteración vocal normal, tiene una autopercepción de normal de su voz y se considera como una persona normalmente habladora; el TMF no se encuentra dentro de los parámetros normales secundario a respiración oral, se percibe voz hipernasal de tipo cerrada; los parámetros de análisis acústico: frecuencia fundamental, intensidad, medidas de perturbación e índices de armonía se encuentran dentro de los parámetros normales, excepto el índice de ruido</p>	Eufonía
----	---	---------

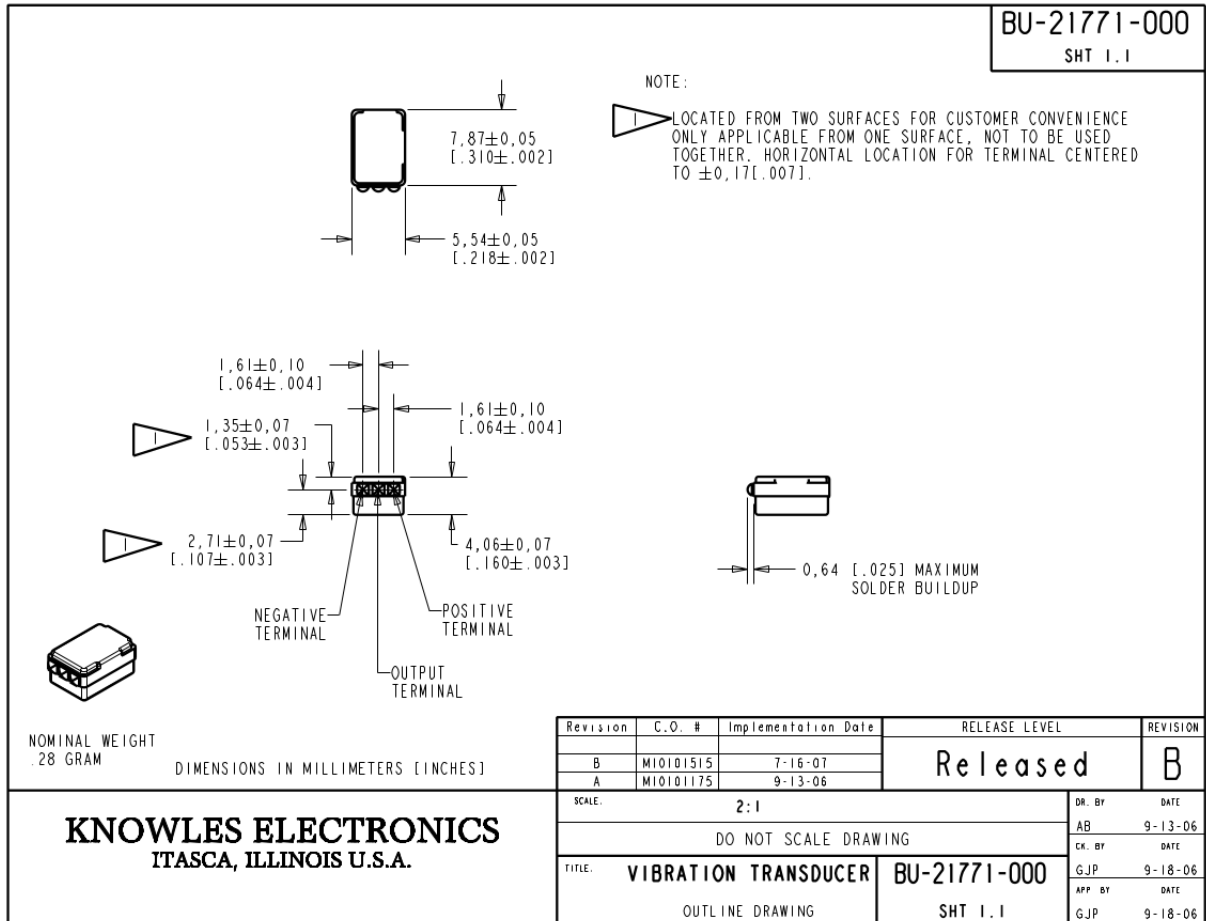
- Evaluador: Nombre y firma

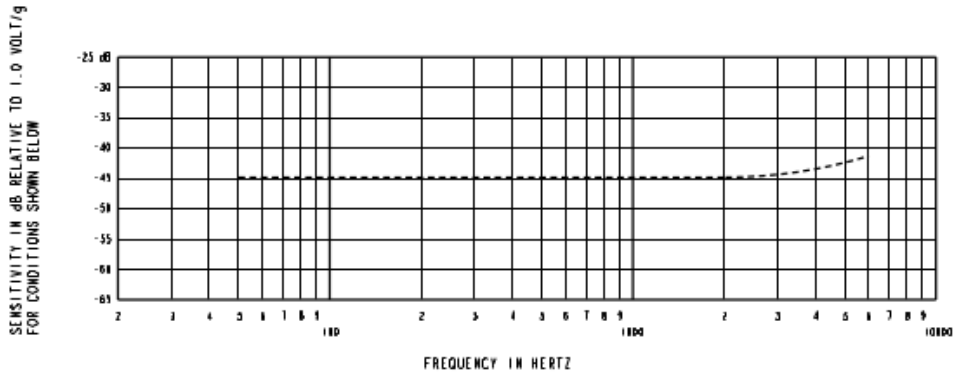
Lic. Sofia Cedillo
Licenciada en fonoaudiología

