



CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS
AVANZADOS DEL INSTITUTO POLITÉCNICO
NACIONAL

UNIDAD ZACATENCO

DEPARTAMENTO DE INGENIERÍA ELÉCTRICA
SECCIÓN DE BIOELECTRÓNICA

**Sistema de evaluación foniatría para la detección temprana de anomalías en
la producción de la voz**

Tesis que presenta

Ing. Ana Laura Cazarín Vázquez

para obtener el Grado de

Maestra en Ciencias

en la Especialidad de

Ingeniería Eléctrica

Directores de Tesis

Dra. Laura Ivoone Garay Jiménez
Dr. Pablo Rogelio Hernández Rodríguez

Ciudad de México.

Diciembre 2019

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo económico brindado durante mi estancia de maestría. Al Centro de Investigación y de Estudios Avanzados de Instituto Politécnico Nacional (CINVESTAV-IPN) por permitirme realizar mi posgrado en esta institución y complementar mi formación académica.

A mis asesores, el Dr. Pablo Rogelio Hernández Rodríguez y la Dra. Laura Ivoone Garay Jiménez por su inmensa paciencia, apoyo y consejos.

A mi familia por el constante apoyo en este proceso y aunque lejos siempre estuvieron ahí cuando los necesite.

Por la ayuda en el desarrollo del proyecto al auxiliar técnico de investigación del laboratorio 4 el Ing. Eladio Cardiel. A todas las personas del laboratorio 4.

A mis amigos Jorge , por su amistad única e incondicional, y a Antonio por su constante apoyo y amistad.

A mis sinodales la Dra. Blanca Tovar Corona y al Dr. Arturo Vera Hernández por su tiempo y colaboración.

Resumen

El sistema que se presenta en esta tesis tiene como principal objetivo ser un dispositivo para evaluar de manera objetiva la calidad de la voz mediante el cálculo de parámetros específicos a partir del análisis de una señal acústica. Se realizaron análisis acústicos a un banco de señales proporcionado por el Instituto Nacional de Rehabilitación (INR), el cual estaba dividido en dos grupos, uno con personas sanas y el otro con pacientes diagnosticados con alguna patología los cuales fueron utilizados como referencia para realizar una clasificación de señales nuevas. Con los resultados obtenidos se realizó un análisis estadístico para seleccionar los parámetros de mayor utilidad e incluirlos en una red neuronal para su clasificación.

Se desarrolló un dispositivo electrónico portátil con el objetivo de que sea transportable y su manejo sea sencillo, el cual integra tres posibles fuentes de caracterización de la voz. Este dispositivo realiza la tarea de registrar señales de audio, vibraciones en la superficie del cuello y flujo de aire producido durante una fonación sostenida.

Se implementó una interfaz gráfica de usuario (GUI) para darle al sistema la característica de poder ser interactivo con el usuario. El sistema cuenta con cuestionarios de información clínica. Finalmente, la clasificación se realizó mediante una red neuronal para reconocimiento de patrones desarrollada en Matlab.

El sistema final fue probado con 8 sujetos realizando 5 repeticiones, se midió la desviación estándar para determinar la repetibilidad de las mediciones del sistema y evaluar su desempeño en general. La comparación de los resultados de parámetros obtenidos previamente con un nuevo conjunto de datos obtenidos con el sistema desarrollado muestra que las nuevas condiciones de grabación mejoraron la repetibilidad.

Abstract

The system presented in this thesis consists in the development of a device for the objective evaluation of the quality of the voice using specific parameters for the analysis of acoustic signals generated during phonations. This analysis was first applied to a bank of data proportionated by the National Institute of Rehabilitation (INR), which was divided into two groups and used as a reference in order to perform the classification of new signals. Then statistical analysis was performed in order to identify the most useful parameters to be included in the design of a neural network that allows classification in health condition.

A portable electronic device was developed in order to facilitate transportation and handle, which integrates three possible sources for voice characterization. This device performs the task of record the audio signals, the vibrations in the surface of the neck, and the airflow generated during sustained phonations.

A Graphic User Interface (GUI) was implemented to interact with the user. The interface includes questionnaires and clinical data collected from the subjects. The classification was carried out by a neural network developed in MATLAB for pattern recognition.

The final system was tested with eight subjects performing five repetitions. The standard deviation was used to determine the repeatability of the system and evaluate its performance. The comparison of the results of the previously obtained parameter with a new dataset obtained with the developed system shows that the new recording conditions improved the repeatability.

Contenido

Capítulo 1	1
1.1 Introducción.....	1
1.2 Objetivos de la tesis.....	2
1.2.1 Objetivo general.....	2
1.2.2 Objetivos específicos	2
1.3 Estructura de la tesis.....	3
Capítulo 2	4
2.1 Antecedentes	4
2.1.1 La voz.....	4
2.1.2 Uso de redes neuronales.....	4
2.1.3 Análisis acústico	5
2.1.4 Parámetros acústicos.....	6
2.1.5 Métodos para análisis estadístico.....	8
2.1.6 Interfaz gráfica de usuario	10
2.1.7 Sensores	10
Capítulo 3	14
3.1 Estado del arte	14
3.1.1 Evaluación de la voz mediante medidas aerodinámicas y mecánicas	14
3.1.2 Análisis acústico	14
Capítulo 4	16
4.1 Solución propuesta	16
4.1.1 Señal acústica de la fonación	16
4.1.2 Vibraciones de la fonación.....	16
4.1.3 Flujo de aire en la fonación.....	16

4.1.4	Acústica de la fonación.....	17
4.1.5	Instrumentación.....	17
4.2	Desarrollo	17
4.2.1	Descripción del desarrollo	17
4.2.2	Obtención de parámetros acústicos a partir de las señales de audio proporcionadas por el INR	18
4.2.3	Análisis estadístico.....	23
4.2.4	Diseño del clasificador.....	25
4.2.5	Diseño a bloques del sistema propuesto	26
4.2.6	Diseño del sistema electrónico.....	27
4.2.7	Sistemas complementarios en el protocolo de registro.....	36
4.2.8	Montaje de sensores.....	37
4.2.9	Adquisición y almacenamiento de datos.....	38
4.2.10	Interfaz gráfica	41
Capítulo 5	46
5.1	Resultados y discusiones.....	46
5.1.1	Análisis acústico	46
5.1.2	Análisis estadístico.....	47
5.1.3	Red neuronal	56
5.1.4	Prototipo del dispositivo de obtención de señales	58
5.1.5	Obtención de señales del circuito electrónico.....	60
5.1.6	Interfaz gráfica	66
5.1.7	Pruebas del sistema.....	72
5.1.8	Evaluación de parámetros	74
Capítulo 6	80

6.1	Conclusiones	80
6.2	Trabajo a futuro	81
	Referencias	83
	Anexos.....	87
A.1	<i>Script</i> para obtención de parámetros en Praat	87
A.2	<i>Script</i> de la función de la red neuronal generado en Matlab	89
A.3	Script para adquisición de datos en tarjeta de desarrollo ESP32	92
A.4	Función para importar el archivo	96
A.5	Diseño del circuito impreso.....	98
A.6	Carta de consentimiento informado	99

Índice de figuras

Fig. 1. Estructura básica de una red neuronal para reconocimiento de patrones. ...5	5
Fig. 2. Esquema de principio de funcionamiento del sensor..... 10	10
Fig. 3. Respuesta en frecuencia de micrófono electret..... 11	11
Fig. 4. Diagrama de flujo de los procesos seguidos en el desarrollo. 18	18
Fig. 5. Ventana de inicio para análisis acústico en Praat. 20	20
Fig. 6. Diseño de la red neuronal en Matlab®. 25	25
Fig. 7. Diagrama de flujo del bloque 1 de la Fig. 4..... 26	26
Fig. 8. Montaje de sensores y distribución grafica del sistema..... 27	27
Fig. 9. Diagrama a bloques del sistema electrónico propuesto. 27	27
Fig. 10. Sensor de flujo Thermal Mass Flow Sensor FS5..... 28	28
Fig. 11. Circuito de acoplamiento propuesto por el fabricante. 29	29
Fig. 12. Gráfica de respuesta del sensor proporcionada por el fabricante..... 30	30
Fig. 13. Circuito restador para eliminación de <i>offset</i> 30	30
Fig. 14. Esquema de circuito amplificador no inversor. 31	31
Fig. 15. Circuito para acoplamiento del sensor de flujo..... 32	32
Fig. 16. Batería tipo Li-Po de tres celdas de 11.1Volts..... 32	32
Fig. 17. Micrófono Electret CMC5044PF..... 33	33
Fig. 18. Circuito de medición para micrófono $R1=2.2k\Omega$; $C1=1\mu F$ 33	33
Fig. 19. Preamplificador para el micrófono Electret [36]. 34	34
Fig. 20. Módulo Max4466 para preamplificación del micrófono. 34	34
Fig. 21. Acelerómetro MMA7361L. 35	35
Fig. 22. Circuito de conexión para el acelerómetro. 36	36
Fig. 23. Sensor de humedad y temperatura DHT11..... 36	36
Fig. 24. Mascarilla con extensión y sensores de audio y flujo montados. 38	38
Fig. 25. Tarjeta de desarrollo ESP32..... 39	39
Fig. 26. Módulo microSD para almacenamiento de datos. 40	40
Fig. 27. Conexión de módulo microSD a ESP32..... 41	41
Fig. 28. Diagrama a bloques para el desarrollo de la interfaz gráfica. 42	42
Fig. 29. Cuestionario VHI-Índice de capacidad vocal [38]..... 43	43

Fig. 30. Diagrama de flujo de bloque 5 de la Fig. 28.....	44
Fig. 31. Diagrama de flujo de bloque 6 de la Fig. 28.....	45
Fig. 32. Distribución de valores obtenidos del análisis acústico del grupo 1.	47
Fig. 33. Distribución de valores obtenidos del análisis acústico del grupo 2.	48
Fig. 34. Distribución de valores obtenidos del análisis acústico del grupo 3.	49
Fig. 35. Distribución de valores obtenidos de análisis acústico del grupo 4.....	49
Fig. 36. Distribución de valores obtenidos del análisis acústico del grupo 5.	50
Fig. 37. Distribución de valores obtenidos del análisis acústico del grupo 6.	51
Fig. 38. Distribución de valores obtenidos del análisis acústico del grupo 7.	51
Fig. 39. Distribución de valores obtenidos del análisis acústico del grupo 8.	52
Fig. 40. Desempeño de la red neuronal sobreentrenada.....	56
Fig. 41. Desempeño de la red neuronal.....	57
Fig. 42. Diagrama de confusión de la red neuronal.....	58
Fig. 43. Montaje de los elementos que integran el sistema de obtención de señales.	59
Fig. 44. Dispositivo resultante con referencia de medida.	59
Fig. 45. Prototipo del dispositivo electrónico para registro.	60
Fig. 46. Toma de registro del sujeto de prueba.	61
Fig. 47. Respuesta del sensor de flujo a tres impulsos de aire.	62
Fig. 48. Respuesta del sensor de flujo durante una vocalización sostenida de una "a".....	62
Fig. 49. Señal de micrófono durante una vocalización de la letra "a" del sujeto de prueba.....	63
Fig. 50. Repetición de vocalización del sujeto de prueba.	63
Fig. 51. Señal del acelerómetro en vocalización de una "a" de sujeto de pruebas.	64
Fig. 52. Señal del acelerómetro durante una vocalización de una "a" del sujeto de pruebas.	64
Fig. 53. Respuesta en tiempo de la humedad con sensor DHT11.	65
Fig. 54. Respuesta en el tiempo de la temperatura con el sensor DHT11 y con el termómetro comercial.	66

Fig. 55. Menú de inicio de la interfaz gráfica.	67
Fig. 56. Ventana para ingreso de datos de nuevo expediente.....	68
Fig. 57. Cuestionario VHI.	69
Fig. 58. Ventana para listado de pacientes con expediente.	70
Fig. 59. Ventana de registro.	71
Fig. 60. Ventana para el despliegue de resultados.	72
Fig. 61. Comparativa entre grupos de estudio con base a la frecuencia fundamental (F0).....	75
Fig. 62. Comparativa entre grupos de estudio con base a los parámetros de CPPS y HNR.	75
Fig. 63. Comparativa entre grupos de estudio con base a la desviación estándar de F0.	76
Fig. 64. Comparativa entre grupos de estudio con base a los parámetros de <i>Shimmer</i>	76
Fig. 65. Comparativa entre grupos de estudio con base a los parámetros de <i>Jitter</i>	77
Fig. 66. Comparativa entre grupos de estudio con base al <i>Shimmer local dB</i>	78
Fig. 67. Diseño PCB del sistema final.	98

Índice de tablas

Tabla 1. Características técnicas del sensor de flujo FS5.....	12
Tabla 2. Características del acelerómetro MMA7361.....	13
Tabla 3. Lista de los parámetros obtenidos en Praat.	24
Tabla 4. Características técnicas del módulo Max4466.....	35
Tabla 5. Características técnicas del acelerómetro MMA7361L.	35
Tabla 6. Características técnicas del sensor DHT11 [37].	37
Tabla 7. Características técnicas de la tarjeta de desarrollo EP32.	39
Tabla 8. Valores promedio obtenidos del análisis acústico en Praat.....	46
Tabla 9. Resultados del análisis de normalidad Shapiro-Wilk.....	53
Tabla 10. Resultados del análisis <i>t test</i> de manera individual de los parámetros.	54
Tabla 11. Resultados del análisis <i>Two-way ANOVA</i>	55
Tabla 12. Resultados de la encuesta VHI.....	72
Tabla 13. Desviaciones estándar de los valores obtenidos de cada sujeto.	73
Tabla 14. Comparativa de promedios y desviación estándar entre los grupos de estudio.	74
Tabla 15. Resultados de la clasificación de sujetos de prueba.....	79

Capítulo 1

1.1 Introducción

La voz es el principal medio de comunicación y su uso más frecuente como herramienta de trabajo, ha propiciado la incidencia de entre un 3 a 9% de diversas patologías que aquejan a la población en general [1]. De entre la población en edad laboral más afectada se encuentran los profesionales docentes, vendedores, oradores, actores, cantantes y locutores, entre otros, que usan la voz como herramienta de trabajo.

Aunado a lo anterior, el uso de la voz sin el adecuado entrenamiento y cuidado puede provocar problemas y lesiones persistentes [2]. La Organización Internacional del Trabajo (OIT) describe que los docentes conforman la principal categoría profesional que está expuesta a padecer patologías de la voz por el uso recurrente de voz proyectada, lo que afecta de manera directa su desempeño en las aulas de clases, así como en su vida cotidiana[3]. Cabe destacar que toda la población es susceptible a sufrir algún problema de la voz, por lo cual resulta necesario el acceso y atención de servicio médico especializado en foniatría. Sin embargo, estas dos acciones, no son suficientes para poder proporcionar los diagnósticos y tratamientos efectivos, en algunos casos por la falta de equipos portátiles para trabajo de campo que limita la obtención de resultados significativos, tanto para la detección de problemas como para la creación de una base de datos para su estandarización.

En México existen pocas bases de datos que reporten de manera porcentual el daño que pueden presentar personas con anomalías foniátricas complementado con datos estadísticos como edad, ocupación y sexo. Por esta condición, se presenta un problema al momento de clasificar la condición de la voz y así determinar su estado de

salud. Además, una clasificación como ésta se puede utilizar como referencia para la obtención de identificaciones primarias con propósitos preventivos, o casos que requieran atención inmediata.

Los equipos comerciales actuales en foniatría no poseen una base de medidas complementarias objetivas de calidad de la voz como referencia en las pruebas de fonación que sirvan para interpretar los resultados del estado de las estructuras [4]. Debido a esto, el diagnóstico que se emite en diferentes instituciones de medicina foniátrica se apoya fundamentalmente en un conjunto de estudios clínicos basados en protocolos que incluyen técnicas invasivas, ocasionando incomodidad, miedo, molestias y en algunas ocasiones lesiones a los pacientes.

1.2 Objetivos de la tesis

1.2.1 Objetivo general

Desarrollar un sistema portable y no invasivo para la medición de flujo de aire, audio y vibración en la superficie del cuello durante una fonación, para la detección temprana de anomalías en la producción de la voz.

1.2.2 Objetivos específicos

- Realizar el análisis estadístico de un banco de señales foniátricas obtenidas en clínica para la selección de parámetros útiles en el estudio diferencial de casos patológicos y sanos, utilizando los programas Praat y Prisma.
- Seleccionar los sensores para la medición de flujo de aire, audio y vibraciones producto de una fonación.
- Implementar una estructura portátil no invasiva para el montaje de los sensores que asegure el registro de las señales de estudio de manera confiable y cómoda.
- Realizar registro y procesamiento de las señales foniátricas.
- Diseñar una red neuronal para clasificación por reconocimiento de patrones que permita obtener información inicial y diferenciar entre casos sanos y patológicos.

- Diseñar una interfaz gráfica en Matlab para registro de participantes y toma de muestras.

1.3 Estructura de la tesis

En esta tesis se desarrolló un sistema de evaluación fonológica para uso ambulatorio y así ofrecer una detección temprana de patologías, ocasionadas por personas que usan la voz de manera intensiva para el desempeño de sus actividades laborales.

En el capítulo uno se aborda la problemática en cuestión y se presentan datos estadísticos correspondientes al tema, así como una breve descripción de la propuesta de solución la cual se retomará más a detalle en el capítulo 4. También se da a conocer el objetivo general del proyecto, así como los objetivos específicos a realizar. Posteriormente en el capítulo 2 se plantean conceptos generales referentes al proyecto que fueron de apoyo en el desarrollo de esta tesis.

En el capítulo 3 se presentan los trabajos más relevantes en el campo. Se mencionan sus aportaciones y se realiza un análisis en conjunto para destacar los aportes de este proyecto. El diseño a modo de diagrama a bloques, así como la solución propuesta a la problemática, presentada en la introducción en el primer capítulo, se presentan en el capítulo 4, al igual que el desarrollo de la metodología empleada con los procesos y las herramientas requeridas para la realización del objetivo general.

El capítulo 5 enlista los resultados obtenidos y la discusión de estos. En el capítulo 6 se presentan las conclusiones obtenidas al finalizar el trabajo realizado en la tesis y las perspectivas para el trabajo futuro de este proyecto.

Capítulo 2

2.1 Antecedentes

2.1.1 La voz

La función vocal no es imprescindible para la vida humana pero necesaria para la comunicación con sus semejantes. La voz es el sonido producido voluntariamente por el aparato fonatorio humano, no solo transmite mensajes orales, sino que a través de sus características muestra el estado emocional de una persona.

La voz humana es el sonido producido por la vibración de las cuerdas vocales y la resonancia de los espacios por los que atraviesa el aire expelido por los pulmones. Cuando nos referimos al habla, no es sólo la emisión del sonido, también incluye la modificación del sonido por los resonadores y la emisión de la palabra por los articuladores. Asimismo, la voz puede reflejar aspectos importantes con respecto a su salud, además de dar indicios de la edad aproximada, ya que la voz a lo largo de la vida se ve modificada [5].

En el proceso de fonación hay diferentes teorías para explicar la vibración necesaria de las cuerdas vocales, una de ellas es el efecto de Bernoulli. En el cual se propone que las cuerdas vocales se aproximan unos 3 mm, dejando un pequeño vacío que produce la columna de aire y, al pasar causa una explosión que es un tono fundamental laríngeo, el cual luego se modifica en las paredes laríngeas al ponerlas en movimiento y las convierte en un sonido [6].

2.1.2 Uso de redes neuronales

Una red neuronal es un conjunto interconectado de elementos, unidades o nodos de procesamiento simples, cuya funcionalidad se basa en la neurona. La capacidad de procesamiento de la red se almacena en los pesos de conexión entre unidades, obtenidos

mediante un proceso de adaptación o aprendizaje de un conjunto de patrones de entrenamiento [7]. Las redes neuronales artificiales se han usado ampliamente clasificación, es decir, en los que se trata de determinar el tipo de categoría a la que pertenece un elemento.

Actualmente existen diversos tipos de redes neuronales, se pueden clasificar según su tipología o su estructura ya sean monocapa o multicapa, o bien su algoritmo de aprendizaje de acuerdo con si éste es supervisado o no, su diseño y configuración es adaptada a la aplicación específica. Un sistema de reconocimiento de patrones que se construye a partir de analizar y manipular instancias o ejemplos de un problema se conoce como adaptivo. La construcción de este tipo de sistemas tiene dos etapas principales: diseño y evaluación. En la primera se generan reglas de decisión para clasificar instancias del problema y en la segunda se verifica de forma intensiva que dichas reglas sean apropiadas para tomar buenas decisiones [8].

El entrenamiento de la red busca identificar el valor del conjunto de pesos o sinapsis que, logra reproducir las salidas incluidas en el conjunto de datos de entrenamiento, mediante la minimización de la función de error de predicción.

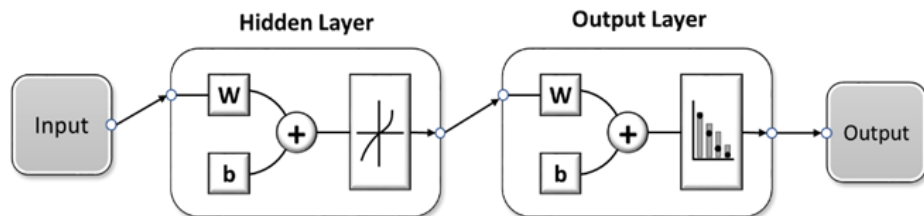


Fig. 1. Estructura básica de una red neuronal para reconocimiento de patrones.

2.1.3 Análisis acústico

El análisis acústico de la voz es una herramienta objetiva y no invasiva de exploración vocal, que emplea el registro y análisis de la señal acústica obtenida a partir de una tarea vocal.

Aplicaciones del análisis acústico

El análisis acústico tiene diversas aplicaciones; una de ellas es en el área de lingüística en donde se estudian fonemas individuales de coarticulación, además de la articulación vocálica y la prosodia.

Otra aplicación muy importante es el estudio de patologías vocales. En esta rama de aplicación se busca la detección de patrones objetivos de las emisiones vocales además de la descripción detallada de dichas patologías. Para ello siempre se considera la comparación de los patrones vibratorios de emisiones normales para mejorar la discriminación de eventos anormales. El seguimiento de los tratamientos fonoaudiológicos es otro ejemplo de aplicaciones, este permite evaluar el desempeño del tratamiento sobre un paciente. Para el área de seguridad o biometría.

El reconocimiento e identificación de voces de personas en específico.

Programas para análisis acústico

Praat es un programa computacional de libre distribución, de código abierto, multiplataforma y, además, gratuito, diseñado especialmente para hacer investigaciones en fonética y análisis acústico ya que ofrece una amplia gama de procedimientos estándar y no estándar, incluyendo análisis espectrográfico, síntesis articulatoria y redes neuronales.

2.1.4 Parámetros acústicos

La determinación de parámetros acústicos de manera objetiva, mediante programas de análisis, sirve de apoyo al diagnóstico. Los más representativos y seleccionados para este trabajo de tesis se enlistan a continuación [7,8,9].

Parámetros de la frecuencia

F0 (Frecuencia fundamental): Este parámetro se refiere a la cantidad de veces que las cuerdas vocales se abren y cierran por segundo durante una fonación. El programa Praat emplea el método de autocorrelación para el cálculo de este parámetro.

F0 SD (Desviación estándar de la frecuencia fundamental): Hace referencia a la variabilidad de la frecuencia fundamental.

Parámetros de la perturbación de la frecuencia

Jitter local absolute: Es la diferencia absoluta media entre periodos consecutivos.

Jitter local: Representa la variabilidad absoluta promedio de la frecuencia periodo a periodo, dividida por el período promedio de la señal.

Jitter rap (Perturbación media relativa): Corresponde a la diferencia relativa de la frecuencia entre un periodo y la media de los dos periodos contiguos, dividido por la frecuencia promedio total de los periodos de la señal.

Jitter ppq5 (Cociente de perturbación del período de cinco puntos): Corresponde a la variabilidad absoluta promedio entre un período y el promedio de este y sus cuatro vecinos más cercanos, dividido por el período promedio.

Jitter ddp: Esta es la diferencia absoluta promedio entre las diferencias consecutivas entre períodos consecutivos, dividida por el período promedio.

Parámetros de la perturbación de la amplitud

Shimmer local: Representa la variabilidad absoluta entre la amplitud de periodos consecutivos, divididos por la media total de la amplitud.

Shimmer dB: Se refiere a la diferencia entre picos expresada en decibelios que es el logaritmo en base 10 de la media absoluta de la diferencia entre la amplitud de periodos consecutivos, multiplicado por 20.

Shimmer apq3 (Cociente de perturbación de amplitud de tres puntos): Corresponde a la diferencia absoluta entre la amplitud de un periodo y el promedio de las amplitudes de los dos periodos consecutivos, divididos por el total de la amplitud promedio.

Shimmer apq5 (Cociente de perturbación de amplitud de cinco puntos): Corresponde a la diferencia absoluta entre la amplitud de un periodo y el promedio de las amplitudes de los cuatro periodos consecutivos, divididos por el total de la amplitud promedio.

Shimmer ddp: Esta es la diferencia absoluta promedio entre diferencias consecutivas entre las amplitudes de períodos consecutivos.

Parámetro de ruido

HNR (Relación armónico-ruido): La relación armónico-ruido es la medida que cuantifica la cantidad de ruido aditivo en la señal de la voz. Es el logaritmo en base 10 de la relación entre la energía periódica y la energía correspondiente al ruido, multiplicado por 10.

HNR SD (Desviación estándar de la relación armónico-ruido): Hace referencia a la variabilidad de la relación armónico-ruido.