Estefanía Mejía
Egresada de licenciatura en
fonoaudiología

G. Datasheet

G.1. Acelerómetro BU-21771





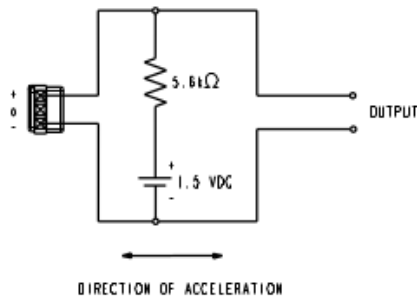
FREQUENCY	SENSITIVITY			DEVICE CONFORMITY	
	MIN	NOM	MAX	RANGE OF DEVIATION FROM 1 kHz	
300	---	-45.0	---	-1.5	+1.5
1000	-49.5	-45.0	-40.5	0	0
3000	---	-44.5	---	-1.0	+2.0

NOTES:

1. ALL DATA MEASURED WITH CIRCUIT CONFIGURATION SHOWN BELOW.
2. CASE CONNECTED TO NEGATIVE TERMINAL.

DC SUPPLY	BATTERY CURRENT DRAIN	"A" WEIGHTED NOISE (RE 1.0V)	OUTPUT IMPEDANCE OHMS (AT 1 kHz)		
			MIN.	NOM.	MAX.
1.5 V	50 μ A MAX.	-103 dB MAX.	4900	5200	5500