Parámetro del Cepstrum

El cepstrum es la transformada inversa de Fourier del logaritmo de la potencia espectral de la señal. En otras palabras, consiste en aplicar al sonido 2 transformadas de Fourier. El análisis cepstral permite determinar la F0 a partir de la periodicidad del componente armónico de un sonido, representado por el pico cepstral de la región alta del cepstrum.

CPPs (Prominencia del pico cepstral suavizado): Este parámetro representa la diferencia entre el pico cepstral más prominente, que corresponde al primer armónico (*rahmonic*), y el punto con la misma frecuencia (*quefreny*) sobre la línea de regresión a través del cepstrum suavizado.

2.1.5 Métodos para análisis estadístico.

Prueba Shapiro-Wilk

La prueba Shapiro-Wilk, desarrollada por Shapiro y Wilk, es la prueba más efectiva en la mayoría de las situaciones. En los últimos años, la prueba SW se ha convertido en la prueba preferida de normalidad debido a sus buenas propiedades de potencia en comparación con una amplia gama de pruebas alternativas. La prueba SW depende de la correlación entre los datos y sus puntajes normales correspondientes [12].

La prueba de Shapiro-Wilk es una prueba de distribución normal que exhibe alta potencia de diferenciación, lo que lleva a buenos resultados incluso con un pequeño

número de observaciones. A diferencia de otras pruebas de comparación, la prueba de Shapiro-Wilk solo es aplicable para verificar la normalidad [13].

Esta prueba de la hipótesis de normalidad se requiere para determinar qué tipo de comparación es la más adecuada, ya que los métodos estadísticos como las pruebas t y el análisis de varianza suponen que las variables se distribuyen normalmente. Si las distribuciones muestrales no se pueden modelar como una distribución gaussiana es necesario usar métodos no paramétricos.

Prueba Mann-Whitney

La U de Mann-Whitney es una prueba no paramétrica utilizada para evaluar las diferencias significativas en una escala o variable dependiente ordinal por una sola variable independiente dicotómica. Es el equivalente no paramétrico de la prueba t de muestras independientes. Esto significa que la prueba no asume ninguna propiedad con respecto a la distribución de la variable dependiente en el análisis. Esto hace que la prueba U de Mann-Whitney sea el análisis apropiado para usar al analizar variables dependientes en una escala ordinal. La prueba de Mann-Whitney es también la base matemática para la prueba Kruskal Wallis H [14].

Análisis de la varianza de dos factores.

En un análisis de la varianza de dos vías compara las diferencias de medias entre los grupos de muestras para dos variables independientes. El propósito principal de un análisis de la varianza de dos vías es comprender si existe una interacción entre las dos variables independientes en la variable dependiente considerando niveles o rangos en las muestras [15].

Este tipo de análisis nos expresa si el efecto de una de las variables independientes en la variable dependiente es el mismo para todos los valores de su otra variable independiente y viceversa. que va acompañado con estudios post-test de las diferencias entre pares de variables para terminar su diferencia estadísticamente significativa.

2.1.6 Interfaz gráfica de usuario

Una interfaz gráfica (GUI) es un elemento gráfico que nos ayuda a comunicar de manera visual las opciones de un sistema o estructura [16]. Es una herramienta que consiste en un modelo computacional y una herramienta de presentación gráfica que le permite al usuario ajustar fácil y naturalmente los parámetros del modelo, ejecutar el cálculo y ver los nuevos resultados [17].

Los componentes de una interfaz gráfica son elementos gráficos y a través de ellos el usuario puede interactuar con la aplicación. En esta interacción el usuario introduce datos, mediante imágenes de ventanas, botones, etc., que el programa necesita para llevar a cabo su funcionalidad y obtener los resultados de procesar dichos datos [18].

2.1.7 Sensores

Un sensor es un elemento que modifica sus características físicas y/o químicas para conocer y registrar la presencia de algún fenómeno.

El principio se basa en la detección en forma cuantitativa de un componente con características imperceptibles por los sentidos, a los cuales se le asignará un valor y una unidad capaz de representar su magnitud.

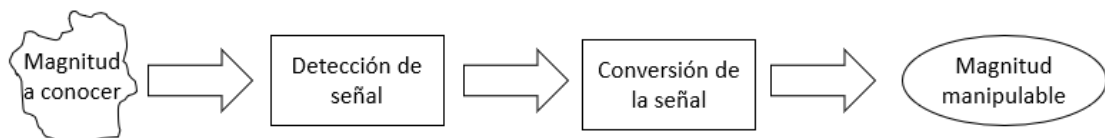


Fig. 2. Esquema de principio de funcionamiento del sensor.

Micrófono electret

El micrófono electret es una variante del micrófono de condensador. Los dispositivos actuales vienen con una carga permanente incorporada en las placas que polarizan al condensador durante la fabricación, por lo que no necesita alimentación externa para operar [19].

El micrófono electret puede ser omnidireccional o unidireccional, son robustos y de tamaño reducido, lo cual facilita su manipulación en aplicaciones donde se requiere un dispositivo fácil de ocultar.

Los micrófonos electret poseen buena respuesta en frecuencia en el rango audible de 50 a 15,000Hz, como se muestra en la Fig. 3 y una sensibilidad entre -50 dB y -70 dB. Se alimenta con una tensión de entre 2 y 12V, la señal que registra es amplificada internamente para entregar una señal mayor y de mejor calidad, es resistente, poco ruidoso y muy sensible [20]. Cubren el rango requerido para voz que va desde 80 hasta 525 Hz con una medición constante en el ancho de banda [21].

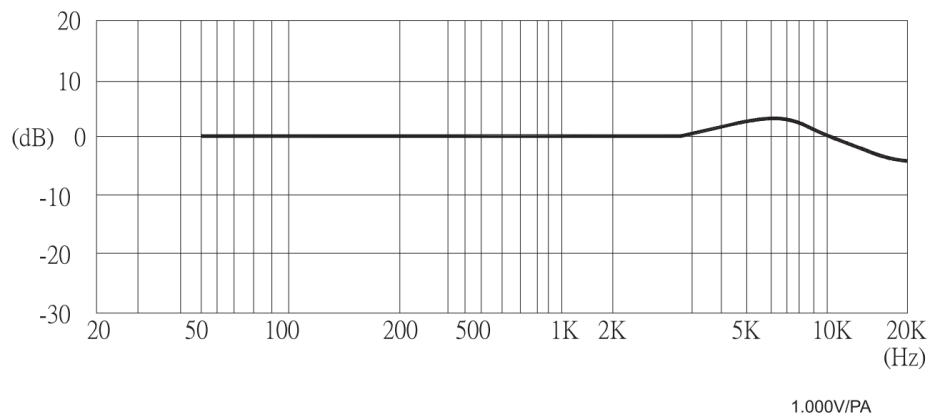


Fig. 3. Respuesta en frecuencia de micrófono electret.

Flujómetro FS5 thermal mass flow sensor

El sensor de flujo FS5 de la empresa suiza *Innovative Sensor Technology*, utiliza la relación entre la tasa de flujo y el efecto de enfriamiento de medición directa de la masa de flujo. Es el sucesor de la línea FS1 y se caracteriza por una estructura de sensado térmicamente más eficiente. Esto se debe a un chip que soporta temperaturas más altas, el cual es más sensible. A diferencia de otros sensores de flujo basados en turbinas, éste tiene un funcionamiento por diferencia de temperatura además de su reducido tamaño.

Este sensor se basa en el principio conductimétrico e incluye 2 resistencias de platino sobre un chip. Una de ellas es utilizada como calentador y la otra, de valor alto, para la medición de la temperatura del fluido. Los sensores conductimétricos térmicos de flujo se basan en coeficientes de transferencia de calor, los cuales están en función

de la velocidad de flujo. El rango de medición de flujo es muy amplio, puede ser ajustado para aplicaciones específicas.

Con ayuda de un circuito electrónico, es posible el incremento de la temperatura del calentador con respecto a la temperatura del medio. La velocidad del flujo cambia la pérdida de energía térmica por el conductor propiciando su enfriamiento, el cual está en función de la masa del fluido.

Por la pequeña masa del sensor se cuenta con una respuesta rápida y poco tiempo de calentamiento, lo que aunado al tamaño de 6 mm de diámetro, son características atractivas para múltiples aplicaciones [22]. Sus características principales se mencionan en la Tabla 1.

Tabla 1. Características técnicas del sensor de flujo FS5.

Principio de medición	Térmico
Rango de medición	0 - 100 m/s
Sensibilidad	0.01 m/s
Precisión	< 3 % del valor medido
Tiempo de respuesta	< 2 s
Rango de temperatura	-20...+150 °C
Sensibilidad de temperatura	< 0.1 % /K
Conexión eléctrica	3 pines de conexión
Voltaje de alimentación	2 - 5 V

Acelerómetro Modelo 7361

Para el registro de las vibraciones de las cuerdas vocales que van desde los 80 a 300 Hz en hombres y 135 a 525 Hz en mujeres [21] se eligió el acelerómetro MMA7361, el cual es un dispositivo inercial micromaquinado de baja energía con perfil capacitivo, con acondicionamiento de señal integrado, filtro pasa bajas de 1 polo, compensador de temperatura, autoprueba, detector de caída lineal con capacidad de selección de 2 sensibilidades. Estas dos últimas cualidades del acelerómetro no necesitan de dispositivos externos para su uso. Algunas de las ventajas de utilizar este dispositivo son:

- Detección de autoprueba para diagnóstico.
- Señal acondicionada con filtro paso-bajas.
- Diseño robusto, resistencia a grandes impactos.
- Producto de adaptación ambiental.
- Bajo costo.

Incluye también un modo de descanso lo que lo hace ideal cuando se maneja con dispositivos electrónicos de batería, lo cual es importante cuando se pretende diseñar un dispositivo portátil [23]. Algunas de sus características se enlistan en la Tabla 2.

Tabla 2. Características del acelerómetro MMA7361.

Dimensiones	3 mm x 5 mm x 1.0 mm
Consumo de corriente	400 μ A
Modo de descanso	3 μ A
Voltaje de operación	2.2 V – 3-6 V
Sensibilidad	800 mV/g @ 1.5g
Tiempo de respuesta	5 ms

Capítulo 3

3.1 Estado del arte

3.1.1 Evaluación de la voz mediante medidas aerodinámicas y mecánicas

El uso de medidas aerodinámicas no invasivas ha sido de gran apoyo para la detección de patologías en voz y mejorar la comprensión clínica en acústica.

Se ha reportado que los parámetros obtenidos de las medidas aerodinámicas son de gran utilidad para la discriminación entre algunos tipos de hiperfuncionalidad vocal y en general han demostrado ser potenciales diferenciadores entre patologías vocales [24].

Por otro lado, se han realizado estudios correspondientes al mapeo de la actividad vibratoria en toda la superficie del cuello a diferentes frecuencias de vocalizaciones, utilizando acelerómetros en la superficie del cuello. A este tipo de estudios se le han dado diversas aplicaciones, la más sobresaliente es la de utilizar las vibraciones laríngeas como un canal de comunicación [25].

3.1.2 Análisis acústico

El análisis acústico es de gran ayuda para el diagnóstico de patologías en voz, ya que aporta información objetiva de la voz de manera indirecta y no invasiva. Se han reportado diversos métodos de análisis acústico para la detección de patologías basadas en coeficientes cepstrales de frecuencia de Mel y variaciones en las frecuencias y amplitudes como características (*Jitter* y *Shimmer*) y Modelos de Mezcla Gaussiana [26].

Desafortunadamente, actualmente no se cuentan con estudios normativos en México para los parámetros obtenidos del análisis acústico. Sin embargo, se ha reportado que dentro del análisis acústico, las medidas temporales, espectrales y cepstrales son las de mayor aportan al análisis para diagnóstico [27]. Además, se han identificado y anexado parámetros a este análisis, siendo el pico cepstral de mayor prominencia en el cepstrum suavizado el que posee una mayor precisión en el diagnóstico de disfonía, además de evaluar el grado de severidad de dicha alteración en habla conectada. Para este estudio se analizaron vocalizaciones de sujetos hablantes en español [9].

Actualmente existen diversos equipos comerciales que se utilizan en el apoyo de la clínica en foniatría o bien con propósitos de investigación. Un ejemplo de esto es el sistema *The Phonatory Aerodynamic System (PAS)* [28], el cual es un sistema con el propósito de apoyar la práctica clínica basada en evidencia, ya que este proporciona medidas clave del flujo de aire, la presión del habla y de la emisión de voz, incluso información gráfica y cuantitativa para monitorizar y reportar el desempeño del paciente. Este mismo sistema ha sido utilizado con el fin de establecer una base de datos en adultos con 41 medidas aerodinámicas, además de observar el efecto de estas mediciones con respecto a la edad y género de los participantes en dicho estudio [29].

Reportes posteriores han demostrado que no siempre es posible una clasificación entre voces patológicas y sanas utilizando bases de datos de sistemas comerciales, al depender en gran medida del personal que realiza el protocolo de registro [30], y esto no permite la comparación de los resultados y por lo tanto no es confiable establecer modelos multimodales a partir de la conjunción de estas. Por lo cual aún es un campo de oportunidad generar sistemas integrales de evaluación foniátrica de manera simultánea.

Capítulo 4

4.1 Solución propuesta

El sistema diseñado permite la medición de parámetros fundamentales en foniatría para realizar un análisis automático de la señal acústica, para detectar anomalías tempranas de la voz en personas que por su trabajo tengan una mayor probabilidad de experimentar alguna patología foniátrica. De manera complementaria se registran señales de flujo y vibraciones en la superficie del cuello, esto con el objetivo de conformar una base de datos para un análisis a futuro.

4.1.1 Señal acústica de la fonación

La señal acústica es un indicador muy utilizado para el análisis de voz ya que permite una evaluación vocal objetiva y es no invasiva para establecer diagnósticos, ser monitor en terapia y generar líneas de investigación en el modelado multidimensional del evento foniátrico. Con este tipo de señales se obtiene información altamente integrada del desempeño conjunto de las tres cavidades foniátricas involucradas en el resultado final.

4.1.2 Vibraciones de la fonación

Otra de las mediciones utilizadas para apoyar un diagnóstico es la de las vibraciones producidas por el trabajo mecánico de las cuerdas vocales. Para este propósito se utilizan acelerómetros colocados en el cuello. De este modo se pueden alizar si existen irregularidades en esta etapa del proceso en la fonación [30].

4.1.3 Flujo de aire en la fonación

El tercer parámetro que se va a incluir en el sistema será el registro de flujo de aire, el cual es un evento controlado que permite conocer la correlación entre la capacidad de emisión de aire de los pulmones y la operación de las diversas estructuras durante las

fonaciones. Alteraciones en los valores del flujo de aire durante patrones de fonación pueden estar relacionadas con anomalías en las cuerdas vocales, incluso en pacientes con tasas de flujo promedio aparentemente normales. Además, la tasa de flujo promedio durante la fonación y especialmente el grado de fluctuación del flujo de aire proporcionan medidas cuantitativas útiles de disfunción laríngea [31].

4.1.4 Acústica de la fonación

Con el procesamiento de las señales acústicas se obtienen los valores de la frecuencia fundamental (F0), *Shimmer* y *Jitter*. La F0 es el parámetro relacionado con las características físicas naturales de las estructuras que participan en el proceso de fonación. El *Jitter* es un índice que sirve para medir la regularidad en la frecuencia fundamental y la variabilidad en la duración de las ondas en ciclos consecutivos. En el caso del *Shimmer* se evalúa la variabilidad de la amplitud del sonido ciclo a ciclo. Ambos parámetros de perturbación se obtienen por análisis informático de fonaciones vocales prolongadas o muestras de voz y cuantifican irregularidades no intencionales en las ondas acústicas generadas por la laringe. Los parámetros *jitter* y *shimmer* se utilizan en entornos clínicos y científicos con fines diagnósticos y descriptivos, así como para documentar y evaluar el resultado del tratamiento de la voz[30,31].

4.1.5 Instrumentación

El sistema está previsto para ser un dispositivo portátil que permita la realización de trabajo de campo. Deberá cumplir con: 1) ser no invasivo; 2) no comprometer la integridad física de los usuarios; 3) Registro y almacenamiento de datos; Procesamiento de los datos; Análisis de información; 4) amigable en su uso; 5) ligero y pequeño para su transportación; 6) alimentación basado en baterías.

4.2 Desarrollo

4.2.1 Descripción del desarrollo

En el diagrama de flujo la Fig. 4 se presentan las actividades a realizar en el desarrollo de este trabajo de tesis, las cuales se detallan más adelante.

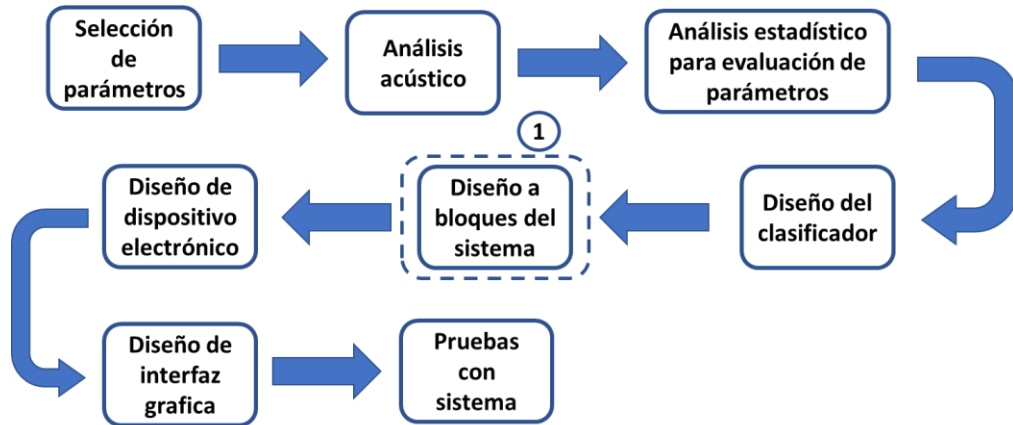


Fig. 4. Diagrama de flujo de los procesos seguidos en el desarrollo.

4.2.2 Obtención de parámetros acústicos a partir de las señales de audio proporcionadas por el INR

Para la obtención de parámetros acústicos mediante el análisis de audio se empleó el software Praat el cual es especializado en voz creado por Paul Boersma y David Weenink en la universidad de Amsterdam [34].

De acuerdo con la bibliografía revisada y algunos de los trabajos actuales más sobresalientes [1, 20, 21], se seleccionaron la frecuencia fundamental, la perturbación en amplitud y en frecuencia, parámetros respecto a la relación ruido-armónico y medidas espectrales.

Parámetros obtenidos

- Frecuencia fundamental

Para la determinación de la frecuencia fundamental se empleó el programa Praat que se basa en el método de autocorrelación el cual se representa en la ecuación 1.

$$r_x(\tau) \approx r_{xw}(\tau) / r_w(\tau) \quad (1)$$

Donde $x(\tau)$ es la señal, $r_{xw}(\tau)$ representa la señal en ventana y $r_x(\tau)$ es la función de autocorrelación de la señal. Se utilizaron los valores determinados en Praat para este análisis con ventanas de 10 milisegundos en *time step* y 75-500 Hertz en el *pitch range*.

- Medidas de *Jitter*

$$Jitter(local, absoluto) = \frac{1}{N-1} \sum_{i=1}^{n-1} |T_i - T_{i+1}| \quad (2)$$

$$Jitter(local) = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i+1}|}{\frac{1}{N} \sum_{i=1}^N T_i} \quad (3)$$

$$Jitter(rap) = \frac{\sum_{i=2}^{N-1} |T_i - \frac{T_{i-1} + T_i + T_{i+1}}{3}|}{\frac{N-2}{\sum_{i=1}^N \frac{T_i}{N}}} \quad (4)$$

donde T_i es la longitud del período de F0 extraída y N es el número de períodos extraídos de F0.

- Medidas de *Shimmer*

$$Shimmer(dB) = \frac{1}{N-1} \sum_{i=1}^{n-1} |\log_{20}(A_{i+1}/A_i)| \quad (5)$$

$$Shimmer(local) = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A_i - A_{i+1}|}{\frac{1}{N} \sum_{i=1}^N A_i} \quad (6)$$

donde A_i son los datos de amplitud pico a pico extraídos y N es el número de periodos extraídos de F0.

Se emplearon scripts para procesar vocales sostenidas provenientes de una muestra de 19 casos de pacientes previamente diagnosticados por un médico especialista procedente del Instituto Nacional de Rehabilitación con patologías de diferente origen y 19 casos de sujetos sanos. Entre ellas se encuentran patologías funcionales causadas por un mal manejo vocal que puede complicarse con lesiones orgánicas de la laringe, provocadas por el esfuerzo vocal (nódulos) o por alteraciones transitorias. Las patologías consideradas son producidas por lesiones en los órganos de la fonación o por causas congénitas. Se utilizó la configuración estándar de Praat para el manejo de archivos, con la ubicación, tal como se muestra en la Fig. 5.

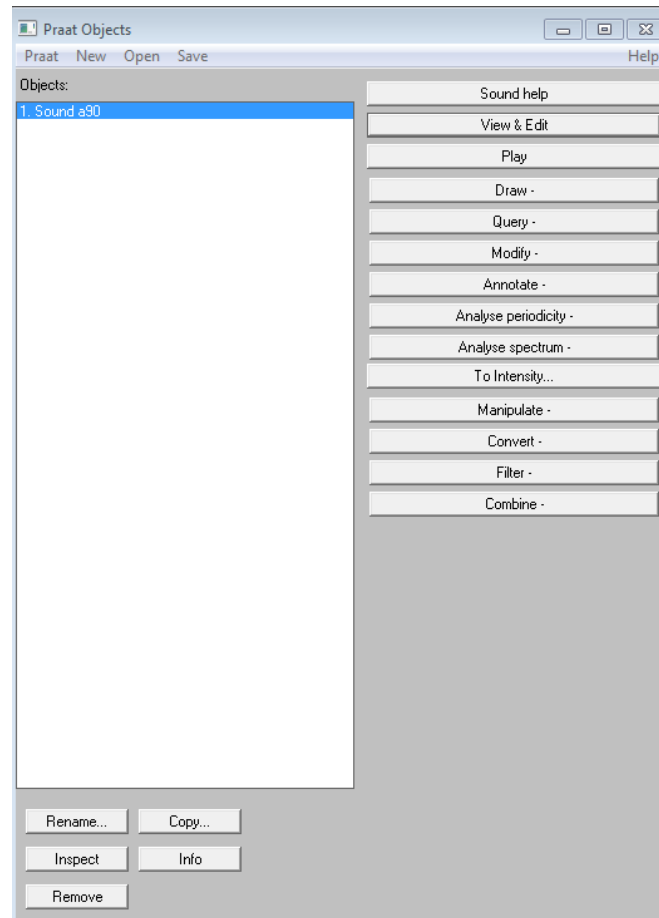


Fig. 5. Ventana de inicio para análisis acústico en Praat.

Las configuraciones que se utilizaron dentro de los *scripts* en Praat para la obtención de los parámetros fueron los siguientes:

- **Mediciones en frecuencia**

Select Sound [*Filename.wav*]

To Pitch... 0.0 50 500

Get mean... 0 0 Hertz

Get standard deviation... 0 0 Hertz

Con el comando *Select Sound [Filename.wav]* se selecciona el archivo de audio a analizar y para la obtención de la frecuencia fundamental se utiliza el comando *To Pitch*, el primer dato hace referencia al tiempo de inicio y fin del audio a analizar, el siguiente dato se refiere al límite inferior de la banda de frecuencia a analizar seguido

del límite superior. Para la obtención de la desviación estándar de la frecuencia fundamental se emplea el comando *Get estándar desviation*, el primer dato corresponde al tiempo de inicio y fin del audio a analizar seguido de la unidad de medida en la cual se solicita el parámetro.

- **Mediciones de *Shimmer***

```
Select Sound [Filename.wav]
```

```
To PointProcess (periodic, cc)... 75 500
```

```
Get shimmer (local)... 0 0 0.0001 0.02 1.3 1.6
```

```
Get shimmer (local_dB)... 0 0 0.0001 0.02 1.3 1.6
```

```
Get shimmer (apq3)... 0 0 0.0001 0.02 1.3 1.6
```

```
Get shimmer (apq5)... 0 0 0.0001 0.02 1.3 1.6
```

```
Get shimmer (dda)... 0 0 0.0001 0.02 1.3 1.6
```

Para la obtención de los parámetros de *Shimmer* se introduce el comando *Select Sound [Filename.wav]* para ingresar el archivo de audio a analizar, con el comando *To PointProcess (periodic, cc)* seguido de los límites superior e inferior de la banda de frecuencia a analizar este comando arroja las secciones periódicas que el programa interpreta como la secuencia de cierres glóticos en la vibración de las cuerdas vocales. Para la obtención del *Shimmer local* se utiliza el comando *Get shimmer (local)*, para la obtención del *Shimmer local dB* se emplea el comando *Get shimmer (local_dB)*, la obtención de *Shimmer apq3* se realiza a través del comando *Get shimmer (apq3)*, el *Shimmer apq5* se adquiere por medio de la instrucción *Get shimmer (apq5)*, por medio del comando *Get shimmer (dda)* todos estos seguidos del inicio y fin del audio a analizar, el intervalo más corto y largo posible que se utilizará en el cálculo en segundos y la mayor diferencia posible entre intervalos consecutivos.