3. TEST CONDITIONS:

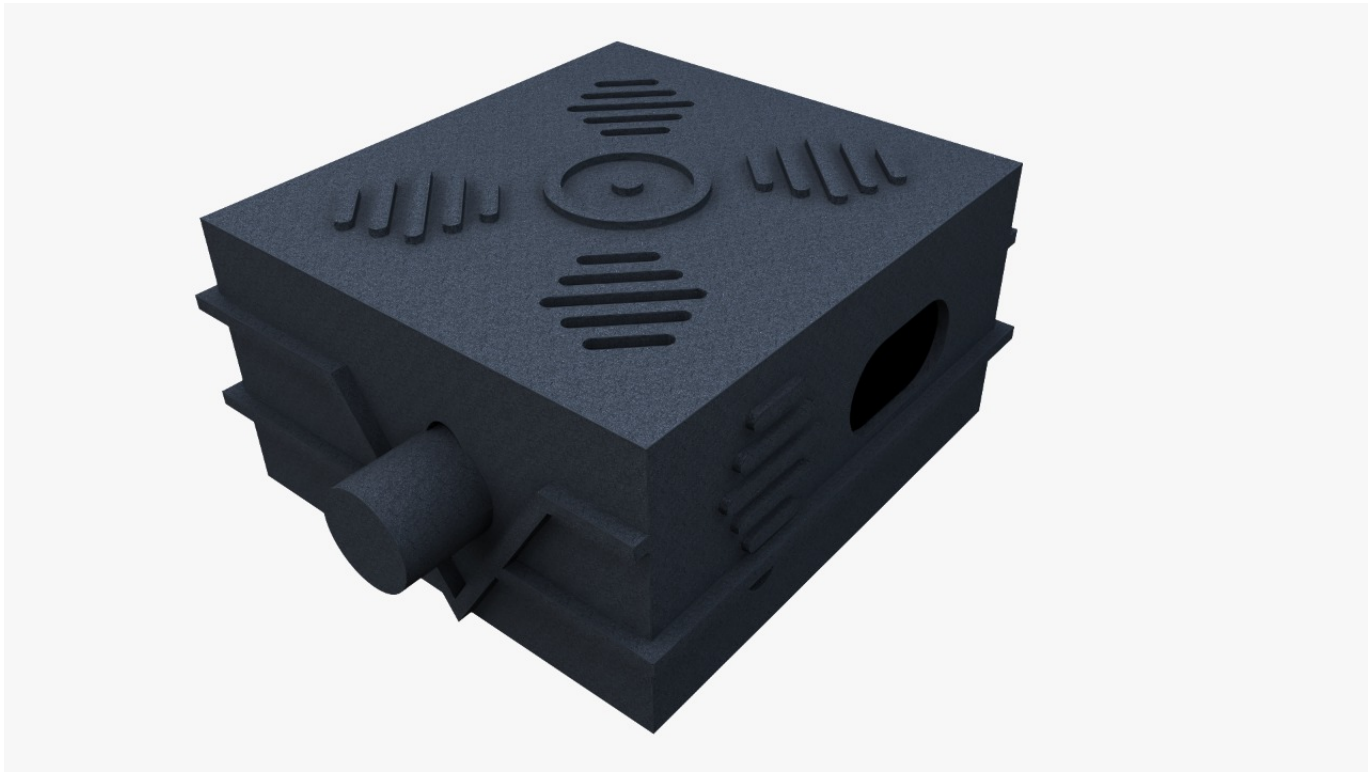


KNOWLES ELECTRONICS
ITASCA, ILLINOIS U.S.A.

Revision	C.O.D	Implementation Date	RELEASE LEVEL	REVISION
B	N10101513	7-16-07	Released	B
A	N1010113	9-13-06		
WHEN TEST LIMITS ARE USED TO ESTABLISH INCOMING INSPECTION ACCEPTANCE/REJECTION CRITERIA, CORRELATION OF TEST EQUIPMENT WITH KNOWLES IS ALSO REQUIRED FOR ELIMINATION OF EQUIPMENT AND TEST METHOD VARIATION			DR DT	DATE
THE VIBRATION TRANSDUCER			BU-21771-000	AB
PERFORMANCE SPECIFICATION			SHT 2.1	9-13-06
			DR DT	DATE
			GJP	9-18-06
			DR DT	DATE
			GJP	9-18-06

H. Diseños

H.1. Caja para circuito



H.2. Sticker 1



H.3. Sticker 2

**Guerrero Sarmiento Mauricio Esteban
Llerena Orellana Paúl Vladimir**

Bibliografía

- [1] G. A. Alzamendi, “Modelado estocástico de la fonación y señales biomédicas relacionadas. Métodos en espacio de estados aplicados al análisis estructural, al modelado de la fonación y al filtrado inverso,” jun 2016. [En línea]. Disponible: <https://bibliotecavirtual.unl.edu.ar/handle/11185/844>
- [2] G. J. Tortora y B. Derrickson, “Principios de Anatomía y Fisiología 15a edición,” *Panamericana*, p. 1236, 2018.
- [3] J. Jiang, E. Lin, y D. G. Hanson, “VOCAL FOLD PHYSIOLOGY,” *Otolaryngologic Clinics of North America*, vol. 33, num. 4, pp. 699–718, aug 2000.
- [4] C. Rivera, “¿Cómo se produce la voz?” [En línea]. Disponible: <https://www.doctorchozas.es/noticias/82-como-se-produce-la-voz>
- [5] “Tubos Endotraqueales y Endobronquiales – Alcimars Medic.” [En línea]. Disponible: <https://www.alcimarsmedic.com/materiales-e-instrumentos-quirurgicos/tubos-endotraqueales-y-endobronquiales/>
- [6] Comaudi, s/f. [En línea]. Disponible: <https://www.comaudi.com/anatomia-del-oido/>
- [7] J. M. M. de la Fuente y L. Muñoz-Repiso, “La percepción acústica: física de la audición,” *Revista de ciencias*, num. 2, pp. 19–26, 2013.
- [8] “Velocidad del sonido.” [En línea]. Disponible: <https://cnx.org/contents/kb7gVsTk@7/Velocidad-del-sonido>
- [9] SEORL, *Evaluación del paciente con disfonía*. IMC, 2018.
- [10] E. Celdrán, “Naturaleza fonética de la consonante ‘ye’ en español,” *Normas*, 08 2015.
- [11] S. Ripoll Gimeno y D. Joan Martínez Mora, “.Evolución de la contaminación acústica provocada por el tráfico de la N-332 en.”
- [12] “Esp 32 Modulo Wifi- Bluetooth Esp32 30 Pines - MEGATRONICA.” [En línea]. Disponible: <https://megatronica.cc/producto/esp-32-modulo-wifi-bluetooth-esp32-30-pines/>
- [13] “Esp 32 Modulo Wifi- Bluetooth Esp32 30 Pines - MEGATRONICA.” [En línea]. Disponible: <https://megatronica.cc/producto/esp-32-modulo-wifi-bluetooth-esp32-30-pines/>

- [14] “Modulo Lector Micro SD.” [En línea]. Disponible: http://www.arduinove.com/index.php?route=product/product&product_id=194
- [15] “BU-21771-000 Knowles — Mouser Ecuador.” [En línea]. Disponible: <https://www.mouser.ec/ProductDetail/Knowles/BU-21771-000?qs=3unH%2FDqlvl%252BPeVBYA0xFIQ%3D%3D>
- [16] M. Nolan, B. Madden, y E. Burke, “Accelerometer based measurement for the mapping of neck surface vibrations during vocalized speech,” in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2009, pp. 4453–4456.
- [17] K. A. Coutts, “Dysphagia services in the era of COVID-19: Are speech-language therapists essential?” *South African Journal of Communication Disorders*, vol. 67, num. 1, pp. 1–6, 2020.
- [18] S. F. González, M. M. Girbau, F. V. de la Iglesia, y C. G. Vilar, “Discapacidad de vocal,” *Revista de Medicina de la Universidad de Navarra*, pp. 73–80, 2006.
- [19] A. Biddle, L. Watson, y C. Hooper, “Using evidence-based research in speech-language pathology. a paper presented to the meeting of the american speech,” *Language, and Hearing Association, Atlanta, GA*, 2002.
- [20] P. Farias, “Mejoría del cierre glótico y de la onda mucosa con terapia vocal en sulcus vergeture: reporte de un caso,” *Areté*, num. 14, pp. 6–19, 2014.
- [21] M. Guzmán, “Evaluación funcional de la voz,” *Santiago, Chile*, 2010.
- [22] M. S. A. Moreno Méndez, M. Á. Tenorio, M. Alejandra, B. Páez, C. Alejandra, y P. Garzón, “PARÁMETROS ACÚSTICOS DE LA VOZ EN EL ADULTO MAYOR.”
- [23] C. C. Gimeno-Vilar, F. F. Vázquez-de-la Iglesia, M. M. Marques-Girbau, y S. S. Fernández-González, “Discapacidad vocal,” *REV MED UNIV NAVARRA*, vol. 50, num. 3, pp. 73–80, 2006. [En línea]. Disponible: <https://dadun.unav.edu/handle/10171/35899>
- [24] R. D. Tema, C. Heredia Lima, P. Ruiz Gallo Lambayeque, P.édico Infectólogo@, D. Profesor principal, L.ía Microbiología Parasitología@, C.ña Vargas@, R. Gastelo Acosta, y A. Tequen Bernilla, “El nuevo Coronavirus y la pandemia del Covid-19,” *Revista Medica Herediana*, vol. 31, num. 2, pp. 125–131, jul 2020. [En línea]. Disponible: http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1018-130X2020000200125&lng=es&nrm=iso&tlng=eshttp://www.scielo.org.pe/scielo.php?script=sci_abstract&pid=S1018-130X2020000200125&lng=es&nrm=iso&tlng=es

- [25] “Enfermedad del coronavirus 2019 (COVID-19) - Síntomas y causas - Mayo Clinic.” [En línea]. Disponible: <https://www.mayoclinic.org/es-es/diseases-conditions/coronavirus/symptoms-causes/syc-20479963#>
- [26] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu y otros, “Clinical features of patients infected with 2019 novel coronavirus in wuhan, china,” *The lancet*, vol. 395, num. 10223, pp. 497–506, 2020.
- [27] M. J. Rouhani, G. Clunie, G. Thong, L. Lovell, J. Roe, M. Ashcroft, A. Holroyd, G. Sandhu, y C. Al Yaghchi, “A Prospective Study of Voice, Swallow, and Airway Outcomes Following Tracheostomy for COVID-19,” *The Laryngoscope*, vol. 131, num. 6, pp. E1918–E1925, jun 2021. [En línea]. Disponible: <https://onlinelibrary.wiley.com/doi/full/10.1002/lary.29346><https://onlinelibrary.wiley.com/doi/abs/10.1002/lary.29346><https://onlinelibrary.wiley.com/doi/10.1002/lary.29346>
- [28] A. S. Freitas, G. M. Zica, y C. L. De Albuquerque, “Pandemia de coronavirus (COVID-19): o que os fonoaudiólogos devem saber,” *CoDAS*, vol. 32, num. 3, jun 2020. [En línea]. Disponible: <http://www.scielo.br/j/codas/a/GgQqqrRtJHw4vDX5YfgJBrs/?lang=pt>
- [29] J. C. Sousa, “PREVALÊNCIA DE SINTOMAS VOCAIS EM PACIENTES PÓS COVID-19: REVISÃO DE LITERATURA,” dec 2021. [En línea]. Disponible: <https://repositorio.pucgoias.edu.br/jspui/handle/123456789/2922>
- [30] M. Behlau, *Voz: o livro do especialista: volume II*. Revinter, 2005.
- [31] A. de Melo Cesar y M. D. Lima, *Fonoaudiologia e COVID-19: Guia de Intervenção*. Thieme Revinter, 2021.
- [32] A. S. y J. M. V. Graciela Oliva. (2013, dec) Herramientas de evaluación logopédica por patologías. [En línea]. Disponible: <https://www.clc.cat/pdf/publicacions/documents/es/HERRAMIENTASdeEvaluacionDeLaVoz.pdf>
- [33] J. H. Van Stan, J. Gustafsson, E. Schalling, y R. E. Hillman, “Direct Comparison of Three Commercially Available Devices for Voice Ambulatory Monitoring and Biofeedback.” [En línea]. Disponible: <http://sig3perspectives.pubs.asha.org/>
- [34] M. Rivas Reyes, M. A. Bastanzuri Rivas, y M. Olivera ValdÃ, “El cuidado de la voz en la actividad docente,” *Revista Habanera de Ciencias MÃ*, vol. 12, pp. 74 – 81, 00 2013. [En línea]. Disponible: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1729-519X2013000500010&nrm=iso