- **Mediciones de *Jitter***

```
Select Sound [Filename.wav]
```

```
To PointProcess (periodic, cc)... 75 500
```

```
Get jitter (local)... 0 0 0.0001 0.02 1.3
```


Get jitter (local, absolute)... 0 0 0.0001 0.02 1.3

Get jitter (rap)... 0 0 0.0001 0.02 1.3

Get jitter (ppq5)... 0 0 0.0001 0.02 1.3

Get jitter (ddp)... 0 0 0.0001 0.02 1.3

Para la obtención de los parámetros de *Jitter* se introduce el comando *Select Sound [Filename.wav]* para ingresar el archivo de audio a analizar, con el comando *To PointProcess (periodic, cc)* seguido de los límites superior e inferior de la banda de frecuencia a analizar este comando arroja las secciones periódicas que el programa interpreta como la secuencia de cierres glóticos en la vibración de las cuerdas vocales. Para la obtención del *Jitter local* se utiliza el comando *Get jitter (local)*, para la obtención del *Jitter local absolute* se emplea el comando *Get jitter (local, absolute)*, la obtención de *Jitter rap* se realiza a través del comando *Get jitter (rap)*, el *Jitter ppq5* se adquiere por medio de la instrucción *Get jitter (pq5)*, por medio del comando *Get jitter (ddp)* todos estos seguidos del inicio y fin del audio a analizar, el intervalo más corto y largo posible que se utilizará en el cálculo en segundos y la mayor diferencia posible entre intervalos consecutivos.

- **Medición de ruido.**

Select Sound [Filename.wav]

To Harmonicity (cc)... 0.01 75 0.1 1

Get mean... 0 0

Get standard deviation... 0 0

Para la obtención de los parámetros de ruido se introduce el comando *Select Sound [Filename.wav]* para ingresar el archivo de audio a analizar, se emplea el comando *To Harmonicity (cc)* esta instrucción realiza una detección de periodicidad acústica sobre la base de un análisis de correlación cruzada directa este comando es seguido por el valor del rango en tiempo del audio a analizar, la longitud de la ventana de análisis en Hertz, el umbral de silencio que hace referencia a tramas que no contienen amplitudes por encima de este umbral y por último el número de períodos por ventana. Para el cálculo del parámetro de la relación armónico-ruido (HNR) se

emplea el comando *Get mean* seguido del rango en tiempo del audio a analizar. Para la obtención de la desviación estándar del HNR se utiliza la instrucción *Get standard deviation* seguido de la ventana en tiempo del audio a analizar.

- **Medición de la prominencia del pico cepstral suavizado**

```
Select Sound [Filename.wav]
```

```
To PowerCepstrogram: 50, 0.002, 5000, 40
```

```
Get CPPS: "yes", 0.02, 0.0005, 50, 330, 0.05, "Parabolic", 0.001, 0, "Exponential decay",  
"Robust"
```

Para la obtención del parámetro CPPS se introduce el comando *Select Sound [Filename.wav]* para ingresar el archivo de audio a analizar. Se obtiene el cepstrograma a través del comando *To PowerCepstrogram*, el cual es una representación en tiempo de *quefreny* en un sonido, este comando va seguido de la longitud de la ventana de análisis, la distancia entre los centros de ventanas posteriores y la frecuencia máxima sujeta a análisis. Seguido de este comando se aplica la instrucción *Get CPPS* seguido de si se sustraerá la línea de regresión suavizada, el tiempo promedio de la ventana, el rango de búsqueda de picos en frecuencia, la tolerancia, el tipo de interpolación, el rango de frecuencia de la línea de regresión, tipo de trazo y método de ajuste.

4.2.3 Análisis estadístico

Para el análisis estadístico se utilizó una muestra de 19 sujetos masculinos sanos y 19 pacientes con patología asociadas a las cuerdas vocales.

Se obtuvieron 15 parámetros del análisis acústico para posteriormente realizar un análisis estadístico los cuales se muestran y describen en la Tabla 3.

Tabla 3. Lista de los parámetros obtenidos en Praat.

Etiqueta	Nombre	Unidades	Definición
F0	Frecuencia fundamental	Hz	Cantidad de veces que vibra la cuerda vocal en un segundo.
F0 SD	Frecuencia fundamental	Hz	Variación en la frecuencia fundamental.
Jitter local absoluto	Variabilidad absoluta de la f0	µs	Diferencia absoluta media entre periodos consecutivos.
Jitter local	Variabilidad de la f0	%	Diferencia absoluta promedio entre períodos consecutivos, dividida por el período promedio.
Jitter rap	Perturbación media relativa	%	Mide los cambios relativos de la F0.
Jitter ppq5	Cociente de perturbación del período de cinco puntos	%	Evaluación relativa de la variabilidad de la f0, periodo a periodo en 5 ciclos.
Jitter ddp	Diferencia de diferencias de períodos	%	Diferencia absoluta promedio entre las diferencias consecutivas entre períodos consecutivos, dividida por el período promedio.
Shimmer local	Perturbación en amplitud	%	Diferencia absoluta promedio entre las amplitudes de períodos consecutivos, dividida por la amplitud promedio.
Shimmer dB	Perturbación en amplitud medida en decibeles	dB	Evaluación en dB de periodo a periodo (tiempos cortos) de la variabilidad de la amplitud pico a pico.
Shimmer apq3	Perturbación en amplitud medida en tres periodos consecutivos	%	Diferencia absoluta promedio entre la amplitud de un período y el promedio de las amplitudes de sus puntos colindantes, dividido por la amplitud promedio.
Shimmer apq5	Perturbación en amplitud medida en cinco periodos consecutivos	%	Diferencia absoluta promedio entre la amplitud de un período y el promedio de las amplitudes de este y sus cuatro puntos más cercanos, dividido por la amplitud promedio.
Shimmer ddp	Diferencia absoluta promedio de amplitud	%	Diferencia absoluta promedio entre diferencias consecutivas entre las amplitudes de períodos consecutivos.
HNR	Relación ruido-armónico	dB	Medición de la relación entre la energía acústica de los armónicos y la energía del ruido agregado del tracto vocal durante la fonación.
HNR SD	Desviación estándar de la relación ruido armónico	dB	Variación del HNR.
CPPs	Pico cepstral de mayor prominencia suavizado	dB	Diferencia entre el pico cepstral más prominente, que corresponde al primer armónico, y el punto con la misma cuefrecuencia sobre la línea de regresión a través del cepstrum suavizado.

Utilizando el programa Graph Pad PRISM®, cada uno de los parámetros fue sometido a un análisis de normalidad y posteriormente se realizó un análisis *Mann Whitney test* de manera individual para conocer las dispersiones de los datos y observar si éstos presentaban diferencias significativas entre los datos de pacientes con alguna patología y los datos de personas sanas.

Se realizó un análisis de la variación de dos vías (*Two-way ANOVA*) para comprobar la efectividad de los parámetros en conjunto de acuerdo con si estos presentaban una variación estadísticamente significativa entre el grupo de sujetos sanos y el de pacientes con patología.

4.2.4 Diseño del clasificador

Posterior al análisis estadístico se empleó una red neuronal para clasificar los parámetros de los dos grupos, implementada en Matlab®. Con base en los grupos identificados en el análisis previo, se entrenó la red para identificar las dos condiciones.

La red neuronal fue diseñada con ayuda del *toolbox* integrado en Matlab®. El procesamiento incluyó la preparación de los datos, definición de las capas y número de capas, así como el método de validación.

La red neuronal diseñada se basa en el reconocimiento de patrones para clasificación y consiste en la capa de entrada con datos de los parámetros obtenidos del análisis en Praat, una capa oculta de 10 neuronas y una capa con una salida binaria.

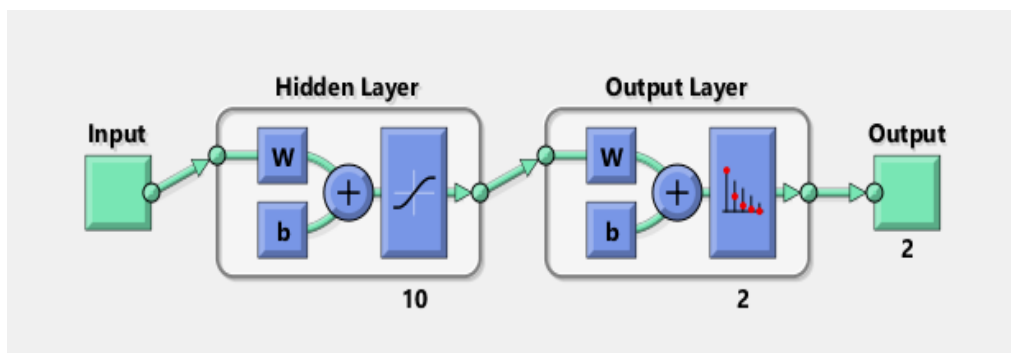


Fig. 6. Diseño de la red neuronal en Matlab®.

La base de datos conformada por 38 sujetos se dividió en 70% para entrenamiento, 15% para validación y 15% para pruebas, seleccionados aleatoriamente

en el arreglo de la red neuronal. La evaluación se llevó a cabo mediante el cálculo de matrices de confusión y permitió evaluar el error en el entrenamiento de la red.

4.2.5 Diseño a bloques del sistema propuesto

En la Fig. 7 se presenta un diagrama a bloques del sistema propuesto para la detección temprana de patologías basado en mediciones acústicas, aerodinámicas y mecánicas de manera ambulatoria, el cual inicia con el llenado de formularios que contendrán la información general del paciente, así como los cuestionarios correspondientes, seguido del posicionamiento de los sensores y la toma de las señales que serán adquiridas y guardadas en una memoria microSD la cual de manera manual deberá introducirse a la computadora para posteriormente descargar el archivo creado para realizar el análisis y la clasificación.

Por último, las gráficas y los resultados obtenidos serán presentados en una ventana de la interfaz gráfica.

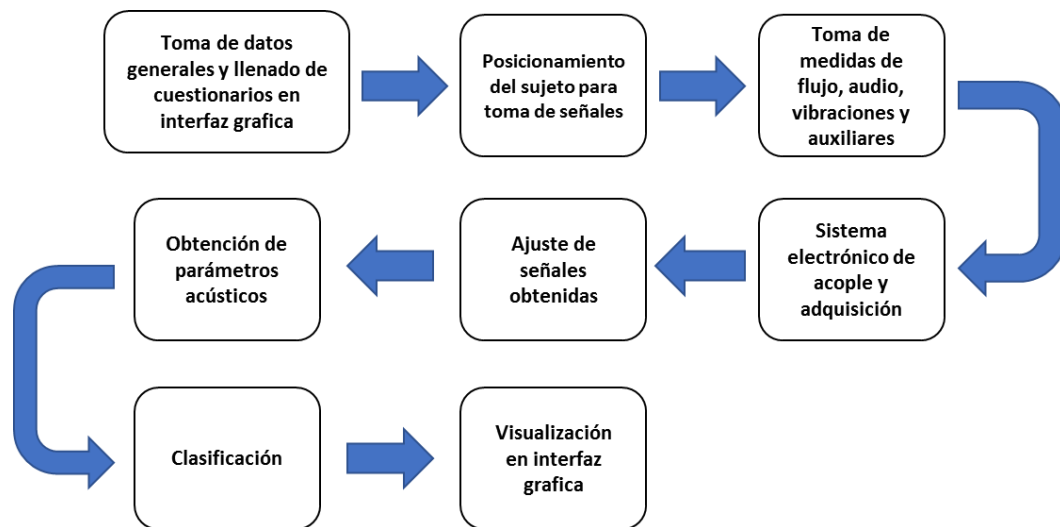


Fig. 7. Diagrama de flujo del bloque 1 de la Fig. 4.

En la Fig. 8 se muestra el montaje de los sensores y su distribución de manera gráfica para la toma de las señales. Los sensores principales (flujo y micrófono) van montados en diferentes puntos de una mascarilla de plástico transparente (Hudson RCI, China). Los sensores se conectan a un sistema de adquisición electrónica, seguido de la transferencia de los datos obtenidos por comunicación serial a una computadora y desplegados en una interfaz gráfica diseñada en Matlab para su análisis.

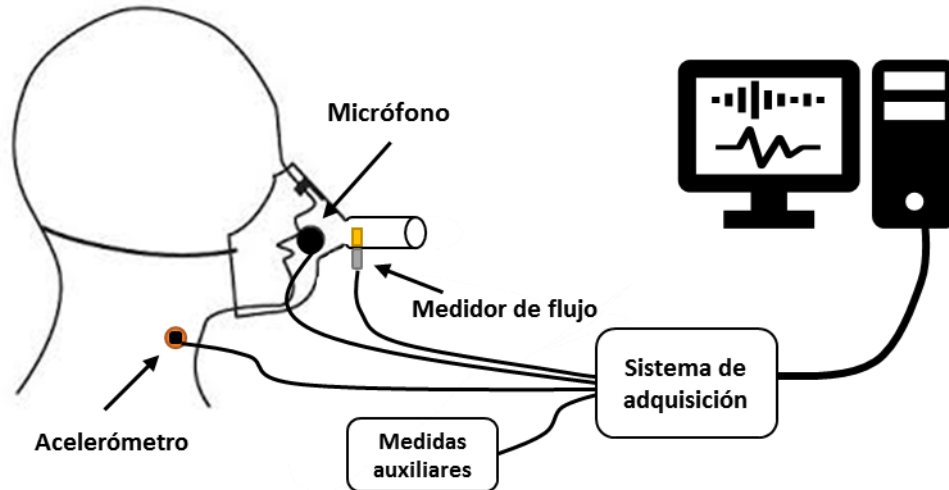


Fig. 8. Montaje de sensores y distribución grafica del sistema.

4.2.6 Diseño del sistema electrónico

Como parte del diseño de un sistema de evaluación se propone la construcción de un prototipo para la obtención y registro de las señales a evaluar. Se elaboró un diagrama con los elementos propuestos para esta aplicación el cual se muestra en la Fig. 9.

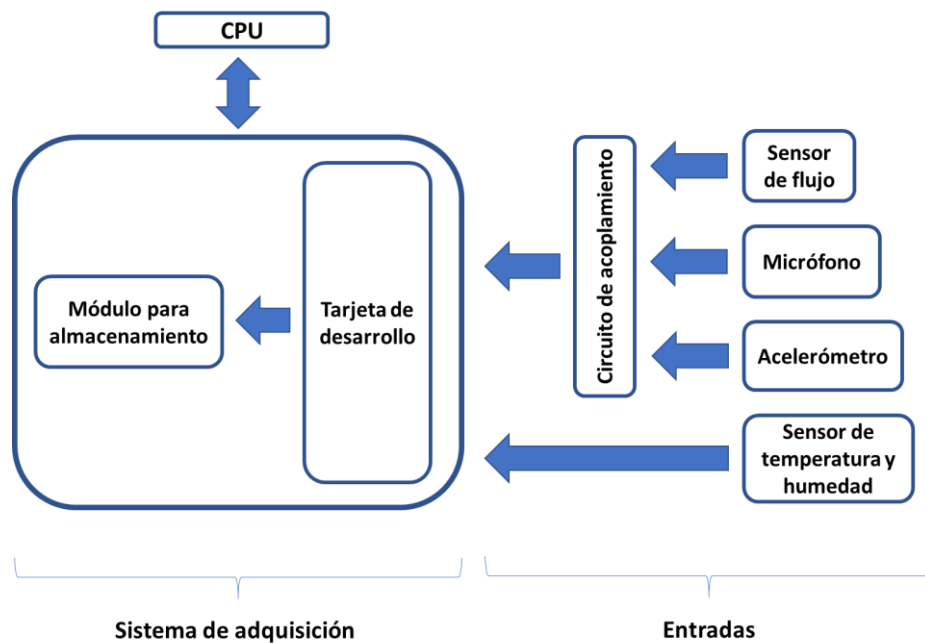


Fig. 9. Diagrama a bloques del sistema electrónico propuesto.

En primera instancia se seleccionaron los sensores para la obtención de las medidas. Se seleccionaron cuatro sensores que se detallan más adelante, que

cumplieran con las características adecuadas para el diseño y con los requerimientos planteados en la solución propuesta.

Sensor de flujo

Para la obtención del flujo de aire se seleccionó el sensor *Thermal Mass Flow Sensor FS5* (Innovative Sensor Technology iST AG, Switzerland) que se muestra en la Fig. 10, debido a sus dimensiones reducidas y fácil manejo.



Fig. 10. Sensor de flujo Thermal Mass Flow Sensor FS5.

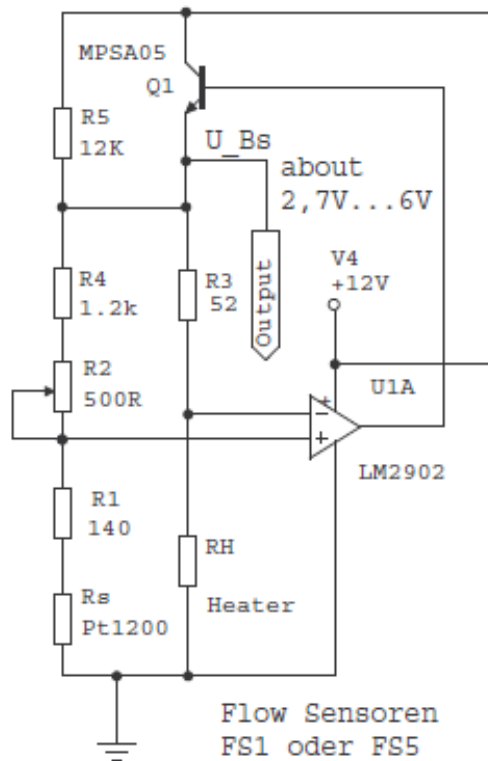


Fig. 11. Circuito de acoplamiento propuesto por el fabricante.

Para el sensor de flujo se comenzó obteniendo una señal a partir del circuito de acoplamiento sugerido por el fabricante Fig. 11.

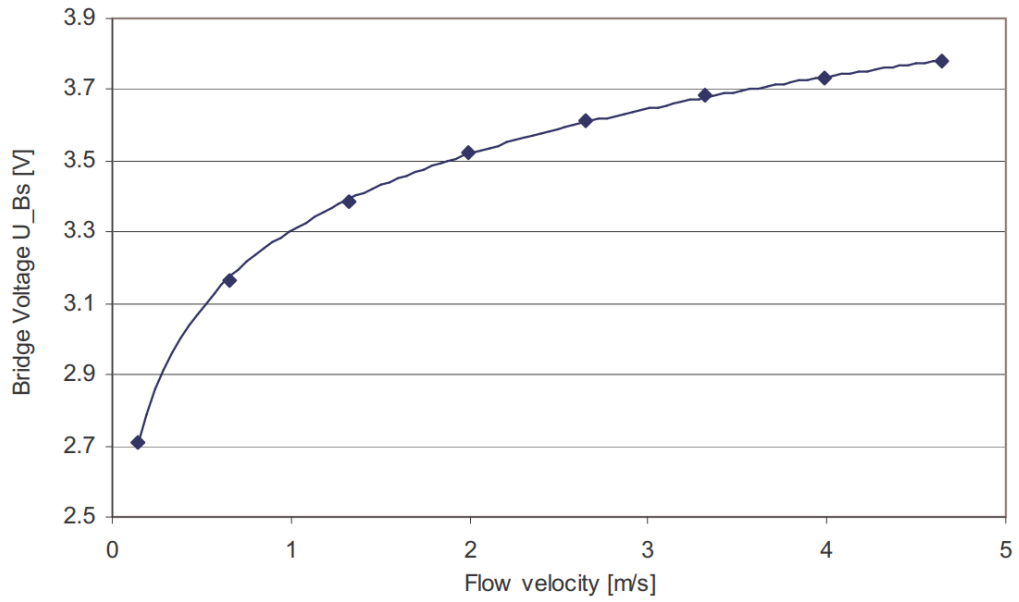


Fig. 12. Gráfica de respuesta del sensor proporcionada por el fabricante.

Sin embargo, el circuito no proporciona la señal deseada, además de tener un valor considerable de offset, por lo que se requirió de una etapa para compensar el valor obtenido para cada sensor.

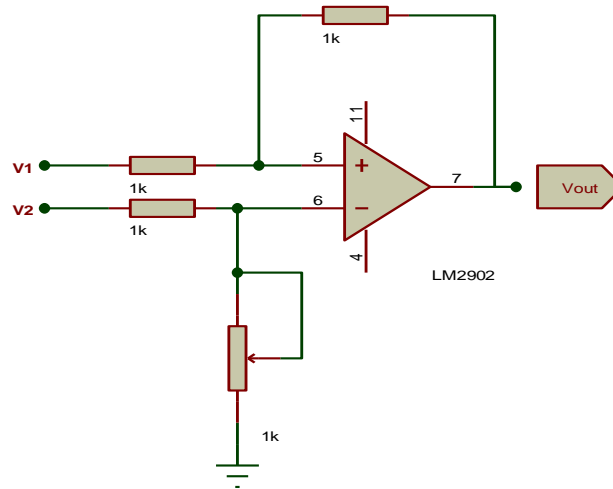


Fig. 13. Circuito restador para eliminación de *offset*.

Se introduce un voltaje de DC a la entrada $V_2 = 2\text{Volts}$ del circuito restador de la Fig. 13, controlado por un potenciómetro para compensar el *offset* que posee la señal. El voltaje de salida responderá de acuerdo con la ecuación (7).

$$V_{out} = V_2 - V_1 \quad (7)$$

Se implementó una etapa de amplificación de la señal. Para esta etapa se utilizó un circuito amplificador no inversor que se muestra en la Fig. 14. La ganancia fue ajustada con resistencias iguales a un valor de $1\text{k}\Omega$ para la obtención de una ganancia de dos basado en la ecuación 8.

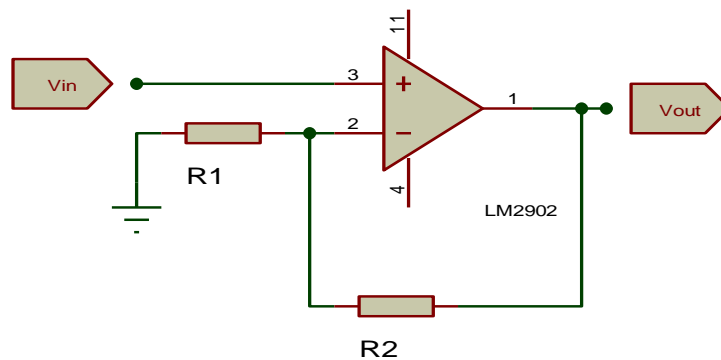


Fig. 14. Esquema de circuito amplificador no inversor.

$$V_{out} = V_{in} \left(1 + \frac{R_2}{R_1} \right) \quad (8)$$

Las etapas se unieron en un circuito final mostrado en la Fig. 15.

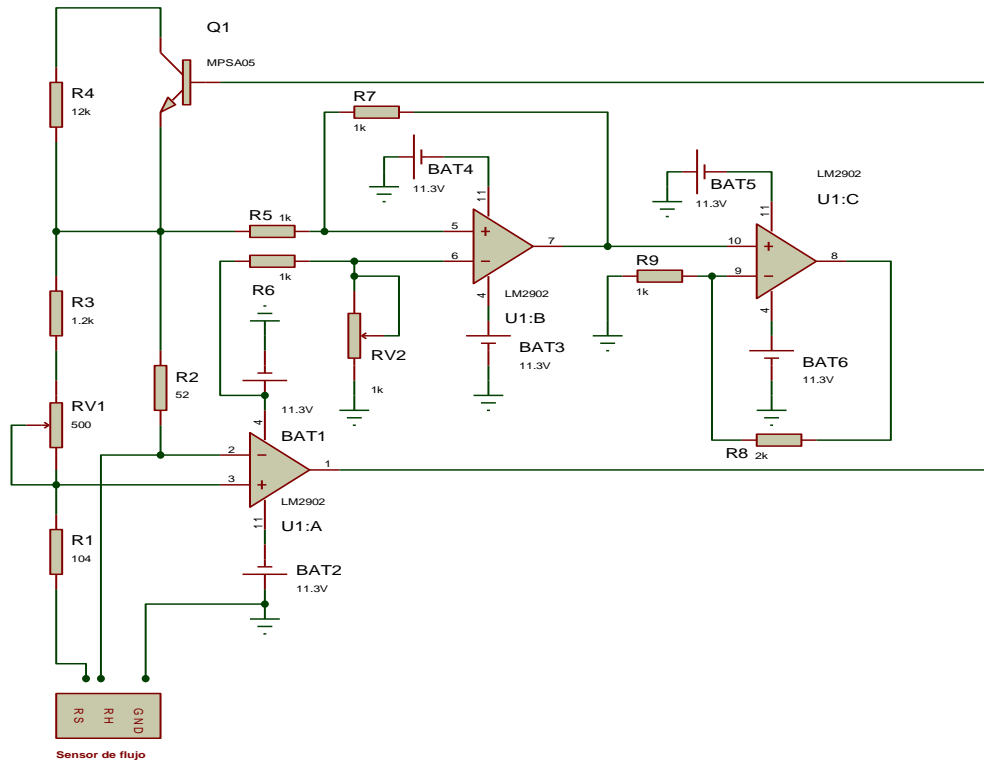


Fig. 15. Circuito para acoplamiento del sensor de flujo.

Para la alimentación del sistema se utilizaron dos baterías tipo Li-Po de tres celdas de 11.1V, 1300mAh.



Fig. 16. Batería tipo Li-Po de tres celdas de 11.1Volts.

Micrófono

Se seleccionó el micrófono Electret modelo CMC5044PF (CUI INC, USA) mostrado en la Fig. 17 el cual presenta un ancho de banda de 50 a 15kHz y una sensibilidad de -70dB, poco sensible a la humedad y calor. Por su tamaño reducido es recomendado en el diseño de sistemas portátiles y de bajo costo.



Fig. 17. Micrófono Electret CMC5044PF.