- [35] P. S. Popolo, J. G. Švec, y I. R. Titze, “Adaptation of a Pocket PC for Use as a Wearable Voice Dosimeter,” *Journal of Speech, Language, and Hearing Research*, vol. 48, num. 4, pp. 780–791, aug 2005. [En línea]. Disponible: <https://pubs.asha.org/doi/10.1044/1092-4388%282005/054%29>
- [36] H. A. Cheyne, H. M. Hanson, R. P. Genereux, K. N. Stevens, y R. E. Hillman, “Development and Testing of a Portable Vocal Accumulator,” *Journal of Speech, Language, and Hearing Research*, vol. 46, num. 6, pp. 1457–1467, dec 2003. [En línea]. Disponible: [https://pubs.asha.org/doi/10.1044/1092-4388\(2003/113\)](https://pubs.asha.org/doi/10.1044/1092-4388(2003/113))
- [37] J. H. Van Stan, D. D. Mehta, S. M. Zeitels, J. A. Burns, A. M. Barbu, y R. E. Hillman, “Average ambulatory measures of sound pressure level, fundamental frequency, and vocal dose do not differ between adult females with phonotraumatic lesions and matched control subjects,” *Annals of Otolaryngology, Rhinology and Laryngology*, vol. 124, num. 11, pp. 864–874, nov 2015. [En línea]. Disponible: <https://journals.sagepub.com/doi/10.1177/0003489415589363>
- [38] B. Ayribas, “Cognitive Difficulties, Somatic Complaints and Anxiety: Psychological Complications in COVID-19 Survivors,” *Clinical Psychiatry*, vol. 7, num. 0, p. 9, nov 2021. [En línea]. Disponible: <https://clinical-psychiatry.imedpub.com/abstract/cognitive-difficulties-somatic-complaints-and-anxiety-psychological-complications-in-covid19-survivors-41476.html>
- [39] A. Castillo-Allendes, F. Contreras-Ruston, L. Cantor, J. Codino, M. Guzman, C. Malebran, C. Manzano, A. Pavez, T. Vaiano, F. Wilder, y M. Behlau, “Terapia de voz en el contexto de la pandemia covid-19; recomendaciones para la práctica clínica,” *Journal of Voice*, vol. 35, num. 5, p. 808.e1, sep 2021. [En línea]. Disponible: <https://pubmed.ncbi.nlm.nih.gov/35442931/>
- [40] A. Crespo y G. Korn, “State-of-the-art in professional voice,” *Brazilian Journal of Otorhinolaryngology (English Edition)*, oct 2021. [En línea]. Disponible: <http://www.bjorl.org/en-state-of-the-art-in-professional-voice-avance-S1808869421001695http://www.bjorl.org/en-state-of-the-art-in-professional-voice-avance-S1808869421001695?ref=busqueda&ant=S2530380521000228&sig=S021361632100063X>
- [41] H. Aljehani, J. H. Van Stan, C. W. Haynes, y D. D. Mehta, “Ambulatory voice monitoring of a muslim imam during ramadan,” *Proceedings of the Voice Foundation Symposium*, 2015.