Para la obtención de una señal acústica, el fabricante propone el circuito mostrado en la Fig. 18. El capacitor tiene un valor de $1\mu\text{F}$ y la resistencia un valor de $2.2\text{k}\Omega$; la alimentación recomendada es de 3 Volts.

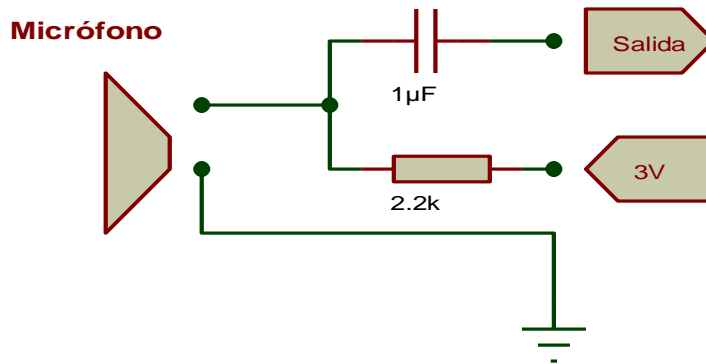


Fig. 18. Circuito de medición para micrófono $R1=2.2\text{k}\Omega$; $C1=1\mu\text{F}$.

Se utilizó un módulo de preamplificación para la optimización de la señal de audio modelo MAX4466 de la compañía *Maxim Integrated Products*.

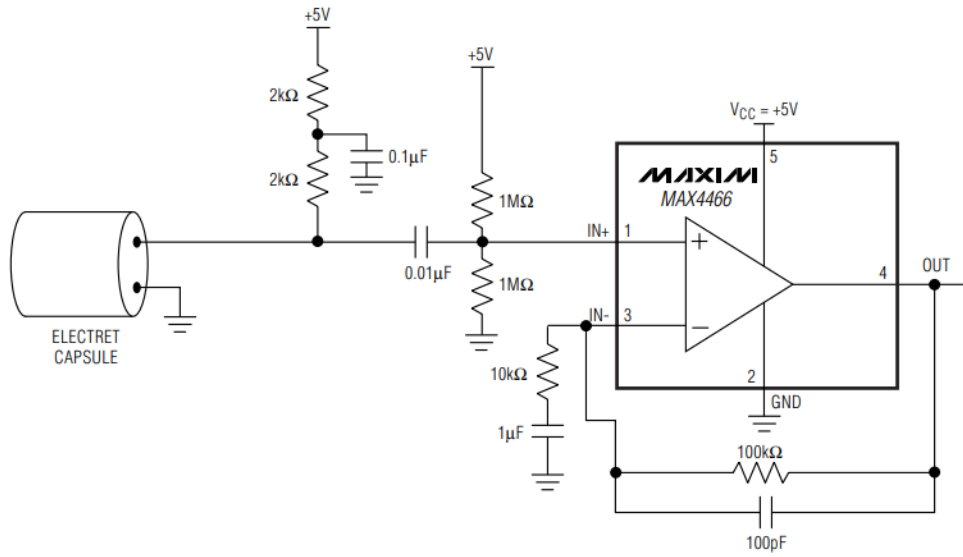


Fig. 19. Preamplificador para el micrófono Electret [36].



Fig. 20. Módulo Max4466 para preamplificación del micrófono.

Este módulo se seleccionó debido a su reducido consumo en corriente, así como un voltaje de alimentación bajo (Tabla 4).

Tabla 4. Características técnicas del módulo Max4466.

Dimensiones	2cmx1.3cm
Consumo de corriente	24 μ A
Voltaje de operación	+2.4V a +5.5V

Sistema de detección de vibración mecánica

Se seleccionó el acelerómetro MMA7361L que se muestra en la Fig. 21 para la detección y registro de las vibraciones en la superficie del cuello.

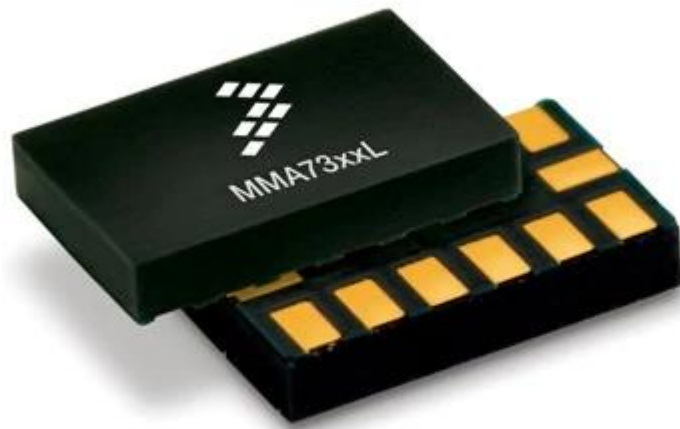


Fig. 21. Acelerómetro MMA7361L.

Se seleccionó el acelerómetro con base a sus características descritas por el fabricante mostradas en la Tabla 5, el cual irá colocado en la parte inferior izquierda de la superficie del cuello durante el estudio [25].

Tabla 5. Características técnicas del acelerómetro MMA7361L.

Dimensiones	3mm x 5mm x 1.0mm
Consumo de corriente	400 μ A
Sensibilidad	800 mV/g @ 1.5g
Voltaje de operación	2.2 V – 3.6 V
Tiempo de respuesta	0.5 ms

Para la obtención de la señal se utilizó la recomendación de conexión del sensor proporcionado por el fabricante, tal como se muestra en la Fig. 22.

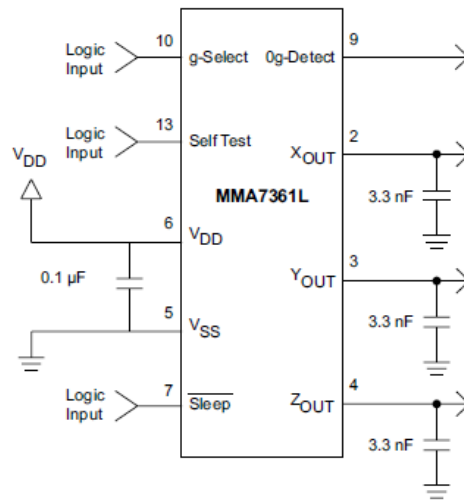


Fig. 22. Circuito de conexión para el acelerómetro.

4.2.7 Sistemas complementarios en el protocolo de registro

Sensor de temperatura y humedad

Con el fin de garantizar el funcionamiento adecuado de los sensores utilizados y para reportar las condiciones de registro, se mantuvieron monitoreadas la temperatura y humedad del ambiente. Para este propósito se seleccionó un sensor de humedad y temperatura digital DHT11 mostrado en la Fig. 23.

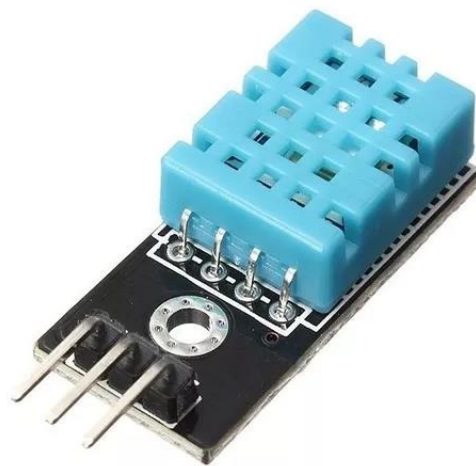


Fig. 23. Sensor de humedad y temperatura DHT11.

Se seleccionó este dispositivo de acuerdo con sus características técnicas mostradas en la Tabla 6.

Tabla 6. Características técnicas del sensor DHT11 [37].

MODELO	DHT11
Alimentación	3,5 V a 5 V
Consumo	2,5 mA
Señal de salida	Digital
Temperatura	
Rango	0°C a 50°C
Precisión	25°C ± 2°C
Resolución	1°C (8-bit)
Humedad	
Rango	20% RH a 90% RH
Resolución	0°C y 50°C ± 5% RH
Resolución	1% RH

4.2.8 Montaje de sensores

El micrófono se colocó en el cuerpo de la mascarilla y el de flujo en una extensión de la boquilla de metacrilato agregada en el tubo de salida de la mascarilla.

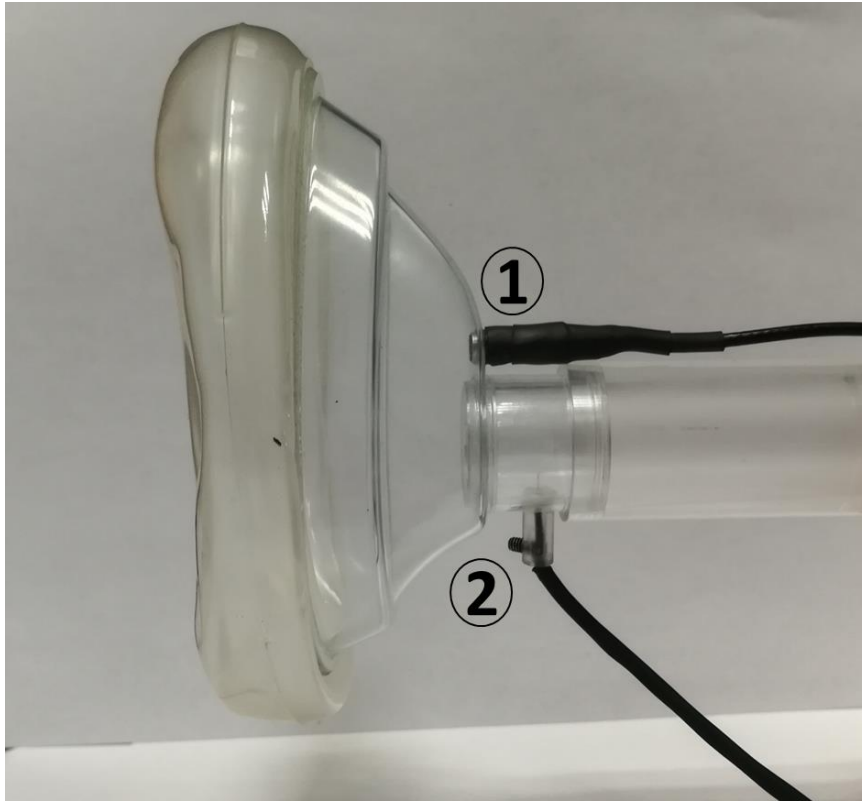


Fig. 24. Mascarilla con extensión y sensores de audio y flujo montados.

En la Fig. 24 se muestra la ubicación de los sensores en la mascarilla; número 1, micrófono alineado con la boca del sujeto; número 2, sensor de flujo introducido al tubo de salida de la mascarilla.

4.2.9 Adquisición y almacenamiento de datos

Para la adquisición de los datos se utilizó la tarjeta de desarrollo ESP32 (Fig. 25), creada y desarrollada por la empresa Espressif Systems. Fue seleccionada por su bajo consumo de corriente y su velocidad (

Tabla 7).

Esta tarjeta de desarrollo emplea un microprocesador Tensilica Xtensa LX6 y está diseñada para dispositivos móviles y dispositivos electrónicos portátiles.



Fig. 25. Tarjeta de desarrollo ESP32.

Tabla 7. Características técnicas de la tarjeta de desarrollo EP32.

Procesador	Tensilica Xtensa X36
Nº bits	32 bits
Nº núcleos	Dual core
Velocidad	160 MHz (hasta 240 MHz)
SRAM	512 kB
SPI FLASH	Hasta 16MiB
Alimentación	2.2 a 3.6V
Rango temperaturas	-40°C a 125°C
Consumo de corriente	80 mA (promedio), 225 mA (máximo)
Consumo en deep sleep	2.5 uA (10 uA RTC + memoria RTC)
ADC	18 (12 bits)
DAC	2 (8 bits)
UART	4
GPIO	36

Para la adquisición se utilizaron tres canales analógicos del manejador de entradas del dispositivo GPIO (13,14,26) para los sensores de flujo, audio, y vibraciones respectivamente, a una frecuencia de muestreo de 8kHz, controlada por un *timer* y un canal digital GPIO (27) para el sensor de temperatura y humedad. Este dispositivo se programó utilizando la IDE de Arduino.

Se empleó un módulo SD para el almacenamiento de datos por medio de transferencia SPI en una memoria microSD con capacidad de almacenamiento de 2GB.

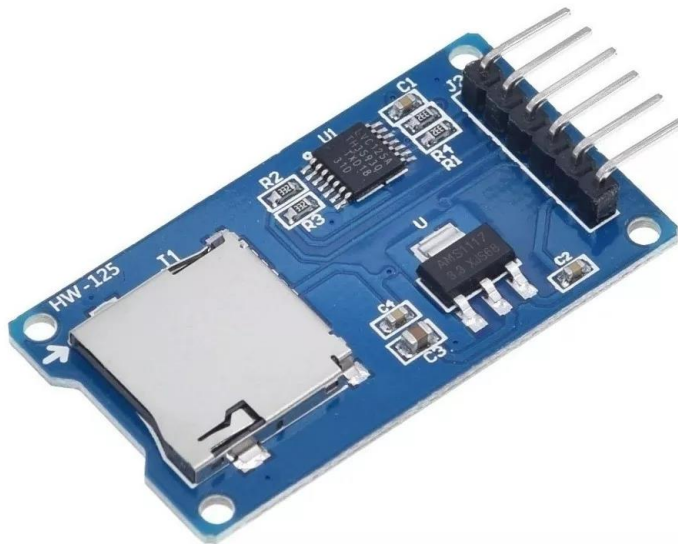


Fig. 26. Módulo microSD para almacenamiento de datos.

El módulo fue conectado a los canales de comunicación SPI como se muestra en la Fig. 27. La información se guarda en archivos de texto con extensión .txt.

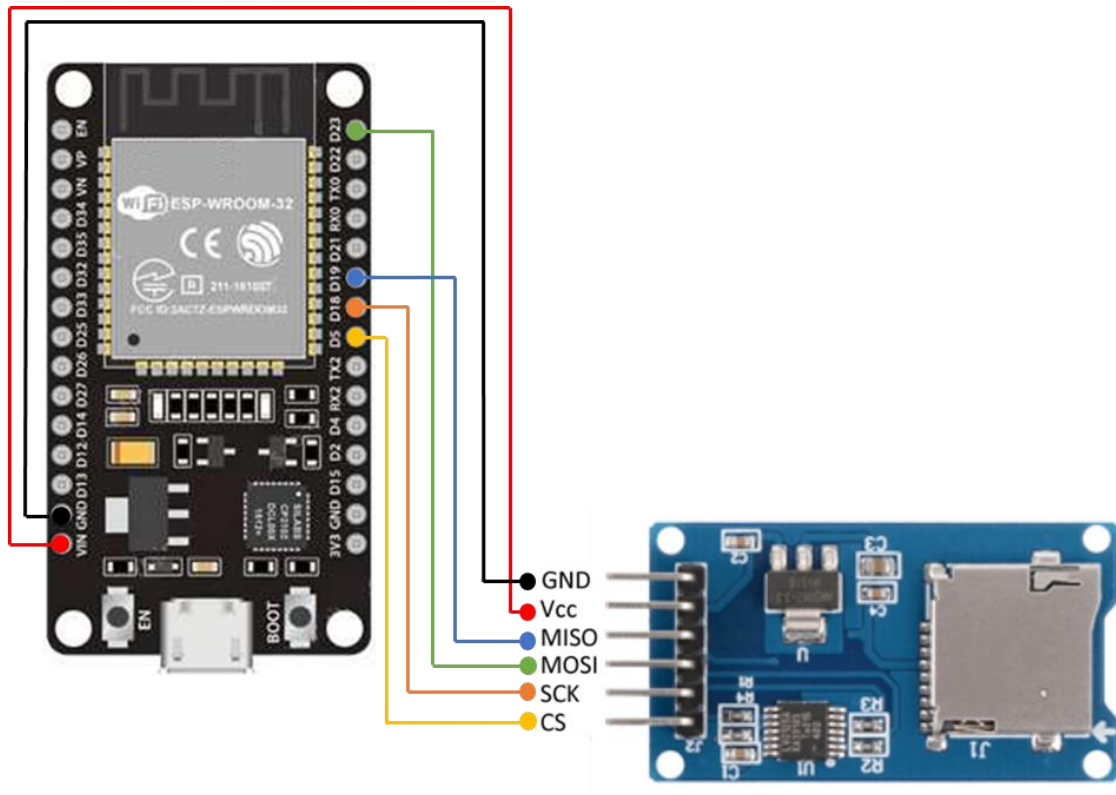


Fig. 27. Conexión de módulo microSD a ESP32.

4.2.10 Interfaz gráfica

La interfaz gráfica se desarrolló en el software de Matlab, siguiendo el diagrama a bloques mostrado en la Fig. 28.

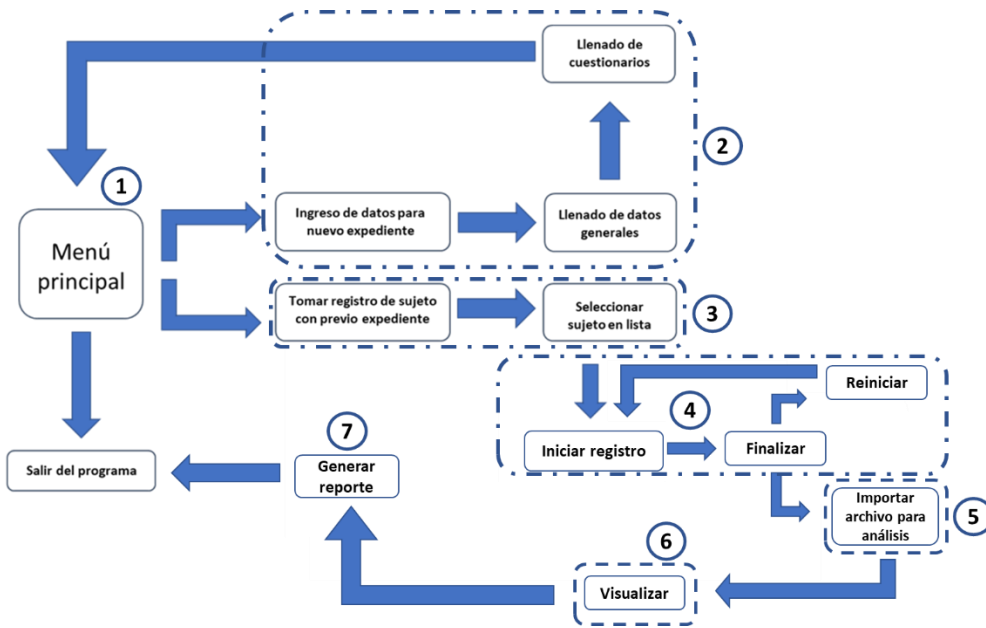


Fig. 28. Diagrama a bloques para el desarrollo de la interfaz gráfica.

A continuación, se presenta una descripción de cada uno de los bloques numerados.

1.- Menú principal: Se diseño un menú principal con tres opciones principales; el ingreso de un nuevo expediente; registro de señales para análisis y salir del programa.

2.- Bloque par creación de expediente: Para la creación de un nuevo expediente se propone la creación de ventanas con cuestionarios diseñados para la recolección de datos generales como nombre completo, edad, sexo, y ocupación. Una nueva ventana en la cual se propone el uso del cuestionario *Voice Handicap Index-10* (VHI-10), Índice de Incapacidad Vocal-10 que se presenta en la Fig. 29, donde: 0=nunca, 1=casi nunca, 2=a veces, 3=casi siempre, 4=siempre, tomado de [38]. El VHI-10 es un instrumento para cuantificar la percepción de los pacientes sobre su discapacidad vocal este cuestionario es una representación del VHI que requiere menos tiempo para que el paciente se complete sin pérdida de validez.

F1. La gente me oye con dificultad debido a mi voz	0	1	2	3	4
F2. La gente no me entiende en sitios ruidosos	0	1	2	3	4
F8. Mis problemas con la voz alteran mi vida personal y social	0	1	2	3	4
F9. Me siento desplazado de las conversaciones por mi voz	0	1	2	3	4
F10. Mi problema con la voz afecta al rendimiento laboral	0	1	2	3	4
P5. Siento que necesito tensar la garganta para producir la voz	0	1	2	3	4
P6. La calidad de mi voz es impredecible	0	1	2	3	4
E4. Mi voz me molesta	0	1	2	3	4
E6. Mi voz me hace sentir cierta minusvalía	0	1	2	3	4
P3. La gente me pregunta: ¿qué te pasa con la voz?	0	1	2	3	4

Fig. 29. Cuestionario VHI-Índice de capacidad vocal [38].

3.- Bloque para selección de sujeto para registro: En la tarea para realizar una nueva toma de señales se diseñó una ventana con el listado de sujetos que se encuentren con un expediente completo.

4.- Bloque para toma de registro: La ventana de registro contiene botones para inicio y fin del proceso de registro. Una vez finalizado se hará visible el botón de reinicio y de importar archivo.

5.- Importa archivo: El diagrama de flujo del botón de importar archivo para análisis se muestra en la

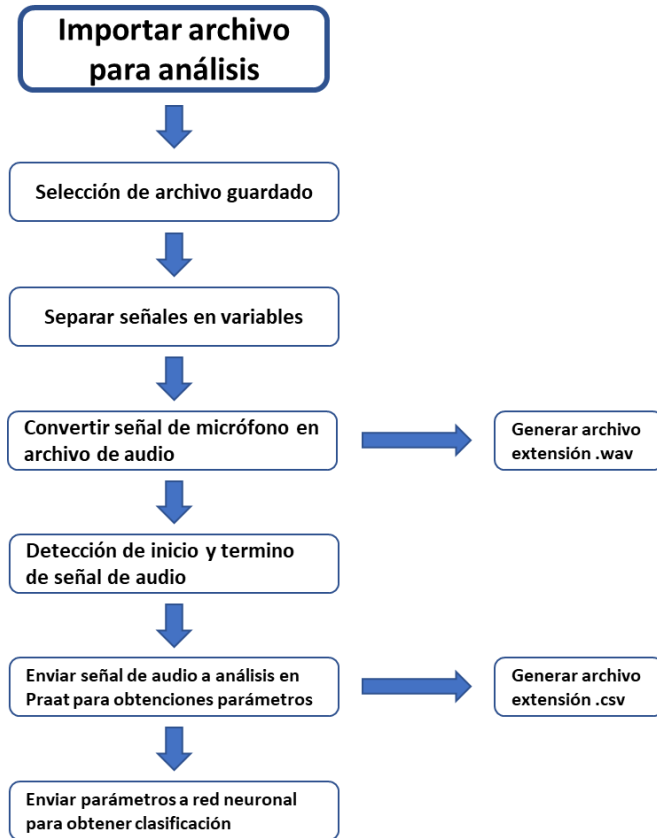


Fig. 30. Diagrama de flujo de bloque 5 de la.Fig. 28.

6.- Visualizar: Se visualizará un botor para abrir una nueva ventana el despliegue de resultados del análisis el diagrama de este proceso se muestra en la Fig. 31.

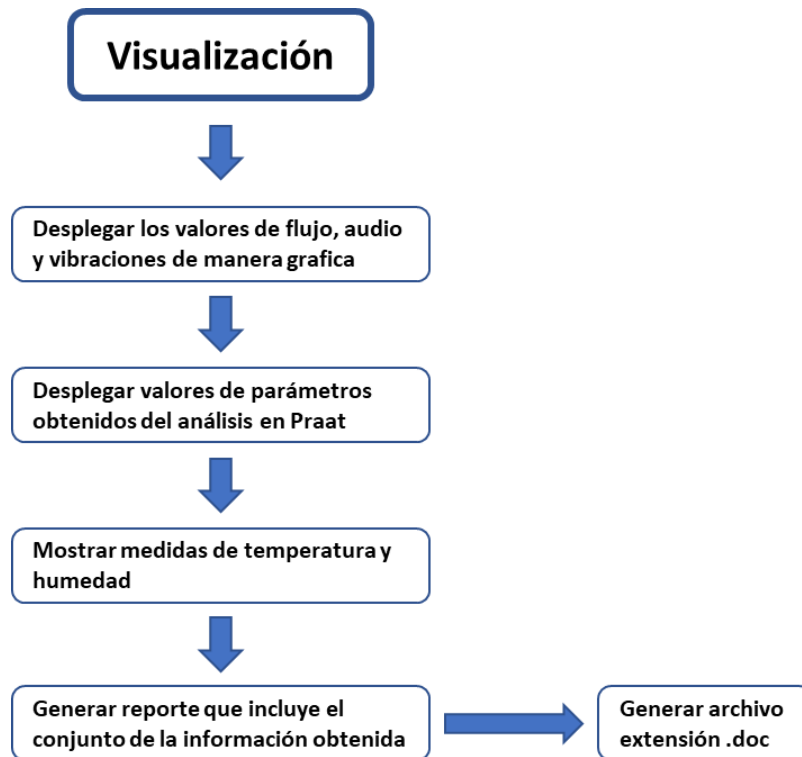


Fig. 31. Diagrama de flujo de bloque 6 de la Fig. 28.

7.- Por último, en la ventana de visualización se anexa la opción de generar un reporte escrito con los datos recolectados y los resultados obtenidos del análisis.

Capítulo 5

5.1 Resultados y discusiones

5.1.1 Análisis acústico

Del análisis acústico realizado en Praat a partir del banco de señales proporcionado, conformado por 19 sujetos con patologías de diferentes orígenes y 19 sujetos sanos, se obtuvieron los siguientes valores de los parámetros seleccionados, observados en la Tabla 8.

Tabla 8. Valores promedio obtenidos del análisis acústico en Praat.