- [42] R. C. Oliveira, A. C. Gama, y M. D. Magalhães, “Fundamental voice frequency: acoustic, electroglottographic, and accelerometer measurement in individuals with and without vocal alteration,” *Journal of Voice*, vol. 35, num. 2, pp. 174–180, 2021.
- [43] P. C. Chwalek, D. D. Mehta, B. Welsh, C. Wooten, K. Byrd, E. Froehlich, D. Maurer, J. Lacirignola, T. F. Quatieri, y L. J. Brattain, “Lightweight, on-body, wireless system for ambulatory voice and ambient noise monitoring,” in *2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 2018, pp. 205–209.
- [44] J. Searl, A. Dietsch, y K. City, “Testing of the VocaLog Vocal Monitor,” 2014. [En línea]. Disponible: <http://dx.doi.org/10.1016/j.jvoice.2014.01.009>
- [45] J. Searl, A. M. Dietsch, y K. City, “Tolerance of the VocaLogtrade; Vocal Monitor by Healthy Persons and Individuals With Parkinson Disease,” 2015. [En línea]. Disponible: <http://dx.doi.org/10.1016/j.jvoice.2014.09.011>
- [46] M. Behlau, “Voz : o livro do especialista.” 2001.
- [47] J.íaz@, C. Sapienza, H. B. Rothman, y Y. Natour, “. INTRODUCCIÓN Algoritmo robusto para la detección de la frecuencia fundamental de la voz basado en el espectrograma,” *REVISTA INGENIERÍA UC*, vol. 10, pp. 7–16, 2003.
- [48] C. García y D. Tapias, “LA FRECUENCIA FUNDAMENTAL DE LA VOZ Y SUS EFECTOS EN RECONOCIMIENTO DE HABLA CONTINUA,” 2021.
- [49] W. C. A. COSTA, “Análise dinâmica não linear de sinais de voz para detecção de patologias laríngeas.” nov 2012. [En línea]. Disponible: <http://dspace.sti.ufcg.edu.br:8080/jspui/handle/riufcg/1416>
- [50] E. Velasco, “Patologías más comunes de la garganta,” *El farmacéutico*, 2020. [En línea]. Disponible: <https://www.elfarmacéutico.es/uploads/s1/22/50/ef584-profesion-patologias-garganta.pdf>
- [51] M. B. Brodsky, M. J. Levy, E. Jedlanek, V. Pandian, B. Blackford, C. Price, G. Cole, A. T. Hillel, S. R. Best, y L. M. Akst, “Laryngeal injury and upper airway symptoms after oral endotracheal intubation with mechanical ventilation during critical care: A systematic review,” *Critical Care Medicine*, vol. 46, num. 12, pp. 2010–2017, 2018. [En línea]. Disponible: https://journals.lww.com/ccmjournal/Fulltext/2018/12000/Laryngeal_Injury_and_Upper_Airway_Symptoms_After.15.aspx
- [52] M. Asiaee, A. Vahedian-azimi, S. S. Atashi, A. Keramatfar, y M. Nourbakhsh, “Voice Quality Evaluation in Patients With COVID-19: An Acoustic Analysis,” *Journal of Voice*, oct 2020.

- [53] V. Despotovic, M. Ismael, M. Cornil, R. M. Call, y G. Fagherazzi, “Detection of COVID-19 from voice, cough and breathing patterns: Dataset and preliminary results,” *Computers in Biology and Medicine*, vol. 138, p. 104944, nov 2021.
- [54] O. Alberto y C. Treviño, “Intubación endotraqueal: importancia de la presión del manguito sobre el epitelio traqueal Endotracheal intubati,” 2013. [En línea]. Disponible: www.elsevier.es
- [55] G. P. López-Herranz, “Intubación endotraqueal: importancia de la presión del manguito sobre el epitelio traqueal,” *Revista Médica del Hospital General de México*, vol. 76, num. 3, pp. 153–161, jul 2013. [En línea]. Disponible: <https://www.elsevier.es/en-revista-revista-medica-del-hospital-general-325-articulo-intubacion-endotraqueal-importancia-presion-del-X0185106313493650>
- [56] P. Zamora Porras, “COMPLICACIONES DE INTUBACIÓN TRANS-LARÍNGEA PROLONGADA N E U M O L O G Í A REVISTA MEDICA DE COSTA RICA Y CENTROAMERICA LXX (605).”
- [57] F. G. Puente y R. R. González, “Secuelas laringotraqueales post intubación endotraqueal prolongada. Laringotracheal Sequelar post a long-term endotracheal intubation,” *Revista de Ciencias Médicas de Pinar del Río*, vol. 2, num. 2, pp. 49–57, 1998. [En línea]. Disponible: <http://www.revcmpinar.sld.cu/index.php/publicaciones/article/view/49/95><http://www.revcmpinar.sld.cu/index.php/publicaciones/article/view/49>
- [58] F. De Medicina, M. C. Idalia, A. Mares, E. En Orl, C. David, N. Fernandez, C. Briscia, E. Delgado, S. E. En Orl, C. Yazmin, y O. Suarez, “UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO HOSPITAL REGIONAL TLALNEPANTLA PARA OBTENER EL DIPLOMA DE POSGRADO DE LA ESPECIALIDAD: OTORRINOLARINGOLOGÍA PRESENTA.”
- [59] Y. G. Droguett, “Aplicaciones clínicas del análisis acústico de la voz,” *Revista de otorrinolaringología y cirugía de cabeza y cuello*, vol. 77, num. 4, pp. 474–483, 2017.
- [60] J. González, T. Cervera, y J. L. Miralles, “Análisis acústico de la voz: fiabilidad de un conjunto de parámetros multidimensionales,” *Acta otorrinolaringológica española*, vol. 53, num. 4, pp. 256–268, 2002.
- [61] I. R. Titze y K. V. Abbott, *Vocology: The science and practice of voice habilitation*. National Center for Voice and Speech, 2012.