Parámetro	Sano		Patológico	
	Promedio	σ	Promedio	Σ
Frecuencia fundamental f0 (Hz)	127.2	14.19	183.7	77.41
Desviación estándar de f0 (Hz)	1.526	0.827	26.39	40.55
<i>Jitter</i> local (%)	0.395	1.713	1.7132	1.507
<i>Jiter</i> local absoluto (μs)	31.47	113.8	113.82	128.1
<i>Jitter</i> rap (%)	0.22	1.01	1.0101	0.94
<i>Jitter</i> ppq5 (%)	0.236	1.023	1.0228	0.889
<i>Jiter</i> ddp (%)	0.66	3.03	3.0304	2.819
<i>Shimmer</i> local (dB)	0.397	1.567	1.5672	0.297
<i>Shimmer</i> local (%)	4.382	18.01	18.010	3.833
<i>Shimmer</i> apq3 (%)	2.337	9.021	9.0209	2.081
<i>Shimmer</i> apq5 (%)	2.812	11.51	11.511	2.97
<i>Shimmer</i> dda (%)	7.012	27.06	27.062	6.242
CPPS (dB)	15.98	8.529	8.529	2.721
HNR (dB)	18.89	6.415	6.415	3.655
Desviación estándar del HNR (dB)	3.257	2.776	2.776	0.608

Se seccionó la muestra entre sujetos sanos y sujetos con alguna patología para obtener los rangos en los que se encontraba cada grupo. Se observó una diferencia entre las medias aritméticas de cada uno de los parámetros, además que la desviación estándar de cada parámetro resultó mayor en el caso de los sujetos con patología en contraste con los sujetos sanos.

5.1.2 Análisis estadístico

Para el análisis estadístico se utilizó el programa Graph Pad PRISM®, los parámetros fueron separados en grupos con respecto a sus rangos y unidades de medida para observar gráficamente la concentración y la distribución de los datos y evaluar las diferencias de acuerdo con la condición de salud reportada.

El grupo uno está conformado únicamente por la frecuencia fundamental (F0), cuyo valor se elevó en los pacientes con patologías por encima de los 350 Hz, como se muestra en la gráfica de la Fig. 32.

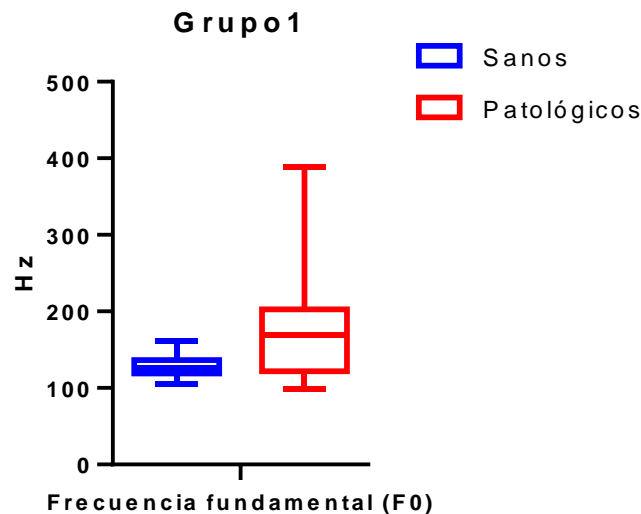


Fig. 32. Distribución de valores obtenidos del análisis acústico del grupo 1.

En el grupo dos mostrado gráficamente en la Fig. 33 incluye la relación armónico-ruido (HNR) y el punto cepstral de mayor prominencia (CPPS) los cuales tienen una escala en dB. Con respecto a la relación armónico-ruido se aprecia que los valores obtenidos de las señales acústicas de personas sanas se concentran por debajo de los 10 dB, por otro lado, la concentración en los sujetos con una patología se coloca

por encima de los 15 dB. Referente al punto cepstral de mayor prominencia se observa que en el caso de las personas sanas el valor se mantiene arriba de los 13 dB, a diferencia de los casos patológicos los cuales muestran que el valor se concentra entre los 5 dB y 10 dB.

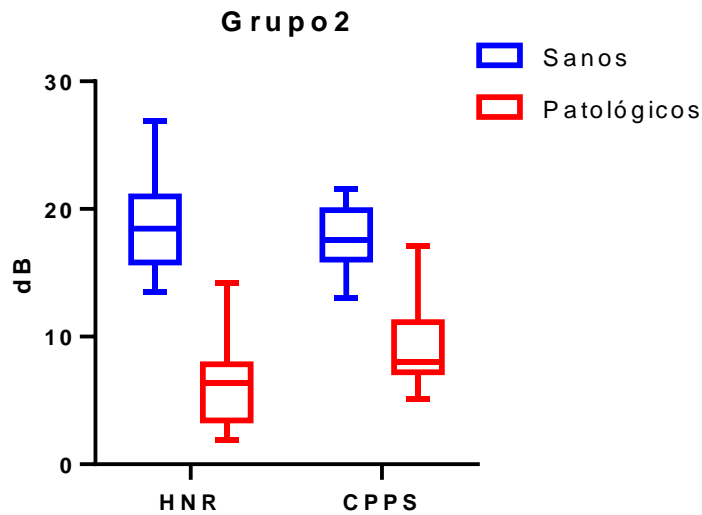


Fig. 33. Distribución de valores obtenidos del análisis acústico del grupo 2.

En el grupo tres se localiza la desviación estándar de la relación armónico-ruido (SD HNR), en este caso no se encontró una diferencia estadísticamente significativa con respecto a los parámetros anteriores como se muestra en la Fig. 34. Esta medición no se agrupo con las de unidades dB, porque es una medida de dispersión sobre la tendencia central, y las anteriores eran valores de tendencia central.

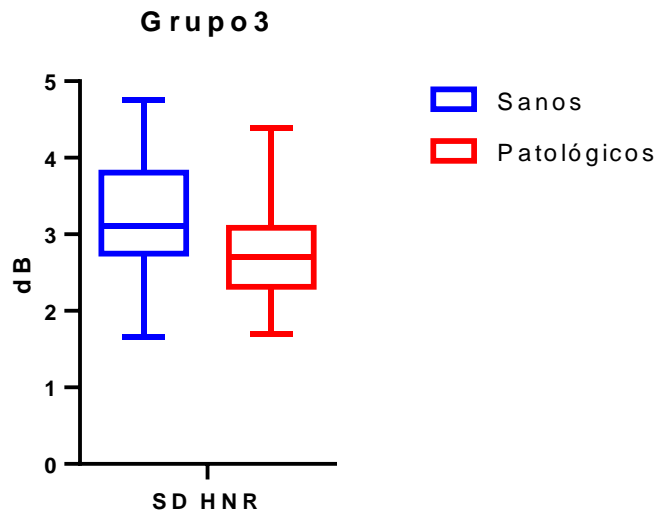


Fig. 34. Distribución de valores obtenidos del análisis acústico del grupo 3.

En la Fig. 35 se muestra el grupo con la desviación estándar de la frecuencia fundamental (F0 SD) la cual en el caso de las personas sanas no se obtiene un valor mayor a los 10 Hz, a diferencia de los pacientes con alguna patología en la cual el valor alcanza valores arriba de 100 Hz, para algunos de los casos.

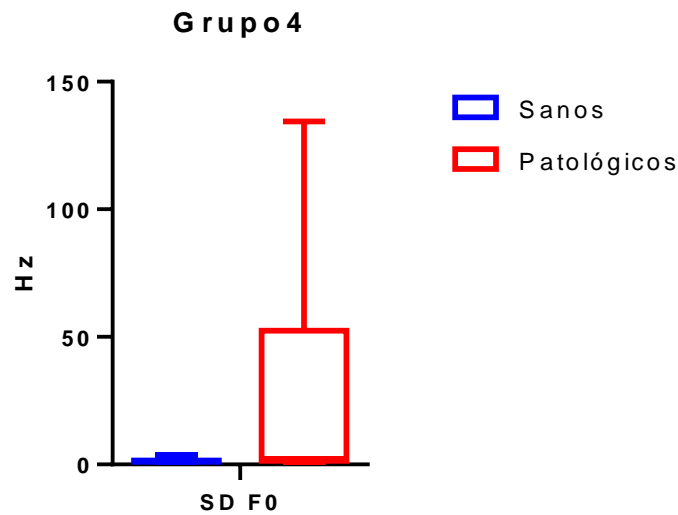


Fig. 35. Distribución de valores obtenidos de análisis acústico del grupo 4.

En el grupo cinco se encuentran las medidas de *Shimmer local*, *apq3*, *apq5* y *dda*, las cuales son medidas de perturbación de la amplitud de la fonación. Estos parámetros muestran diferentes distribuciones lo cual hace a estos parámetros de gran utilidad para la clasificación ya que se observa una diferencia importante en la

concentración de valores de los grupos sanos con respecto a los patológicos como se muestra en la Fig. 36.

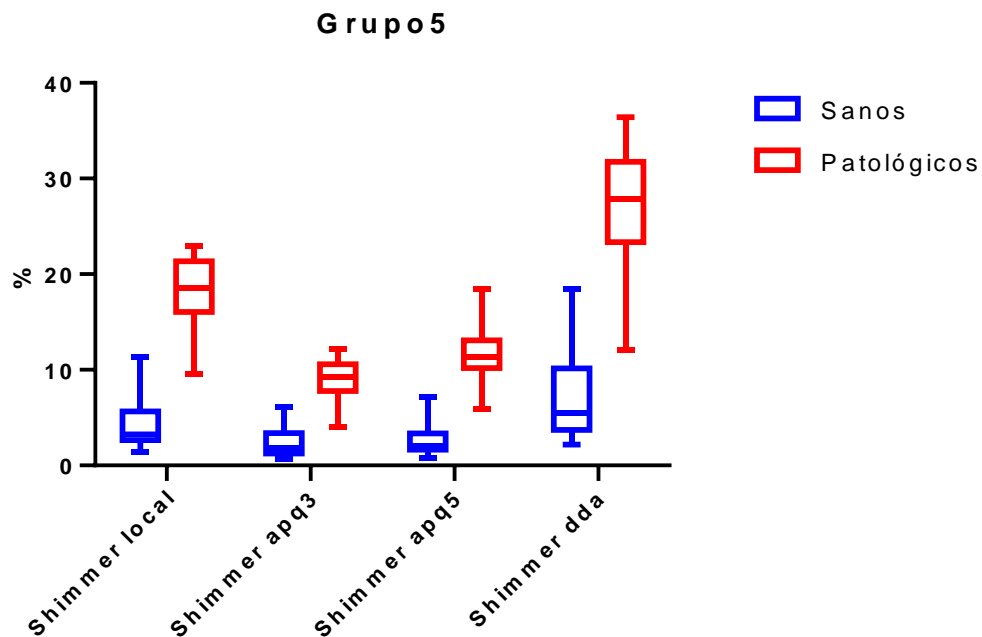


Fig. 36. Distribución de valores obtenidos del análisis acústico del grupo 5.

En general, los sujetos patológicos tienden a mostrar una mayor variación en la magnitud de la fonación y estos parámetros permiten evaluar su cambio bajo para ambos estados de salud.

En el grupo seis se muestran las medidas de Jitter local, rap, ppq5 y ddp. Estos parámetros reflejan la variabilidad en frecuencia de la frecuencia fundamental y presentan cambios evidentes entre los dos estados de salud que pueden ser útiles para la clasificación, ya que se observa una diferencia importante en la concentración de valores de los grupos sanos con respecto a los patológicos como se muestra en la Fig. 37.

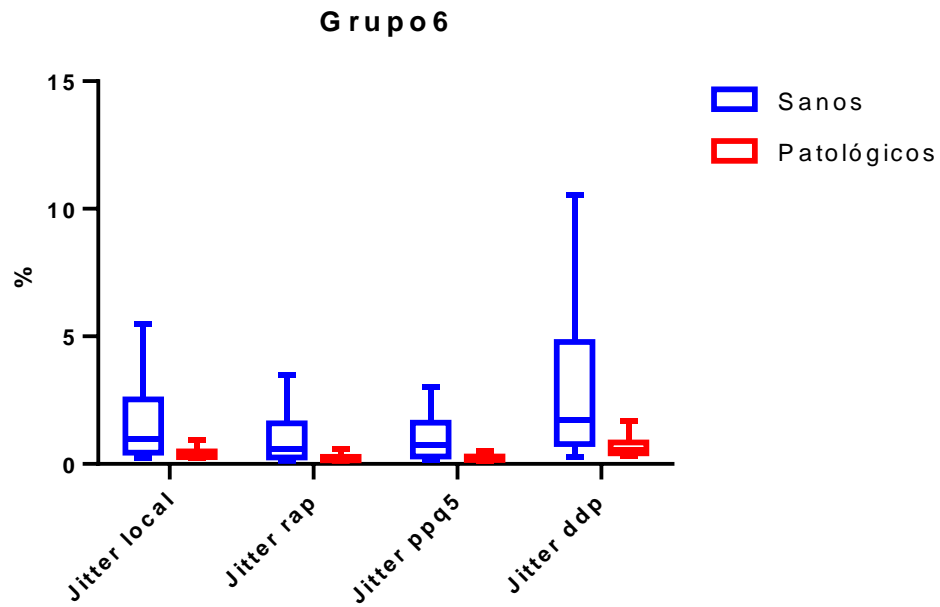


Fig. 37. Distribución de valores obtenidos del análisis acústico del grupo 6.

En el grupo siete se observa la distribución del *Jitter local absolute*. Se distingue que la mayor concentración de los sujetos sanos se mantiene por debajo del límite superior que alcanzan los pacientes con alguna patología como se observa gráficamente en la Fig. 38.

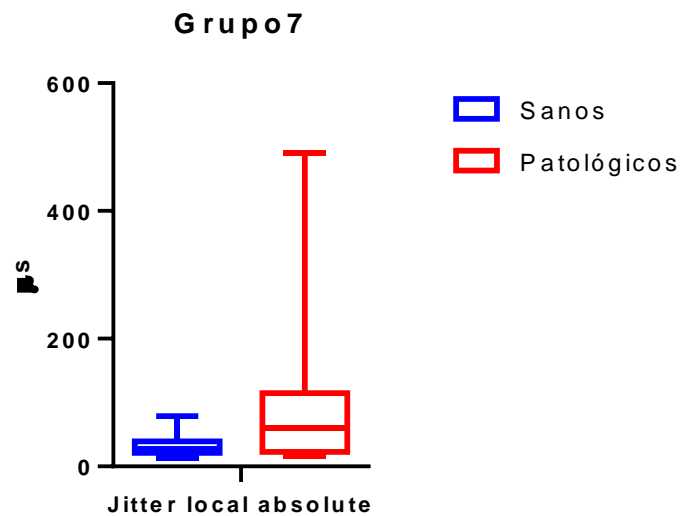


Fig. 38. Distribución de valores obtenidos del análisis acústico del grupo 7.

Sin embargo, no es posible con esta muestra demostrar una diferencia estadísticamente significativa.

Por último, en el grupo ocho donde se tiene el *Shimmer local dB* y que presenta la distribución de la Fig. 39, se observa que el valor máximo obtenido del análisis de las señales de sujetos sanos se posiciona un poco más de 1 dB a diferencia de los pacientes diagnosticados con alguna patología que pueden alcanzar un valor de más de 1.5 dB. Sin embargo, las distribuciones se traslapan los cual nos indica que los grupos no pueden ser linealmente separables.

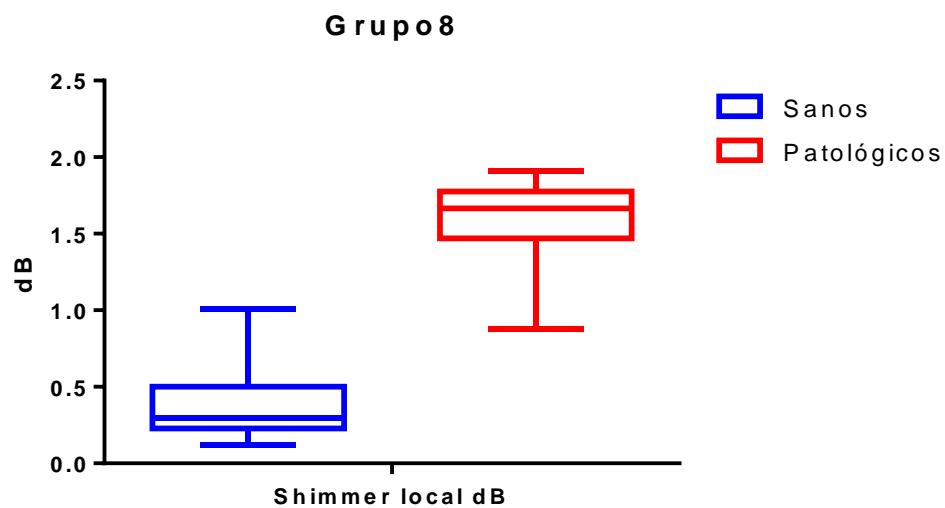


Fig. 39. Distribución de valores obtenidos del análisis acústico del grupo 8.

De acuerdo con las gráficas presentadas anteriormente podemos observar que la mayoría de los parámetros tienen concentraciones diferentes entre grupos. Sin embargo, se realizaron análisis estadísticos de manera individual y conjunta para evaluar la efectividad de los parámetros como elementos de entrada para las neuronas de entrada en el clasificador.

Se realizó la prueba de normalidad Shapiro-Wilk con una $p=0.05$ a cada uno de los parámetros para determinar si la distribución de estos se comportaba de manera normal o difieren de esta distribución.

Tabla 9. Resultados del análisis de normalidad Shapiro-Wilk.

Parámetro	W	P value	Passed normality test (alpha=0.05)?	P value summary
Jitter local	0.6848	< 0.0001	No	****
Jitter local absolute	0.5857	< 0.0001	No	****
Jitter rap	0.6756	< 0.0001	No	****
Jitter ppq5	0.6846	< 0.0001	No	****
Jitter ddp	0.6756	< 0.0001	No	****
Shimmer local	0.8727	0.0005	No	***
Shimmer local dB	0.8603	0.0002	No	***
Shimmer apq3	0.8885	0.0012	No	**
Shimmer apq5	0.9105	0.0051	No	**
Shimmer dda	0.8885	0.0012	No	**
Frecuencia fundamental F0	0.7368	< 0.0001	No	****
DS F0	0.4888	< 0.0001	No	****
CPPS	0.9092	0.0047	No	**
HNR	0.9364	0.0321	No	*
DS HNR	0.9424	0.0502	Yes	Ns

En su mayoría los parámetros no pasaron la prueba de normalidad a excepción de la desviación estándar de la relación-armónico ruido (DS HNR) por lo que se determinó que se utilizarían análisis no paramétricos para evaluar la diferencia entre grupos.

Con base en los resultados del análisis de normalidad y el número de muestras por grupo se seleccionó el análisis *t test* para determinar si existen diferencias significativas entre los estados de cada parámetro. Para esto se seleccionó el análisis *Mann Whitney test* con un valor de $p=0.05$. Los resultados del análisis individual de cada parámetro se muestran en la Tabla 10. Del análisis se obtuvo que en su totalidad los parámetros tenían una diferencia significativa, sin embargo, se encontraron algunos parámetros que tienen un valor bajo de p por lo que se esperaba un buen desempeño

como parámetro identificador del estado de salud, utilizando la red neuronal como evaluador de todos ellos

Tabla 10. Resultados del análisis *t test* de manera individual de los parámetros.

Parámetro	P value	P value summa ry	Significantly different? (p < 0.05)	95.06% ci of difference
CPPS	< 0.0001	****	Yes	-10.83 to -7.797
HNR	< 0.0001	****	Yes	-14.86 to -10.22
Shimmer local dB	< 0.0001	****	Yes	0.1213 to 0.1634
Shimmer local	< 0.0001	****	Yes	1.022 to 1.421
Shimmer apq3	< 0.0001	****	Yes	0.05809 to 0.08147
Shimmer apq5	< 0.0001	****	Yes	0.07171 to 0.1013
Shimmer dda	< 0.0001	****	Yes	0.1743 to 0.2444
Jitter ppq5	0.0003	***	Yes	0.001111 to 0.01056
Jitter ddp	0.0002	***	Yes	0.004660 to 0.02912
Jitter rap	0.0002	***	Yes	0.001553 to 0.009706
Jitter local	0.0001	***	Yes	0.003100 to 0.01628
Frecuencia fundamental F0	0.0075	**	Yes	11.00 to 66.03
DS F0	0.0062	**	Yes	0.2545 to 23.59
DS HNR	0.0428	*	Yes	-0.8662 to -0.02684
Jitter local absolute	0.0118	*	Yes	5.334e-006 to 8.478e-005

Como siguiente paso se realizó un análisis general *Two-way ANOVA* para observar el comportamiento de los parámetros en conjunto y evaluar el efecto de la interacción entre ellos. Para este análisis fue necesario realizar una normalización de los datos, ya que requería que todos los parámetros se encontraran en el mismo rango y de esta manera evitar el enmascaramiento de las diferencias entre ellos debido a los valores absolutos de cada uno de ellos. En este análisis se obtuvieron los resultados que

se muestran en la Tabla 11. Con base en los resultados obtenidos se eligieron los parámetros que se incluirían en el diseño de la red neuronal.

Tabla 11. Resultados del análisis *Two-way ANOVA*.

Parámetro	Mean diff.	95% ci of diff.	Significant?	Summary	Adjusted p value
CPPS	-0.4768	-0.6515 to -0.3022	Yes	****	< 0.0001
HNR	-0.499	-0.6737 to -0.3244	Yes	****	< 0.0001
Shimmer local dB	0.6541	0.4795 to 0.8288	Yes	****	< 0.0001
<i>Shimmer local</i>	0.6332	0.4585 to 0.8078	Yes	****	< 0.0001
<i>Shimmer apq3</i>	0.5857	0.4110 to 0.7604	Yes	****	< 0.0001
<i>Shimmer apq5</i>	0.4931	0.3184 to 0.6677	Yes	****	< 0.0001
<i>Shimmer dda</i>	0.5857	0.4110 to 0.7604	Yes	****	< 0.0001
<i>Jitter ppq5</i>	0.2733	0.09863 to 0.4479	Yes	****	< 0.0001
<i>Jitter local</i>	0.2511	0.07646 to 0.4258	Yes	***	0.0004
<i>Jitter ddp</i>	0.2316	0.05693 to 0.4062	Yes	**	0.0016
<i>Jitter rap</i>	0.2316	0.05693 to 0.4062	Yes	**	0.0016
Frecuencia fundamental F0	0.1945	0.01985 to 0.3692	Yes	*	0.0167
DS F0	0.1859	0.01127 to 0.3606	Yes	*	0.0272
DS HNR	-0.1556	-0.3303 to 0.01904	No	ns	0.1273
<i>Jitter local absoluto</i>	0.1728	-0.001834 to 0.3475	No	ns	0.055

Los parámetros de *Jitter local absoluto* y la desviación estándar de la relación armónico ruido fueron descartados como parámetros útiles para la clasificación ya que no presentaba diferencias significativas y valores de p mayores a 0.05.

La reducción de variables es una etapa necesaria e importante en este proyecto debido a que se espera en un futuro usar métodos ligeros, computacionalmente hablando, para que sean implementados en sistemas de procesamiento portátiles y de bajo costo.

5.1.3 Red neuronal

Con base en los resultados obtenidos del análisis estadístico se diseñó una red neuronal para la clasificación de los sujetos participantes. La red neuronal fue desarrollada en el software Matlab con ayuda del *toolbox* implementado en este programa para este propósito.

Con base en el análisis estadístico se diseñó una red neuronal de reconocimiento de patrones con 13 neuronas de entrada las cuales representan cada uno de los parámetros que serán calculados y posteriormente introducidos a la red neuronal. Siguiendo el diagrama propuesto en la entrada se propusieron 10 neuronas internas y una salida binaria.

En la evaluación del diseño de la red utilizando el método de validación cruzada (70%,15%,15%), se obtuvieron los diagramas de confusión en las diferentes etapas del proceso de desarrollo.

En la Fig. 40 se muestra el desempeño de una red neuronal sobreentrenada, manifestando un aumento de la entropía cruzada en la etapa de prueba.

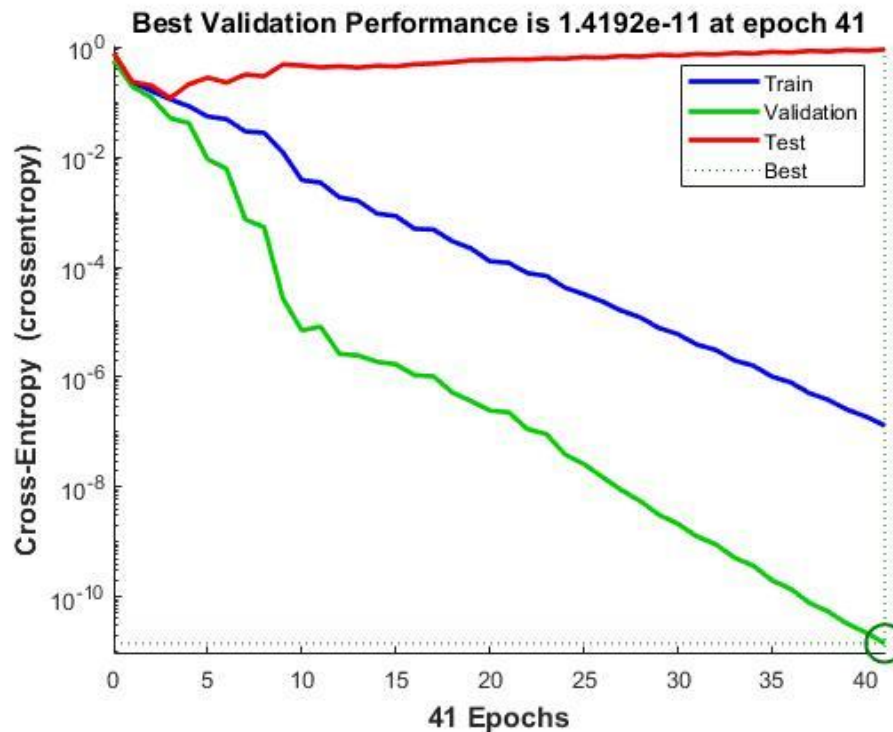


Fig. 40. Desempeño de la red neuronal sobreentrenada.

Se realizaron diversos entrenamientos para seleccionar el que tuviera un mejor desempeño. Se obtuvo la gráfica de desempeño de la red en cada una de las etapas durante cada una de las iteraciones (Fig. 41), en donde la iteración número 31 la que obtuvo un mejor resultado.

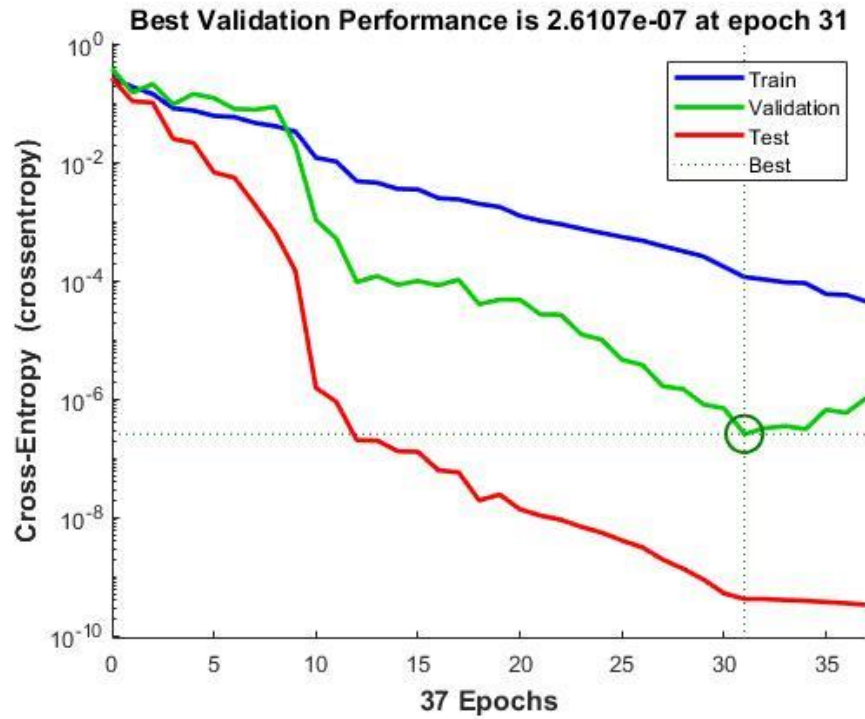


Fig. 41. Desempeño de la red neuronal.

Conforme se realizan las iteraciones del proceso se minimizan los resultados de la entropía cruzada para una buena clasificación. Conforme este resultado se acerca a cero el desempeño del clasificador será mejor. La gráfica demuestra la reducción del valor de entropía cruzada conforme a las etapas del proceso, con lo cual se puede argumentar que no hay un sobreentrenamiento de la red.



Fig. 42. Diagrama de confusión de la red neuronal.

Se puede observar en la (Fig. 42) que la red neuronal tuvo un desempeño favorable en cada una de las etapas, sin presentarse sobreentrenamiento.

5.1.4 Prototipo del dispositivo de obtención de señales

El dispositivo resultante con el circuito impreso en conjunto con el banco de baterías se muestra en la figura Fig. 43.

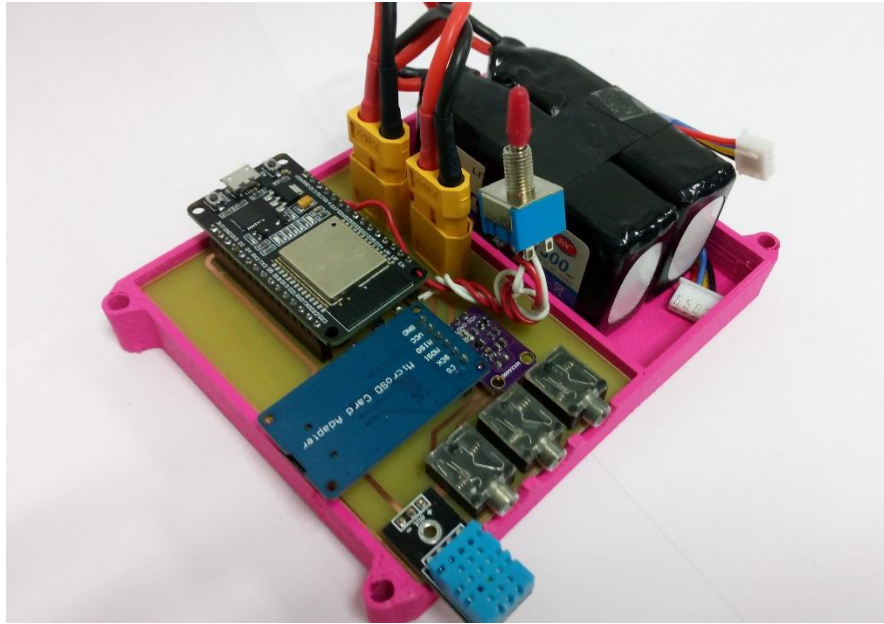


Fig. 43. Montaje de los elementos que integran el sistema de obtención de señales.

Se elaboró un gabinete diseñado en el software Inventor, impreso con filamento ABS en impresora 3D Davinci1.0 xyz.



Fig. 44. Dispositivo resultante con referencia de medida.

El dispositivo es portátil, tiene 5cm de altura, 10cm de ancho y 13.5cm de largo.

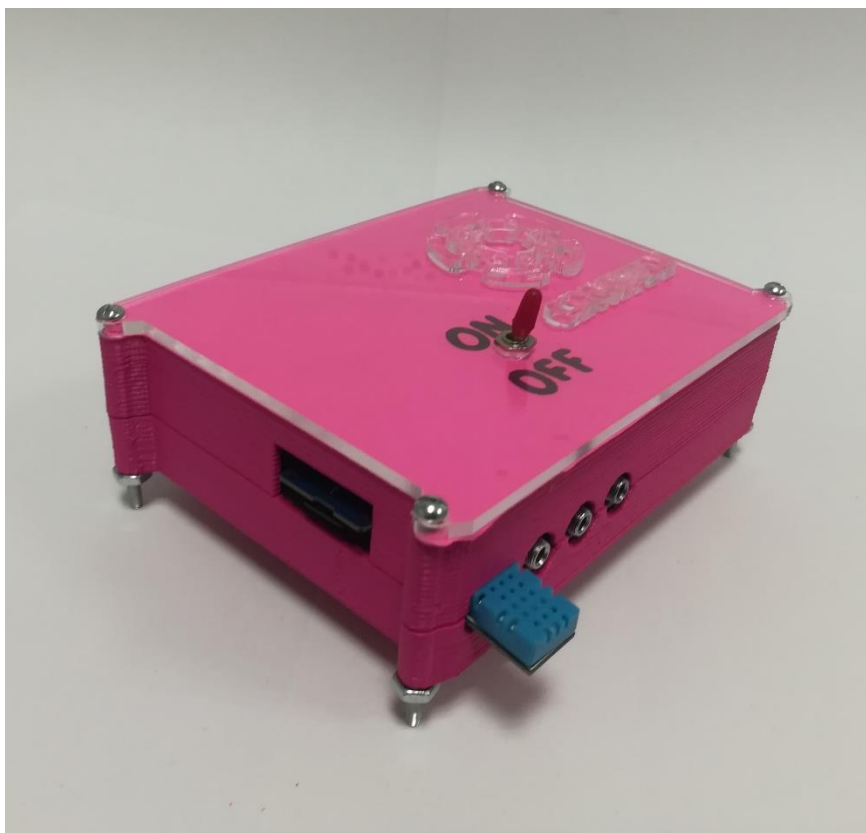


Fig. 45. Prototipo del dispositivo electrónico para registro.

5.1.5 Obtención de señales del circuito electrónico

Se siguió el protocolo de Helsinki mediante la firma de una carta de consentimiento informado para el registro de señales en los sujetos participantes de la prueba (Anexo A.5) y se llenaron los formatos correspondientes de los datos generales y la autoevaluación de incapacidad vocal. Con el sujeto de prueba en posición sedente y con el tronco erecto se colocó el acelerómetro para el registro de las vibraciones, en la parte inferior izquierda del cuello fijado con cinta adhesiva Microporo®. Se le pidió al participante colocarse la mascarilla que fue previamente desinfectada y realizar la fonación de una “a” sostenida el mayor tiempo posible a un volumen cómodo para el sujeto (Fig. 46).



Fig. 46. Toma de registro del sujeto de prueba.

Flujo de aire

Como prueba de funcionamiento del flujómetro de aire, se le pidió al participante generar tres soplos cortos con la mascarilla puesta y dirigidos al sensor. La señal obtenida se muestra en la Fig. 47 la cual muestra que el dispositivo responde a flujos de aire comparables con los generados por las fonaciones establecidas en los protocolos de valoración de la voz. Debido a que el sistema desarrollado ya cuenta con el medidor de flujo de aire, los valores de este parámetro podrán incluirse como trabajo a futuro de esta tesis.

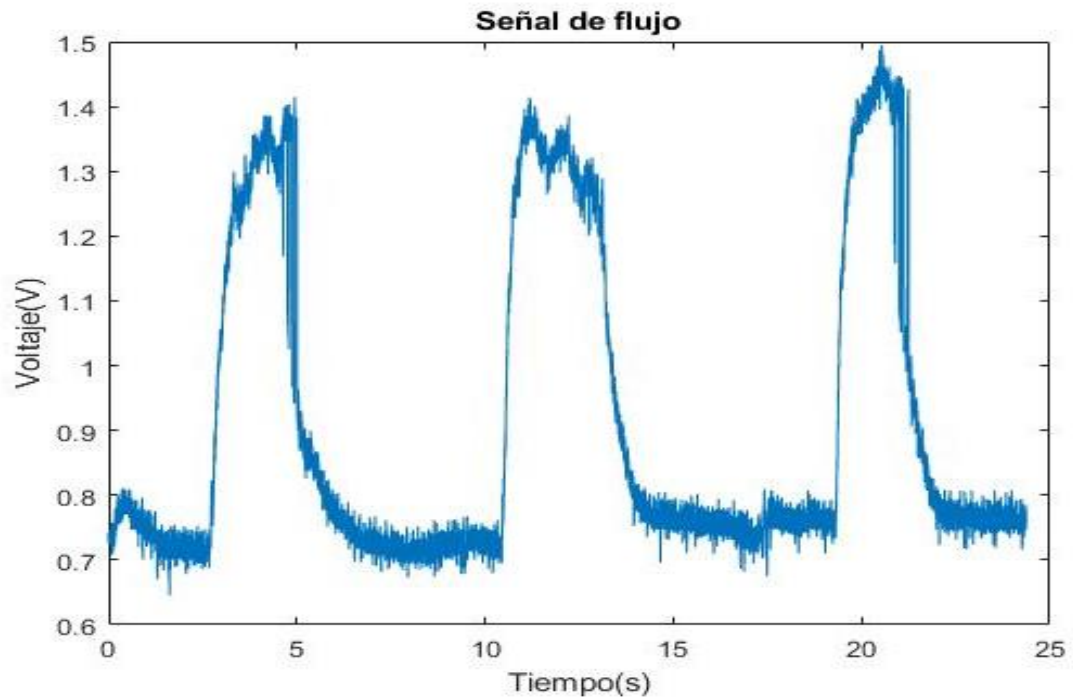


Fig. 47. Respuesta del sensor de flujo a tres impulsos de aire.

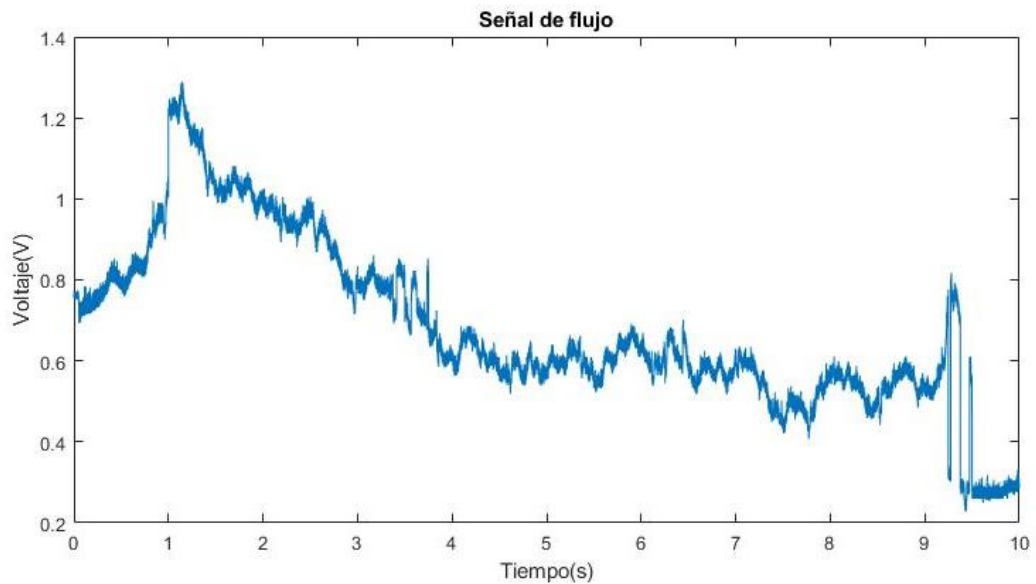


Fig. 48. Respuesta del sensor de flujo durante una vocalización sostenida de una "a".

Audio

Del circuito propuesto y el sensor montado en la mascarilla se obtuvieron las señales mostradas en la Fig. 49 y la Fig. 50, realizando una vocalización de la letra "a".

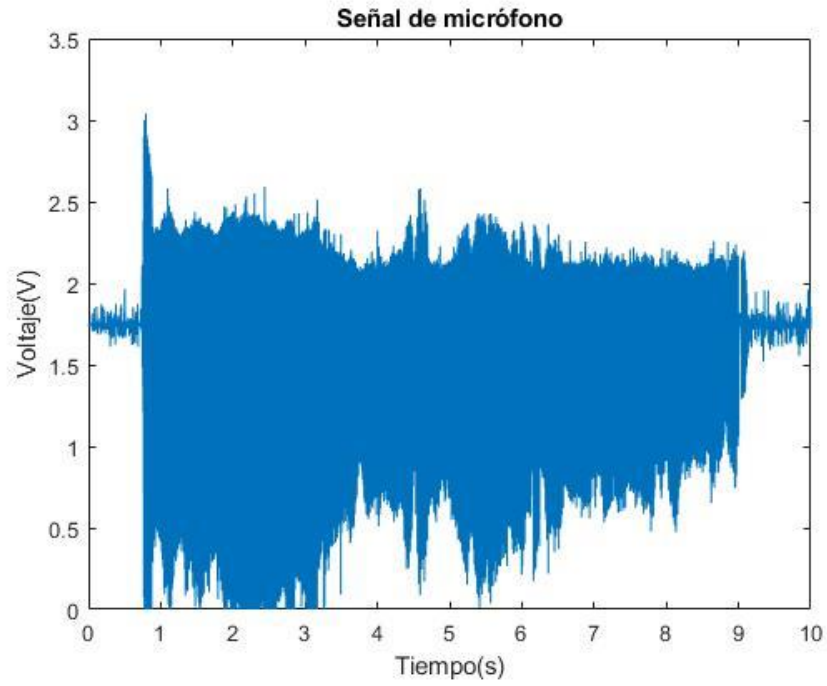


Fig. 49. Señal de micrófono durante una vocalización de la letra “a” del sujeto de prueba.

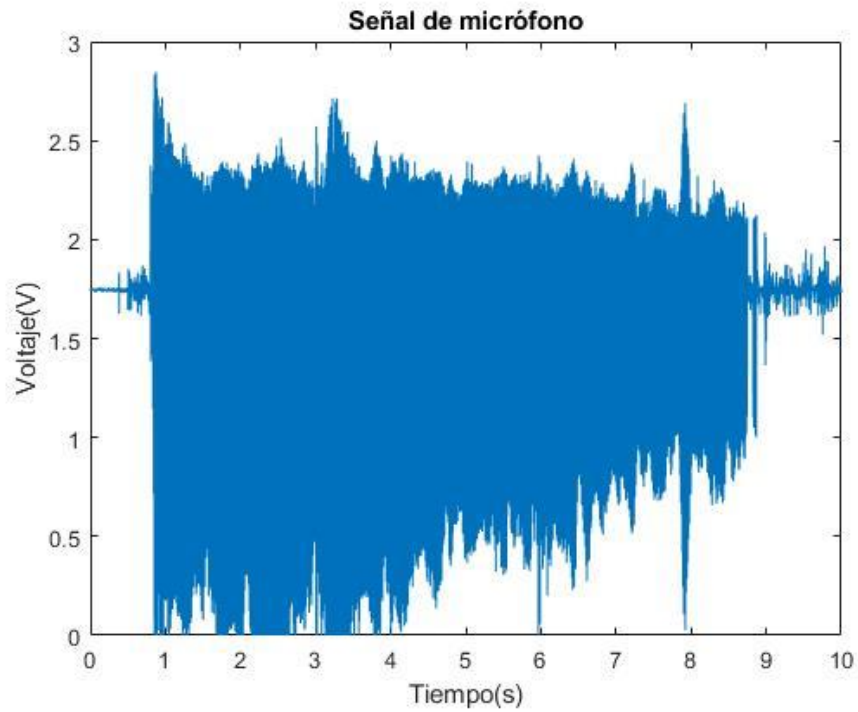


Fig. 50. Repetición de vocalización del sujeto de prueba.

Vibraciones

Del circuito propuesto y el sensor colocado en la superficie del cuello se obtuvieron las señales mostradas en la Fig. 51 y Fig. 52.

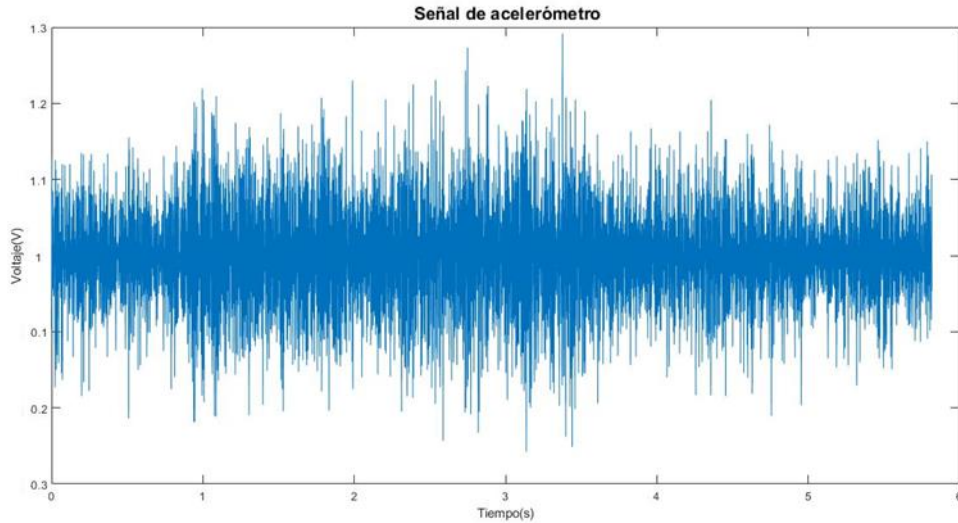


Fig. 51. Señal del acelerómetro en vocalización de una "a" de sujeto de pruebas.

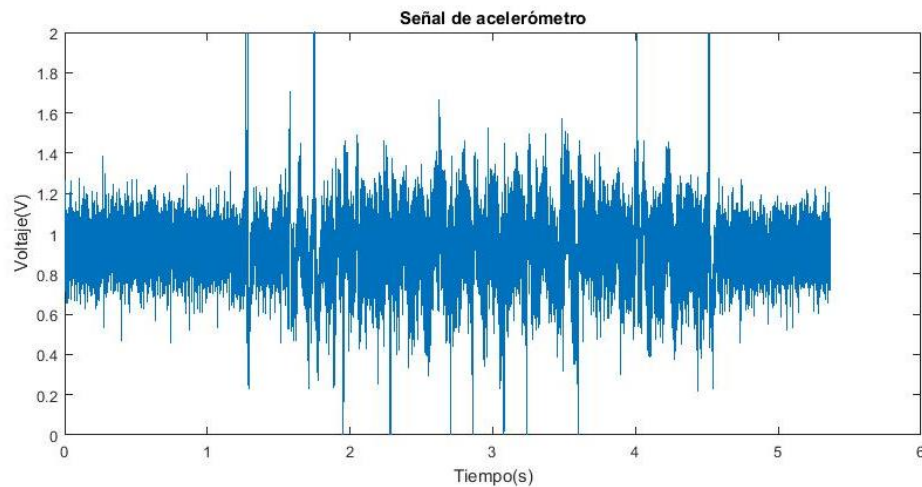


Fig. 52. Señal del acelerómetro durante una vocalización de una "a" del sujeto de pruebas.

En las señales mostradas en la Fig. 51 y Fig. 52 se logra apreciar una actividad en ambas señales sin embargo se propone el uso de filtros en un trabajo a futuro para la reducción de ruido, así como una mayor sujeción a la superficie del cuello.

Temperatura y humedad

Con el sensor de temperatura y humedad digital se realizaron mediciones durante 24 horas para observar su respuesta. Se compararon los valores obtenidos del sensor DHT11 con un medidor analógico de la marca Quartz®. Ambas gráficas se muestran en la Fig. 53. Como se observa las gráficas se comportan de manera similar.

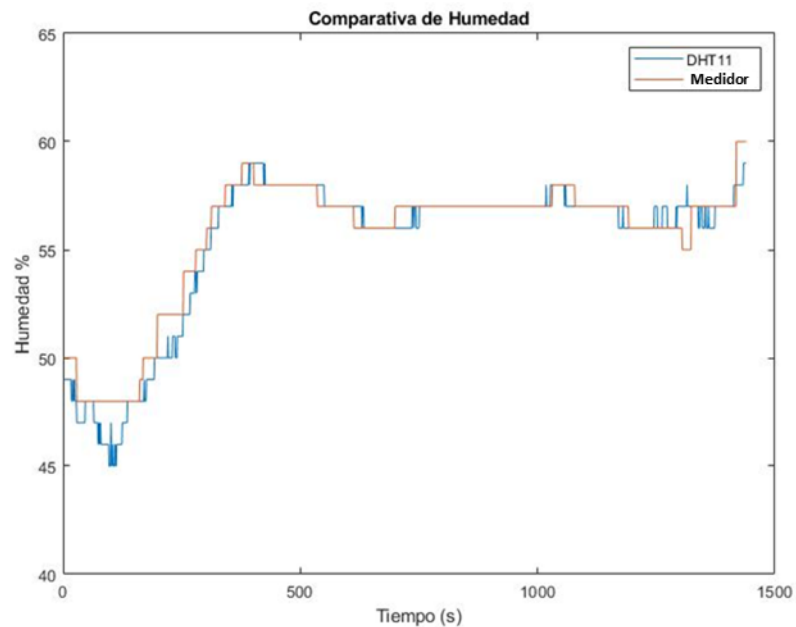


Fig. 53. Respuesta en tiempo de la humedad con sensor DHT11.

Respecto a la temperatura, se realizaron mediciones con un multímetro modelo UT61C de la marca UNI-T® el cual cuenta con un sensor integrado de temperatura y se compararon con las mediciones del sensor DHT11. Los resultados se muestran en las gráficas de la Fig. 54.

Se tuvo una temperatura promedio de 23 con un máximo de 25 a las 2:30 p.m. y una temperatura mínima de 21 a las 7:30 a.m. Esto demuestra que es importante tener esta temperatura como referencia durante la sesión de prueba.

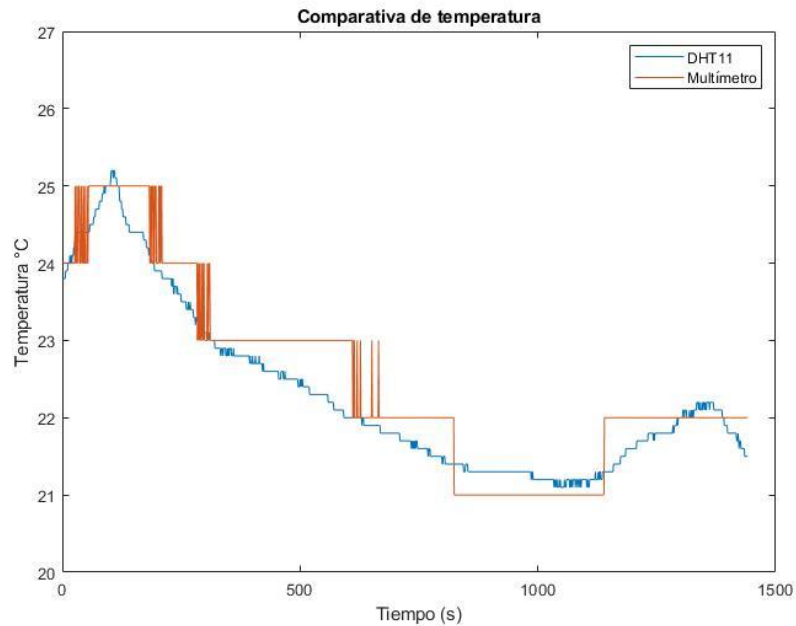


Fig. 54. Respuesta en el tiempo de la temperatura con el sensor DHT11 y con el termómetro comercial.