- [62] J. P. Teixeira, D. Ferreira, y S. M. Carneiro, “Análise acústica vocal-determinação do jitter e shimmer para diagnóstico de patologias da fala,” in *6º Congresso Luso-Moçambicano de Engenharia, 3º Congresso de Engenharia de Moçambique*, num. 6º. INEGI, 2011.
- [63] H. Reymond, J. K. Colton, y L. R. Casper, “Understanding voice problems,” *A Physiological Perspective For Diagnosis and Treatment 3rd Ed Lippincott Williams and Wilkins*, 2006.
- [64] D. D. Mehta, J. H. Van Stan, M. Zañartu, M. Ghassemi, J. V. Guttag, V. M. Espinoza, J. P. Cortés, H. A. Cheyne, y R. E. Hillman, “Using ambulatory voice monitoring to investigate common voice disorders: Research update,” *Frontiers in bioengineering and biotechnology*, vol. 3, p. 155, 2015.
- [65] E. R. Carrasco, G. Oliveira, y M. Behlau, “Análise perceptivo-auditiva e acústica da voz de indivíduos gagos,” *Revista CEFAC*, vol. 12, pp. 925–935, 2010.
- [66] M. C. A. Jackson-Menaldi, *La voz normal*. Ed. Médica Panamericana, 1992.
- [67] T. Masuda, Y. Ikeda, H. Manako, y S. Komiyama, “Analysis of vocal abuse: fluctuations in phonation time and intensity in 4 groups of speakers,” *Acta Oto-Laryngologica*, vol. 113, num. 4, pp. 547–552, 1993.
- [68] M. Virebrand, “Real-time monitoring of voice characteristics using accelerometer and microphone measurements,” 2011.
- [69] V. B. López, “Anatomía y fisiología de los mecanismos del habla,” *Odovtos-International Journal of Dental Sciences*, num. 8, pp. 22–31, 2006.
- [70] M. U. Lacabe, S. F. González, M. M. Girbau, L. Sarrasqueta, y R. Urrutia, “Voz cantada,” *Rev Med Univ Navarra*, vol. 50, num. 3, pp. 49–55, 2006.
- [71] P. G. Farías, *Ejercicios que restauran la función vocal: observaciones clínicas*. Librería Akadia Editorial, 2007.
- [72] L. Neira, *Teoría y técnica de la voz: el método Neira de educación vocal*. Akadia, 2009.
- [73] M. Behlau, *A voz do especialista*. Revinter, 2001.
- [74] E. Yumoto, W. J. Gould, y T. Baer, “Harmonics-to-noise ratio as an index of the degree of hoarseness,” *The journal of the Acoustical Society of America*, vol. 71, num. 6, pp. 1544–1550, 1982.
- [75] L. R. Rabiner y R. W. Schafer, *Introduction to digital speech processing*. Now Publishers Inc, 2007, vol. 1.