La respuesta del sensor en la medición de temperatura es comparable a la medida arrojada por el multímetro. La diferencia máxima de los valores fue de un grado debido a que el multímetro arroja la medida redondeada en enteros mientras que el valor obtenido del sensor DHT11 se muestra con 2 dígitos decimales

Esta medición servirá de referencia para la futura implementación de protocolos en campo donde no siempre se tiene controladas las condiciones de registro y permitirá investigar su efecto en los parámetros en general.

5.1.6 Interfaz gráfica

La interfaz gráfica creada en Matlab cumplió la función de desplegar los resultados obtenidos de forma gráfica e ilustrativa de manera sencilla para el usuario. Como una primera versión, el diseño cumple con las funciones básicas para el despliegue de las señales adquiridas. Algunas de las propuestas de mejora para un trabajo a futuro son que el manejo de la interfaz se personalice de acuerdo con el rol de los usuarios, asignando contraseñas para el acceso de los usuarios y con esto la creación de perfiles para complementar el sistema.

Guía de uso de la interfaz gráfica

Esta primera versión de la interfaz presenta una ventana inicial que despliega el menú principal mostrado en la Fig. 55 con tres botones. El botón para crear el nuevo expediente cierra la venta principal y abre una nueva ventana para el ingreso de datos del paciente con los campos que muestra la Fig. 56.



Fig. 55. Menú de inicio de la interfaz gráfica.

En la ventana de datos generales se llenan los campos de nombre, edad, ocupación y sexo del participante, este proceso se podrá cancelar y regresar al menú principal si así se requiere o se proseguirá con el llenado del cuestionario VHI. Una vez llenado se generará un archivo de texto con extensión .txt que contendrá la información recopilada dentro de una carpeta generada con el nombre ingresado por el participante, en dicha carpeta se guardarán todos los documentos obtenidos durante el proceso.

Datos

Datos generales

Nombre

Edad

Ocupación

Sexo

-- Hombre Mujer

Cancelar Siguiente

Fig. 56. Ventana para ingreso de datos de nuevo expediente.

Una vez obtenidos los datos generales, el participante llena la autoevaluación con el cuestionario VHI que se despliega en la Fig. 57. Durante esta etapa se genera un archivo de texto con extensión .txt dentro de la carpeta personal del participante que contiene las respuestas.

VHI

Formulario

La gente me oye con dificultad debido a mi voz
 -- Nunca Casi nunca A veces Casi siempre Siempre

La gente no me entiende en lugares ruidosos
 -- Nunca Casi nunca A veces Casi siempre Siempre

Mis problemas con la voz alteran mi vida personal y social
 -- Nunca Casi nunca A veces Casi siempre Siempre

Me siento desplazado de las conversaciones por mi voz
 -- Nunca Casi nunca A veces Casi siempre Siempre

Mi problema con la voz afecta al rendimiento laboral
 -- Nunca Casi nunca A veces Casi siempre Siempre

Siento que necesito tensar la garganta para producir la voz
 -- Nunca Casi nunca A veces Casi siempre Siempre

La calidad de mi voz es impredecible
 -- Nunca Casi nunca A veces Casi siempre Siempre

Mi voz es molesta
 -- Nunca Casi nunca A veces Casi siempre Siempre

Mi voz me hace sentir cierta minusvalía
 -- Nunca Casi nunca A veces Casi siempre Siempre

La gente me pregunta: ¿Qué te pasa con la voz?
 -- Nunca Casi nunca A veces Casi siempre Siempre

Siguiente

Fig. 57. Cuestionario VHI.

Una vez finalizada la encuesta VHI continúa otra ventana para el registro de señales en la que se selecciona el nombre del paciente (Fig. 58). En caso de no aparecer el paciente o participante bajo estudio es necesario regresar a registrarlos previamente.

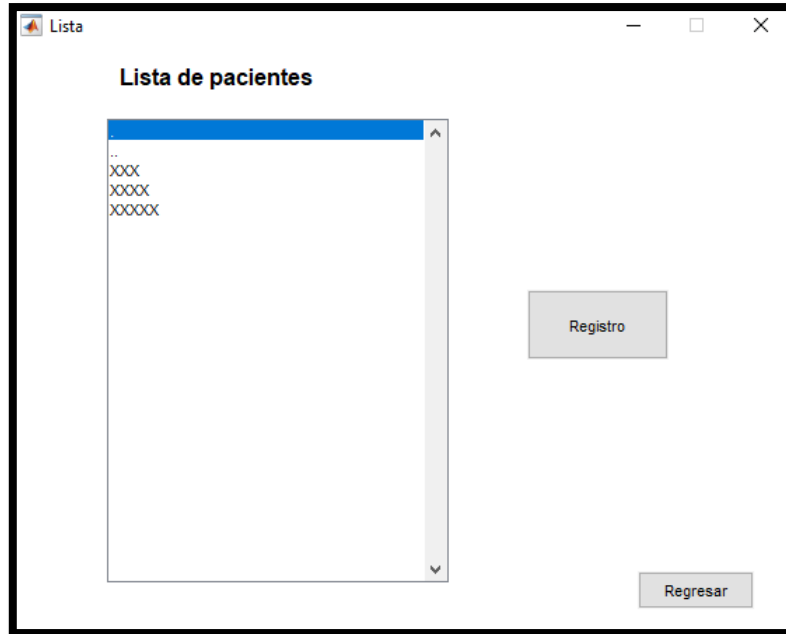


Fig. 58. Ventana para listado de pacientes con expediente.

Una vez seleccionado se presiona el botón registro lo cual abrirá una ventana con la cual se realizará la toma de señales (Fig. 59), se coloca al paciente sentado de manera erguida ya que la espalda debe estar recta para evitar alteración de la postura [39], para registro y se coloca el acelerómetro en la parte inferior izquierda y se le solicita colocar la mascarilla con los sensores montados. Se le solicita al sujeto una vocalización de una “a” sostenida a un volumen cómodo el mayor tiempo posible, siguiendo el protocolo “Estandarización del análisis acústico de la voz en un sector de la población mexicana” [39]. Cuando el sujeto esté listo para iniciar la vocalización, se presiona el botón “Iniciar” de la ventana registro. Cuando este concluya la vocalización se accionará el botón “Finalizar”, si por algún motivo el sujeto no realiza el ejercicio adecuadamente, la toma de las señales puede reiniciarse el número de veces que sea necesario con el botón “Reiniciar”.

Concluida la muestra se detiene el registro y se procede a retirar la memoria microSD del sistema de adquisición y con el adaptador correspondiente introducir la memoria en el CPU y de esta manera importar el archivo con los datos para ser analizado. Lo siguiente en este proceso es la separación de las diferentes señales como variables independientes y trabajarlas por separado.

A partir de la señal acústica se genera un archivo de audio con extensión .wav para después hacer la detección del inicio y fin de la vocalización. Una vez recortado, se envía a Praat para su análisis. Praat genera un archivo de texto separado por comas con extensión .csv y es recibido por Matlab, el cual contiene los parámetros obtenidos del análisis. Enseguida se realizará la clasificación mediante la red neuronal diseñada.

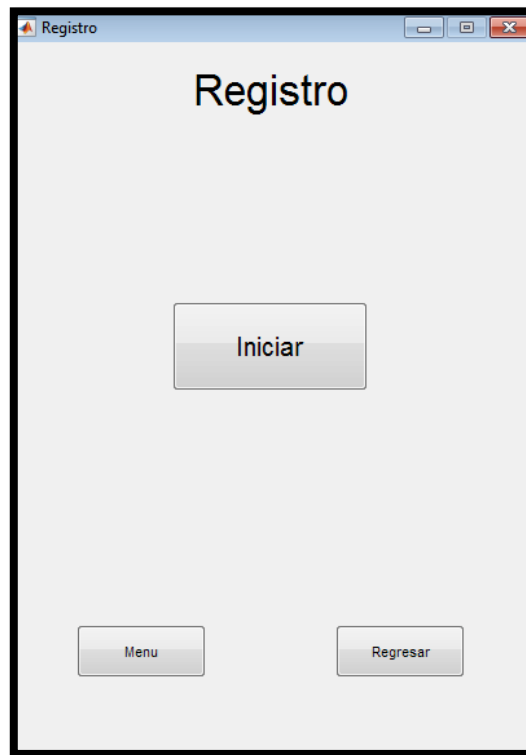


Fig. 59. Ventana de registro.

Una vez importado el archivo se procederá a analizar el archivo de audio creado y se desplegarán las gráficas de las señales obtenidas como se muestra en la Fig. 60. En esta ventana se presentan las gráficas de flujo, audio y vibraciones, así como también las medidas de temperatura y humedad. En la barra de parámetros se puede seleccionar el parámetro que se quiera visualizar. Se despliega el resultado de la clasificación con base a los resultados del análisis. En la parte inferior se encuentran los botones para agregar una nota, si así se requiere, y guardar los resultados, generando un reporte en formato de texto con una extensión .doc. De esta manera se concluye el registro y saldremos del programa al menú principal.

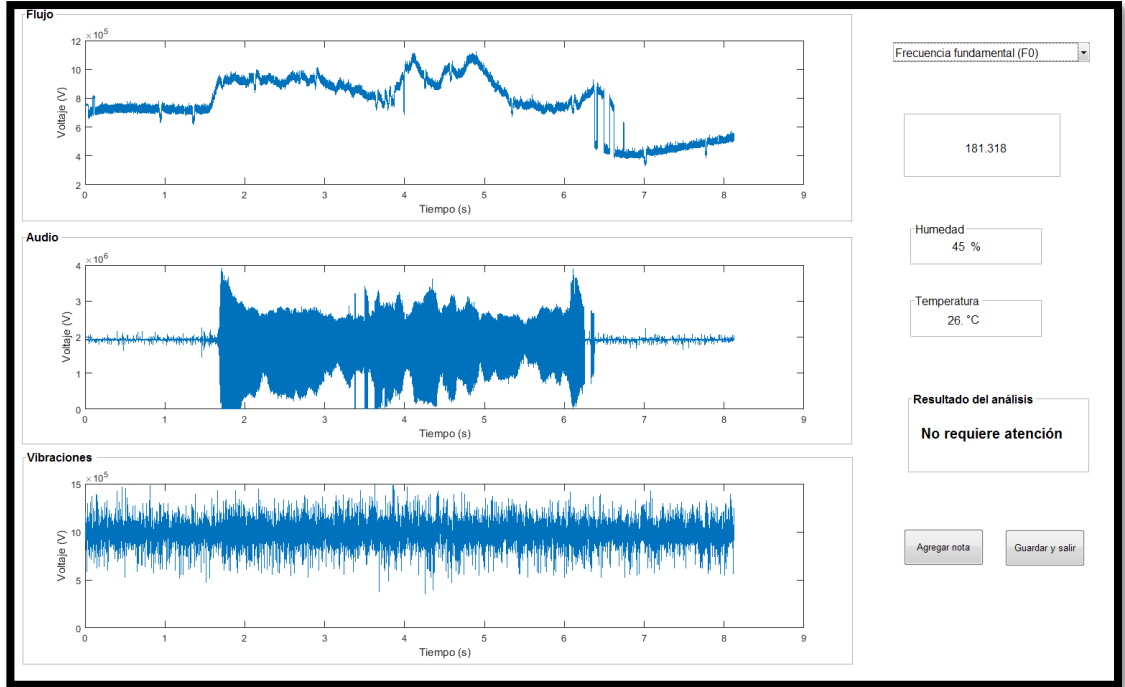


Fig. 60. Ventana para el despliegue de resultados.

5.1.7 Pruebas del sistema

Primero se evaluó el cuestionario de la salud vocal que se resume en la Tabla 12.

Tabla 12. Resultados de la encuesta VHI.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Puntaje total
S1	2	0	0	2	0	0	0	0	0	0	4
S2	2	3	1	1	0	2	2	1	0	1	13
S3	2	0	0	0	0	0	2	0	0	0	4
S4	1	0	0	0	0	0	0	0	0	0	1
S5	0	0	0	0	0	0	0	0	0	0	0
S6	0	2	0	0	0	0	0	0	0	0	2
S7	2	2	2	0	2	0	2	0	0	0	10
S8	0	0	0	0	1	0	1	1	0	0	3

En su totalidad los sujetos no superaron el 32% del puntaje máximo que se puede alcanzar en esta encuesta.

Con fines de observar el comportamiento del sistema se realizaron pruebas en 8 sujetos hombres con 5 repeticiones por sujeto y se realizó una comparativa entre las repeticiones de cada sujeto.

Para evaluar el impacto del cambio en la forma de medición, eliminando las diferencias en el posicionamiento de los sensores dentro protocolo de la medición, se obtuvo la repetibilidad de la toma de los datos con el sistema propuesto para obtener las señales de voz, para ello se obtuvo el promedio y desviación estándar de todos los parámetros considerados de este grupo.

Tabla 13. Desviaciones estándar de los valores obtenidos de cada sujeto.

Parámetro	S1	S2	S3	S4	S5	S6	S7	S8	Promedio
Frecuencia fundamental F0	2.524	1.819	3.024	1.505	2.635	1.863	2.351	3.897	2.452
DS F0	1.153	7.646	1.073	1.533	4.236	0.875	0.630	0.808	2.244
CCPS	0.505	0.883	0.624	0.583	2.669	0.199	1.552	0.534	0.944
HNR	2.015	1.209	1.781	1.636	5.916	1.049	0.838	2.552	2.125
Shimmer local dB	0.073	0.064	0.016	0.027	0.084	0.029	0.079	0.030	0.050
Shimmer local	0.898	0.337	0.198	0.211	0.987	0.298	0.900	0.355	0.523
Shimmer apq3	0.537	0.203	0.146	0.119	0.608	0.136	0.589	0.232	0.321
Shimmer apq5	0.407	0.233	0.134	0.110	0.626	0.231	0.440	0.187	0.296
Shimmer dda	1.612	0.609	0.437	0.356	1.824	0.408	1.767	0.696	0.964
Jitter ddp	0.800	0.236	0.392	0.164	1.303	0.462	0.751	0.299	0.551
Jitter local	0.380	0.107	0.199	0.099	0.679	0.257	0.346	0.175	0.280
Jitter rap	0.267	0.079	0.131	0.055	0.434	0.154	0.250	0.100	0.184
Jitter ppq5	0.138	0.054	0.093	0.073	0.360	0.170	0.140	0.126	0.144

El valor de desviación estándar obtenido en cada uno de los parámetros es considerablemente bajo, por lo que podemos argumentar que el sistema tuvo buena repetibilidad entre los sujetos.

5.1.8 Evaluación de parámetros

Finalmente se realizó una comparación de la desviación estándar y promedio de la muestra obtenida los resultados de este grupo de sanos con respecto a los otros dos grupos de estudio (base de datos previa), para contrastar el comportamiento de los parámetros, sin perder de vista que fueron tomados bajo condiciones diferentes de registro, así como para conocer si el grupo de prueba se encontraba dentro del rango de alguno de los grupos de referencia. Los resultados se muestran en la Tabla 14

Tabla 14. Comparativa de promedios y desviación estándar entre los grupos de estudio.

Parámetros	Patológicos		Sanos		Prueba	
	Promedio	σ	Promedio	σ	Promedio	Σ
F0	183.74	77.414	127.15	14.192	124.53	20.338
F0 DS	26.39	40.55	1.526	0.827	3.6389	1.570
CPPS	8.529	2.721	15.98	2.001	16.255	2.646
HNR	6.415	3.654	18.895	3.397	25.813	2.765
Shimmer local dB	1.5672	0.296	0.397	0.260	0.123	0.023
Shimmer local	18.010	3.832	4.381	2.755	1.330	0.279
Shimmer apq3	9.020	2.080	2.337	1.460	0.733	0.1742
Shimmer apq5	11.511	2.970	2.811	1.889	0.814	0.181
Shimmer dda	27.062	6.241	7.012	4.380	2.200	0.522
Jitter ddp	3.030	2.819	0.66	0.354	1.637	0.560
Jitter local	1.713	1.506	0.395	0.183	0.955	0.286
Jitter rap	1.01	0.94	0.22	0.118	0.545	0.186
Jitter ppq5	1.023	0.889	0.236	0.1	0.551	0.162

El grupo de prueba se asimilaba al grupo de referencia de sujetos sanos en las siguientes gráficas se observa la dispersión de promedios con el SEM (Standard error of the mean) el cual representa la dispersión que tendría la media de una muestra de valores si se continuaran tomando muestras.

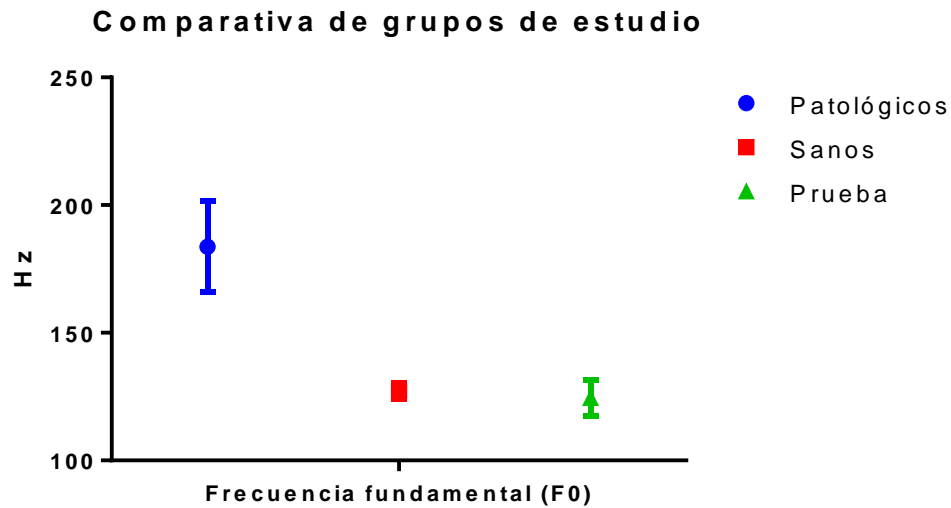


Fig. 61. Comparativa entre grupos de estudio con base a la frecuencia fundamental (F0).

En la gráfica de la Fig. 61 se observa que los valores obtenidos de frecuencia fundamental del grupo de prueba se asemejan en mayor medida con el grupo de referencia de sujetos sanos.

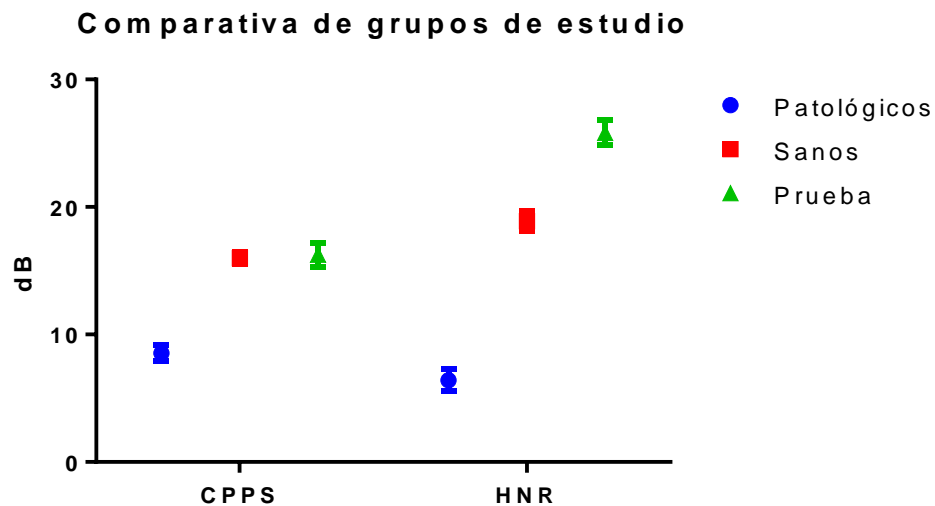


Fig. 62. Comparativa entre grupos de estudio con base a los parámetros de CPPS y HNR.

Con base a la gráfica de comparativa de la Fig. 62 en el parámetro de CPPS los resultados del grupo de pruebas está dentro de los resultados de los sujetos sanos siendo un parámetro que no se ve afectado por las condiciones de medición de las señales de audio por otro lado la relación armónico-ruido (HNR) si mostró una diferencia significativa entre los tres grupos de estudio siendo el grupo de los sujetos de prueba el

que presento un HNR más elevado lo cual significa que hay una mejora en la captación de la señal de audio.

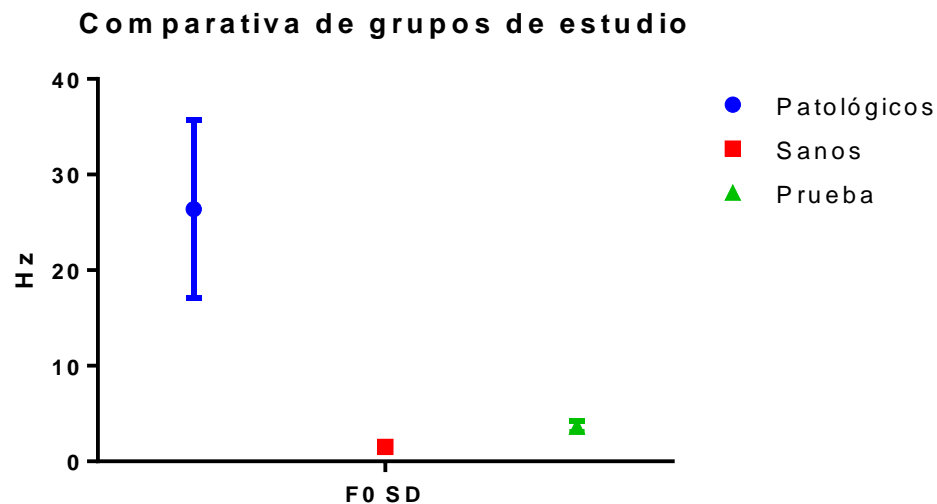


Fig. 63. Comparativa entre grupos de estudio con base a la desviación estándar de F0.

Referente a la gráfica de comparativa de la Fig. 63 la desviación obtenida de las mediciones de frecuencia fundamental (F0) fue menor en el grupo de personas sanas sin embargo el grupo de prueba no alcanza los valores mínimos del grupo de sujetos patológicos.

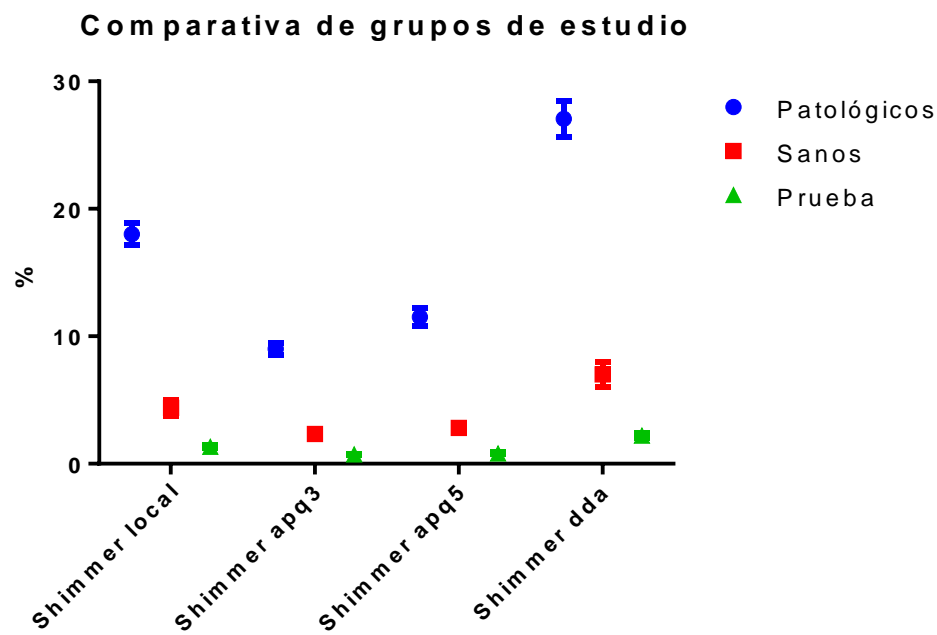


Fig. 64. Comparativa entre grupos de estudio con base a los parámetros de *Shimmer*.

De acuerdo con los resultados mostrados en la Fig. 64 hubo un mejor resultado en las variaciones de amplitud de las señales de audio lo cual se podría argumentar que es debido al posicionamiento del micrófono dentro de la mascarilla.

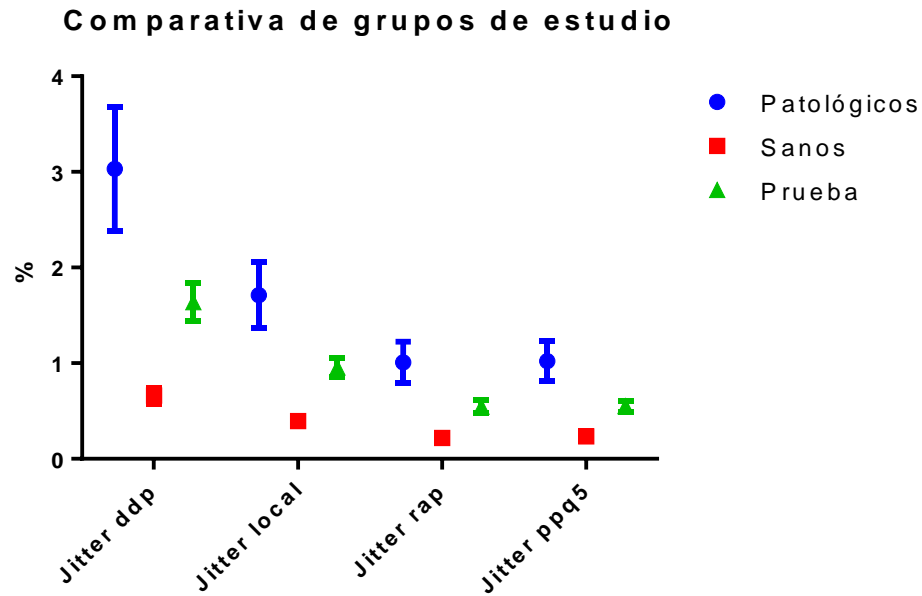


Fig. 65. Comparativa entre grupos de estudio con base a los parámetros de *Jitter*.