- [76] J. S. William Moebs, Samuel J. Ling. (2021, sep) Física universitaria volumen 1. [En línea]. Disponible: <https://openstax.org/books/f%C3%ADsica-universitaria-volumen-1/pages/1-introduccion>
- [77] C. P. Vega, “Sonido y audición,” *UNIVERSIDAD DE CANTABRIA. Dpto. de Ingeniería de Comunicaciones*, pp. 1–2, 2016.
- [78] A. M. J. Jaramillo, *Acústica: la ciencia del sonido*. ITM, 2007.
- [79] A. Gutiérrez, “Las vocales glotalizadas del nivaçle,” in *Proceedings from The 6th Conference on the Indigenous Languages of Latin America (CILLA VI)*, 2013.
- [80] L. D. Estrada Solarte, “El ruido: definición, tipos y efectos por la exposición en ambiente laboral.(alteración auditiva): una revisión de literatura años 2000–2015,” 2017.
- [81] M. F. Yépez Arcentales, “Evaluación de los factores de riesgo físico ocasionado por ruidos en operadores de equipos camineros.” Master’s thesis, Universidad de Guayaquil. Facultad de Ingeniería Industrial. Maestría en . . . , 2016.
- [82] M. Babiuch, P. Foltynnek, y P. Smutny, “Using the ESP32 microcontroller for data processing,” *Proceedings of the 2019 20th International Carpathian Control Conference, ICC 2019*, may 2019.
- [83] “ESP-WROOM-32D/ESP32-WROOM-32U Datasheet.” [En línea]. Disponible: https://synergygallery.renesas.com/media/products/135/288/en-US/esp-wroom-32d_esp32-wroom-32u_datasheet_en.pdf
- [84] “MAX4466 Datasheet.” [En línea]. Disponible: <https://datasheets.maximintegrated.com/en/ds/MAX4465-MAX4469.pdf>
- [85] Luis, “Leer y escribir en una tarjeta sd o micro sd con arduino,” Jan 2018. [En línea]. Disponible: <https://www.luisllamas.es/tarjeta-micro-sd-arduino/#:~:text=Un%20lector%20SD%20es%20un,gran%20capacidad%20y%20peque%C3%B1o%20tama%C3%B1o.>
- [86] “GitHub - MhageGH/esp32_SoundRecorder: ESP32 Sound recorder with simple code in arduino-esp32. (I2S interface).” [En línea]. Disponible: https://github.com/MhageGH/esp32_SoundRecorder
- [87] “GitHub - G6EJD/ESP32-ESP8266-File-Download-Upload-Delete-Stream-and-Directory: ESP (32/8266) File Services in one application, enables management of files on an ESP using SD-Filing System.” [En línea]. Disponible: <https://github.com/G6EJD/ESP32-ESP8266-File-Download-Upload-Delete-Stream-and-Directory>

- [88] F.úñez Batalla@, “PROTOCOLO BASICO PARA LA VALORACION FUNCIONAL DE LA PATOLOGIA VOCAL,” 2015.
- [89] I. Cobeta, F. Núñez, y S. Fernández, *Patología de la Voz*, primera ed. Valencia: Marge Médica Books, 2013. [En línea]. Disponible: [https://seorl.net/PDF/ponenciasoficiales/2013Patolog{x\bgroup\let\unhbox\voidb@x\setbox\@tempboxa\hbox{\OT1\i\global\mathchardef\accent@spacefactor\spacefactor}\let\begin\group\let\typeout\protect\begin\group\def\MessageBreak{\Omega\(Font\)}\let\protect\immediate\write\m@ne{LaTeXFontInfo:\def{ }oninputline587.}\end\group\end\group\relax\let\ignorespaces\relax\accent19\OT1\i\egroup\spacefactor\accent@spacefactor\futurelet\@let@token\protect\penalty\@M\hskip\z@skip}adelavoz.pdf](https://seorl.net/PDF/ponenciasoficiales/2013Patolog{x\bgroup\let\unhbox\voidb@x\setbox\@tempboxa\hbox{\OT1\i\global\mathchardef\accent@spacefactor\spacefactor}\let\begin\group\let\typeout\protect\begin\group\def\MessageBreak{\Omega(Font)}\let\protect\immediate\write\m@ne{LaTeXFontInfo:\def{ }oninputline587.}\end\group\end\group\relax\let\ignorespaces\relax\accent19\OT1\i\egroup\spacefactor\accent@spacefactor\futurelet\@let@token\protect\penalty\@M\hskip\z@skip}adelavoz.pdf)
- [90] P. H. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, P. Van De Heyning, M. Remacle, y V. Woisard, “A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques,” *European Archives of Oto-rhino-laryngology*, vol. 258, num. 2, pp. 77–82, 2001.
- [91] R. Garcia-Tapia y I. Cobeta, “Diagnóstico y tratamiento de los trastornos de la voz. editorial garsi sa, editor. 1^a,” 1996.
- [92] N. Elisei, “Acoustic analysis of normal and pathological voices using two different systems: Anagraf and praat,” *Interdisciplinaria*, vol. 29, pp. 339–357, 2012.
- [93] J. M. Vila-Rovira, J. Valero, L. González Sanvisens *y otros*, “Indicadores fonorrespiratorios de normalidad y patología en la clínica vocal,” *Revista de Investigación en Logopedia*, 2011, Vol. 1, Núm. 1, 2017.
- [94] R. Dinámico, R. Vocal, C. Vocal, P. M-Sin, P. H-Sin, P. H-Con, y P. M-Con, “Fonotograma de la voz hablada: valores de referencia normal para hablantes masculinos y femeninos.”
- [95] F. N. Batalla, R. G. Márquez, M. B. P. González, I. G. Laborda, M. F. Fernández, y M. M. Galán, “Acoustic voice analysis using the praat programme: Comparative study with the dr. speech programme,” *Acta Otorrinolaringologica (English Edition)*, vol. 65, num. 3, pp. 170–176, 2014.
- [96] F. N. Batalla y C. S. Nieto, *Espectrografía clínica de la voz*. Universidad de Oviedo, 1999.