Los resultados obtenidos por los parámetros que miden las variaciones en frecuencia del grupo de prueba resultaron mayores que el grupo de referencia de sujetos sanos; no obstante, se mantuvieron debajo de la distribución del grupo de referencia con patologías. Este grupo de parámetros requerirá de más estudios para poder analizar el efecto de la varianza con respecto al grupo de referencia.



Fig. 66. Comparativa entre grupos de estudio con base al *Shimmer local dB*.

Por último, el parámetro de *Shimmer* en dB tuvo un resultado menor que ambos grupos de referencia.

De manera general se observó una mejora en los parámetros de *Shimmer* que miden los cambios en amplitud y el parámetro de la relación armónico-ruido (HNR). Los parámetros de frecuencia fundamental (F0), desviación estándar de F0 y CPPS se comportaron de manera similar al grupo de prueba de los sujetos sanos. Sin embargo, los parámetros de *Jitter* se vieron afectados por lo que se deberán estudiar más a fondo para poder determinar la causa.

Con base a la comparativa de los grupos de estudio se esperaba que el resultado de la clasificación arrojara que los sujetos del grupo de prueba son sanos y no requieren atención.

Tabla 15. Resultados de la clasificación de sujetos de prueba.

Sujeto	Resultado de la clasificación
1	No requiere atención
2	No requiere atención
3	No requiere atención
4	No requiere atención
5	No requiere atención
6	No requiere atención
7	No requiere atención
8	No requiere atención

De manera general se observó que los parámetros seleccionados para el análisis son efectivos para la evaluación objetiva de la voz y una clasificación apropiada.

Capítulo 6

6.1 Conclusiones

Se desarrollo un sistema portátil y no invasivo para la medición de flujo de aire, audio y vibración en la superficie del cuello, durante una fonación, para detectar tempranamente anomalías en la producción de la voz.

Se realizó el análisis acústico de la base de datos de vocalizaciones de sujetos mexicanos proporcionada por el INR con ayuda del software para análisis de audio Praat. Mediante un análisis estadístico de los resultados obtenidos realizado en Prisma se seleccionaron los parámetros acústicos que mostraban una diferencia significativa entre el grupo de personas sanas y los pacientes con alguna patología, ya que estos serían de mayor utilidad para la clasificación.

La clasificación fue llevada a cabo mediante el desarrollo de una red neuronal para reconocimiento de patrones con la cual se puede obtener la clasificación de pacientes sanos o patológicos.

Se seleccionó la variable de flujo de aire ya que se ha reportado como una variable que brinda información acerca de la periodicidad y estabilidad de la voz [31]. La incorporación de la medición de este parámetro permitirá hacer estudios más detallados de la producción de la voz. Para la medición de esta variable se seleccionó un sensor de flujo el cual utiliza la compensación térmica como principio de funcionamiento, lo cual hizo posible la detección del flujo de aire producido por vocalizaciones sostenidas.

Se optó por el uso de un acelerómetro para la medición de las vibraciones en la superficie del cuello por la naturaleza de las estructuras y de los procesos involucrados

en la producción de la voz, así como reducir interferencias por el ruido ambiental. La incorporación de la medición de este parámetro en el sistema desarrollado en esta tesis permitirá hacer estudios más detallados de la producción de la voz.

En la evaluación de la producción de la voz, uno de los parámetros más utilizados es la medición de la acústica de los sonidos producidos establecidos en protocolos. El sistema desarrollado permite la medición, análisis y clasificación de vocalizaciones sostenida. Como salida da información del estado de salud de la voz.

La interfaz gráfica creada para el registro y toma de muestras fue de fácil manejo por parte del usuario y del paciente. Se incluyeron cuestionarios y ventanas para la presentación de los resultados gráficos y numéricos del análisis del sujeto.

Los sensores fueron colocados de manera que no fuese invasiva y no resultara molesto para el sujeto. Estos fueron integrados en un dispositivo portátil para su registro en conjunto con una interfaz gráfica para su procesamiento.

La modificación de la manera de colocar los sensores tuvo un efecto positivo en los resultados de algunos parámetros. Como siguiente etapa se requerirá realizar estudios con grupos de prueba patológicos y analizar en conjunto con las demás señales obtenidas.

Finalmente, el sistema desarrollado tuvo una respuesta positiva por lo que podrá ser utilizado como una herramienta de primer frente y para trabajo de campo con el propósito de apoyar a los especialistas en el diagnóstico oportuno de anomalías en la producción de la voz.

6.2 Trabajo a futuro

Para mejorar el sistema, se proponen acciones que a continuación se enlistan como perspectiva a futuro:

a) La elaboración de una base de datos extensa con la inclusión de pacientes femeninos en el estudio en la cual se plantea la inclusión del análisis de las señales de flujo y vibraciones para una clasificación más completa.

b) La creación de perfiles de manera jerárquica para la seguridad de datos del paciente. De este modo se delimitará el acceso a la información clínica dependiendo el nivel de acceso del usuario al sistema.

c) Una optimización de la parte electrónica reduciendo consumos de corriente y voltaje, así como el tamaño del sistema electrónico, de esta forma se podrá lograr un mayor tiempo de trabajo del banco de baterías.

d) Implementar un expediente clínico que permita el seguimiento de los casos.

Referencias

- [1] N. G. Elisei, “Análisis acústico de la voz normal y patológica utilizando dos sistemas diferentes: ANAGRAF Y PRAAT,” 2012.
- [2] B. Palomino and G. López, “Las enfermedades de la voz,” *Calid. vida, salud y medio Ambient.*, pp. 207–219, 2000.
- [3] “Organización Internacional del Trabajo.” [Online]. Available: <https://www.ilo.org/global/lang--es/index.htm>.
- [4] M. Brockmann, M. J. Drinnan, C. Storck, and P. N. Carding, “Reliable jitter and shimmer measurements in voice clinics: The relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task,” *J. Voice*, vol. 25, no. 1, pp. 44–53, 2011.
- [5] R. Dosal González, “Producción de la voz y el habla,” p. 27, 2014.
- [6] U. D. C. Rica, U. D. C. Rica, M. De Oca, and C. Rica, “Anatomía y fisiología de los mecanismos del habla,” 2006.
- [7] J. Bruce, *Review: An Introduction to Neural Networks*, vol. 40, no. 4. 1998.
- [8] A. R. L. Francisco, “La Computación en México por especialidades académicas,” *J. Chem. Inf. Model.*, vol. 53, no. 9, 2013.
- [9] J. Delgado-Hernández, N. León-Gómez, and A. Jiménez-Álvarez, “Diagnostic accuracy of the Smoothed Cepstral Peak Prominence (CPPS) in the detection of dysphonia in the Spanish language,” *Loquens*, vol. 6, no. 1, p. 058, 2019.
- [10] F. Núñez Batalla, R. González Márquez, M. B. Peláez González, I. González Laborda, M. Fernández Fernández, and M. Morato Galán, “Análisis acústico de la voz mediante el programa Praat: Estudio comparativo con el programa Dr. Speech,” *Acta Otorrinolaringol. Esp.*, vol. 65, no. 3, pp. 170–176, 2014.

- [11] J. Delgado-Hernández, N. M. León-Gómez, L. M. Izquierdo-Arteaga, and Y. Llanos-Fumero, “Análisis cepstral de la voz normal y patológica en adultos españoles. Medida de la prominencia del pico cepstral suavizado en vocales sostenidas versus habla conectada,” *Acta Otorrinolaringológica Española*, vol. 69, no. 3, pp. 134–140, 2017.
- [12] D. Öztuna, A. H. Elhan, and E. Tüccar, “Investigation of four different normality tests in terms of type 1 error rate and power under different distributions,” *Turkish J. Med. Sci.*, vol. 36, no. 3, pp. 171–176, 2006.
- [13] M. B. W. Samuel Sanford Shapiro, “An analysis of variance test for normality (complete samples),” *Biometrika*, vol. 52, no. 3–4, pp. 591–611.
- [14] C. N. Delmo, J. E. & Refugio, “Empirical Research on Mann-Whitney U-test,” no. December, 2018.
- [15] H. J. Seltman, *Experimental design and analysis*. 2018.
- [16] L. L. González, “El diseño de interfaz gráfica de usuario para publicaciones digitales,” *Rev. Digit. Univ.*, vol. 5, no. 1067–6079, pp. 447–455, 2004.
- [17] C. S. Lent, *Learning to program with Matlab*. .
- [18] M. del S. M. H. Micael Gallego Carrillo, *Interfaces gráficas en Java*. Editorial Universitaria Ramon Areces, 2005.
- [19] S. M. BONGIOVANNI PABLO, CASCINO MARCELO, “Micrófonos,” pp. 1–12, 2011.
- [20] CUI INC®, “MODEL : CMA-4544PF-W | DESCRIPTION : ELECTRET CONDENSER MICROPHONE,” 2013.
- [21] H. y E. A. Traunmüller, *The frequency range of the voice fundamental in the speech of male and female adults*. Stockholm, 1995.
- [22] Innovative Sensor Tecnology, “Application Note. Thermal Mass Flow Sensor.” .
- [23] N. Semiconductors, “Micromachined Accelerometer MMA7361L.” pp. 1–11,

2008.

- [24] S. E. E. Profile, “Evaluación Clínica de la Voz por medio de Mediciones Aerodinámicas y Acústicas,” no. December, 2017.
- [25] M. Nolan, B. Madden, E. Burke, M. Nolan, B. Madden, and E. Burke, “Accelerometer Based Measurement for the Mapping of Neck Surface Vibrations During Vocalized Speech,” 2009.
- [26] M. Fezari, F. Amara, and I. M. M. El-etary, “Acoustic Analysis for Detection of Voice Disorders Using Adaptive Features and Classifiers,” pp. 112–117, 2014.
- [27] J. Delgado, N. León, A. Jiménez, and L. Izquierdo, “Análisis acústico de la voz: medidas temporales, espectrales y cepstrales en la voz normal con el Praat en una muestra de hablantes de español Acoustic analysis of the voice: Temporal, spectral and cepstral measurements in the normal voice with the Praat ,” *Rev. Investig. en Logop.*, vol. 2, no. 2, pp. 108–127, 2017.
- [28] Pentax Medical, “The Phonatory Aerodynamic System (PAS),” 2015.
- [29] L. L. S. Richard I. Zraick, Laura Smith-Olinde, and, “Adult Normative Data for the KayPENTAX Phonatory Aerodynamic System Model 6600,” *J. Voice*, vol. 26, no. 1, pp. 164–176, 2011.
- [30] D. D. Mehta *et al.*, “Using Ambulatory Voice Monitoring to Investigate Common Voice Disorders: Research Update,” *Front. Bioeng. Biotechnol.*, vol. 3, no. October, pp. 1–14, 2015.
- [31] J. R. Orozco-Aroyave *et al.*, “Characterization methods for the detection of multiple voice disorders: Neurological, functional, and laryngeal diseases,” *IEEE J. Biomed. Heal. Informatics*, vol. 19, no. 6, pp. 1820–1828, 2015.
- [32] T. IR, *Workshop on Acoustic Voice Analysis: Summary Statement*. Iowa City, IA: National Center for Voice and Speech, 1995.
- [33] Y. E. Ma E, “Multiparametric evaluation of dysphonic severity,” *J Voice*, 2006.

- [34] D. W. Paul Boersma, “Praat: doing phonetics by computer.” .
- [35] Y. D. Heman-Ackah *et al.*, “Cepstral Peak Prominence: A More Reliable Measure of Dysphonia,” *Ann. Otol. Rhinol. Laryngol.*, vol. 112, no. 4, pp. 324–333, 2014.
- [36] Maxim Integrated Products, “Microphone Preamplifiers with Complete Shutdown,” pp. 1–15, 2001.
- [37] Mouser Electronics, “DHT11 - Humidity and Temperature Sensor,” *Datasheet*, pp. 1–7, 2011.
- [38] C. A. Rosen, A. S. Lee, J. Osborne, T. Zullo, and T. Murry, “Development and Validation of the Voice Handicap Index-10.”
- [39] C. A. A. Jiménez, Víctor Manuel Valadez, Gabriela Pérez Martínez, Laura Ivoone Garay Jiménez, Pablo Rogelio Hernández Rodríguez, Blanca Tovar Corona, Nora Daniela Sánchez Rodríguez, “Estandarización del análisis acústico de la voz en un sector de la población mexicana.” 2018.

Anexos

A.1 *Script para obtención de parámetros en Praat*

```
#Borra el archivo existente con el mismo nombre#
filedelete 'Parametros.csv'
#Construccion de las filas que se llenaran#
header_row$="1"+"", "2"+"", "3"+"", "4"+"", "5"+"", "6"+"", "7"+"", "8"+"", "9"+"", "10"+"", "11"+"", "12"+"", "13"+"", "14"+"", "15"+"", "16"+"", "17"+"", "+" newline$
#Crea el archivo que contendrá los resultados#
header_row$ >Parametros.csv
#Carga el archivo de audio con nombre predeterminado a analizar#
name$ = "P1.wav"
#Separamos el nombre de la extensión del archivo para utilizar el nombre#
Read from file: name$
name$ = name$ - ".wav"

#Mediciones de pitch#
    select Sound 'name$'
    To Pitch... 0.0 75 500
    meanpitch = Get mean... 0 0 Hertz
    meansd = Get standard deviation... 0 0 Hertz
    minpitch = Get minimum... 0 0 Hertz Parabolic
    maxpitch = Get maximum... 0 0 Hertz Parabolic

#Mediciones de jitter & shimmer
    select Sound 'name$'
    To PointProcess (periodic, cc)... 75 500
    meanlocal = Get jitter (local)... 0 0 0.0001 0.02 1.3
    meanlocalabsolute = Get jitter (local, absolute)... 0 0 0.0001
0.02 1.3
    meanrap = Get jitter (rap)... 0 0 0.0001 0.02 1.3
    meanppq5 = Get jitter (ppq5)... 0 0 0.0001 0.02 1.3
    meanddp = Get jitter (ddp)... 0 0 0.0001 0.02 1.3
    select Sound 'name$'
    plus PointProcess 'name$'
    meanlocals = Get shimmer (local)... 0 0 0.0001 0.02 1.3 1.6
    meanlocaldb = Get shimmer (local_dB)... 0 0 0.0001 0.02 1.3 1.6
    meanapq3 = Get shimmer (apq3)... 0 0 0.0001 0.02 1.3 1.6
    meanapq5 = Get shimmer (apq5)... 0 0 0.0001 0.02 1.3 1.6
    #meanapq11 = Get shimmer (apq11)... 0 0 0.0001 0.02 1.3 1.6
    meandda = Get shimmer (dda)... 0 0 0.0001 0.02 1.3 1.6
```

```

#Mediciones de la relación armonico-ruído#
select Sound 'name$'
To Harmonicity (cc)... 0.01 75 0.1 1
meanHNR = Get mean... 0 0
meansdHNR = Get standard deviation... 0 0

#Mediciones del espectro para la obtención de pico cepstral de mayor
prominencia#
select Sound 'name$'
To PowerCepstrogram: 60, 0.002, 5000, 50
To PowerCepstrum (slice): 0.1
Smooth: 0.0005, 1
pscp = Get peak prominence: 60, 333.3, "Parabolic", 0.001, 0,
"Straight", "Robust"

select PowerCepstrogram 'name$'
cpps = Get CPPS: "yes", 0.02, 0.0005, 60, 330, 0.05, "Parabolic",
0.001, 0, "Exponential decay", "Robust"

#Se agregan las variables creadas para los valores obtenidos en el
archivo .csv#
fileappend "'directory$'Parametros.csv" 'meanpitch:10' , 'meansd:10'
, 'cpps:10' , 'meanHNR:10' , 'meansdHNR:10' , 'meanlocaldb:10' ,
'meanlocals:10' , 'meanapq3:10' , 'meanapq5:10' , 'meandda:10' ,
'meanddp:10' , 'meanlocal:10' , 'meanrap:10' , 'meanppq5:10' ,
'meanlocalabsolute:10' , 'minpitch:10' , 'maxpitch:10' 'newline$'

#Seleccionamos los archivos generados en la ventana de objetos y se
eliminan#
select all
Remove

```

A.2 Script de la función de la red neuronal generado en Matlab

```
function [Y,Xf,Af] = myNeuralNetworkFunction(X,~,~)
%MYNEURALNETWORKFUNCTION neural network simulation function.
%
% Generated by Neural Network Toolbox function genFunction, 02-Jul-
2019 13:06:53.
%
% [Y] = myNeuralNetworkFunction(X,~,~) takes these arguments:
%
%   X = 1xTS cell, 1 inputs over TS timesteps
%   Each X{1,ts} = 15xQ matrix, input #1 at timestep ts.
%
% and returns:
%   Y = 1xTS cell of 1 outputs over TS timesteps.
%   Each Y{1,ts} = 2xQ matrix, output #1 at timestep ts.
%
% where Q is number of samples (or series) and TS is the number of
timesteps.
%#ok<*RPMT0>
% ===== NEURAL NETWORK CONSTANTS =====
% Input 1
x1_step1_xoffset =
[0.099128613;0.004706142;0.144973635;0.063862903;0.166428899;0.03972
66;0.013887;0.0072453;0.0082412;0.0217358;0.002943332;0.0022243;0.00
098111;0.0012276;0.0138136];
x1_step1_gain =
[2.30313293803196;2.24357622927889;3.83949957989731;2.39921951853562
;6.4763237552326;3.35399680454662;9.29172868895566;17.5264561856122;
11.3353623631963;5.84215206187073;19.5402818464851;38.1032942203018;
58.6208392393829;69.4955526321093;4.19688247176233];
x1_step1_ymin = -1;
% Layer 1
b1 = [1.617120641184399;0.7775216508604863;-0.7937531910373552;-
0.57295213033219494;-0.063292269437638701;0.27866858534444333;-
0.58623105672140841;0.90510787559763484;-1.0795145543796412;-
1.2198643747165168];
IW1_1 = [-0.57443043760197887 0.10778348687500412
0.19346647309378834 -0.46538416760259133 -0.68143562664723889 -
0.27695773195801338 0.38857437426817198 -0.46140463107479485 -
0.51611247001423621 0.0048316961024381785 0.033375407178340588 -
0.052739397733706429 0.7128925752532056 0.38664560952004001
0.52688972192371653;-0.079984118811014887 0.75827753977991863 -
0.097762541147858803 -1.2614999863794079 -1.0414146009787473
0.34475484098483739 0.11819072311831916 0.048502304836021803 -
0.38824004767525083 0.53087454522134214 -0.14120489595040669 -
0.44370997612430546 0.65399269642563163 -0.40307038793116434
0.65594364753415468;0.23860121010713781 0.3545707406951748
0.59056191364540478 0.60851735428907827 -0.20152632677312668 -
0.49980386162523088 -0.56008930524266087 -0.29417193464628977 -
0.72797073178762817 0.36265722521538923 -0.59902071998063411 -
0.24543175120494179 -0.27919264194415855 0.43690429365458605
0.4447659057341557;-0.044433971131312144 0.44574202493253784
```

```

0.084477099114642809 -0.39007987333682292 -0.23802273723012191 -
0.35821602897426552 0.39438063542935209 0.19147371828187454 -
0.86308033864279465 0.31065263566810869 0.4495098747999926 -
0.33411514564347977 -0.066698533061867724 -0.79264688255620763
0.0089225565280967086;0.18119952143288359 -0.52565631222033926
0.16222521127554956 0.18167480842259148 -0.056048903408794183
0.23640988510932184 0.24528671834390262 -0.95269822387111425 -
0.10321722418576436 -0.48510789308685304 -0.18167430162855971 -
0.94267875109455157 0.20438831517469175 0.1943992869090809 -
0.66256270061323208;0.14238007481692569 -0.58515303145848274 -
0.23811964638400898 -0.37698662249903048 0.23873211275990919
0.4684596467081541 0.32556562324192617 -0.69925252223724488 -
0.43131609273612032 0.61957048334864262 -0.63154569673169769 -
0.59820127357860553 -0.20084117601467988 -0.10417065740522755
0.034886946580020721;-0.65069375618548964 0.58416305488639486
0.02951793953103388 -0.68243403523476542 -0.15743878753105553 -
0.55374980399239393 -0.13065892903656962 -0.41655302358226287
0.036251679915498408 -0.15221867396271307 0.0087017759916012787
0.6137912247018974 0.085542847397049968 -0.22113075476497904
0.66899641060350823;0.61206396656956552 0.19256726707086078
0.091554016664707874 0.3778877130961078 0.39600618858042064
0.54948150289271891 0.57781630295268926 -0.37506221561980235 -
0.39347927862753784 0.46003628483256809 0.49831021227463668
0.28786769095598347 0.55146309351761769 0.31224320968795255 -
0.25865703865728923;-0.64222038085423494 -0.094154568652451892 -
0.48798768208389853 0.59398568849467259 1.1306001510786179 -
1.0545651435305505 -0.57131978482095558 -0.53039075948212033
0.19959605867472802 0.33238583366731356 0.0057534726403546052 -
0.55979911076414468 0.092957199900539994 0.40009163324595753 -
0.18480851592742414;-0.13935339164859667 -0.4717714467206458
0.24719652358441552 0.7689864807359702 0.55137033040467287 -
0.89933895084998938 -0.047434574229999374 -0.22635917494144162 -
1.0239764884088052 0.046587996681898801 0.2698215311879576 -
0.6681911966338091 0.56087033910612472 -0.11562342126953543 -
0.42692118846099053];
% Layer 2
b2 = [-0.78450908497306404;-0.14967199098663994];
LW2_1 = [0.34559904807092451 -1.0849760760265998 -
0.26696796588559424 0.40844599852619351 0.39235275545330034 -
0.56806702897350991 0.24115719633269395 -0.13865288653446667
0.66884335558032071 1.2172684953479196;0.50522714702999028
1.3768195192722978 -1.4258041234463499 0.77821018345871684 -
0.56121564942153246 -0.80292140292599623 0.72200045793360534 -
0.27217287725895406 -1.464158027984185 -0.9129149229414355];
% ===== SIMULATION =====
% Format Input Arguments
isCellX = iscell(X);
if ~isCellX, X = {X}; end;
% Dimensions
TS = size(X,2); % timesteps
if ~isempty(X)
    Q = size(X{1},2); % samples/series
else
    Q = 0;
end
% Allocate Outputs
Y = cell(1,TS);

```

```

% Time loop
for ts=1:TS
    % Input 1
    Xp1 =
mapminmax_apply(X{1,ts},x1_step1_gain,x1_step1_xoffset,x1_step1_ymin
);

    % Layer 1
    a1 = tansig_apply(repmat(b1,1,Q) + IW1_1*Xp1);

    % Layer 2
    a2 = softmax_apply(repmat(b2,1,Q) + LW2_1*a1);

    % Output 1
    Y{1,ts} = a2;
end
% Final Delay States
Xf = cell(1,0);
Af = cell(2,0);

% Format Output Arguments
if ~isCellX, Y = cell2mat(Y); end
end
% ===== MODULE FUNCTIONS =====
% Map Minimum and Maximum Input Processing Function
function y =
mapminmax_apply(x,settings_gain,settings_xoffset,settings_ymin)
y = bsxfun(@minus,x,settings_xoffset);
y = bsxfun(@times,y,settings_gain);
y = bsxfun(@plus,y,settings_ymin);
end

% Competitive Soft Transfer Function
function a = softmax_apply(n)
nmax = max(n, [],1);
n = bsxfun(@minus,n,nmax);
numer = exp(n);
denom = sum(numer,1);
denom(denom == 0) = 1;
a = bsxfun(@rdivide,numer,denom);
end

% Sigmoid Symmetric Transfer Function
function a = tansig_apply(n)
a = 2 ./ (1 + exp(-2*n)) - 1;
end

```

A.3 Script para adquisición de datos en tarjeta de desarrollo

ESP32

```
//////////////////////////////////Inclusion de
librerias//////////////////////////////////
#include <Arduino.h>
#include "FS.h"
#include "SD.h"
#include "SPI.h"
//////////////////////////////////Variables para las entradas
analogicas//////////////////////////////////
int sensor[5];
int TB=0;
int CB=0;
uint16_t Bf1[5][5][1000];
//////////////////////////////////Banderas//////////////////////////////////
boolean band[10]={false};
boolean banderareg=false;
boolean banderaini=false;
boolean banderafin=false;
boolean banderaimp=false;
boolean banderares=false;
//////////////////////////////////Variables para la funcion de adquisicion
con timer//////////////////////////////////
volatile int interruptCounter;
int totalInterruptCounter=0;
hw_timer_t * timer = NULL;
portMUX_TYPE timerMux = portMUX_INITIALIZER_UNLOCKED;
//////////////////////////////////Funciones//////////////////////////////////
void IRAM_ATTR onTimer() {
portENTER_CRITICAL_ISR(&timerMux);
interruptCounter++;
  sensor[0]=analogRead(13);
  sensor[1]=analogRead(12);
  sensor[2]=analogRead(14);
  sensor[3]=analogRead(27);
  sensor[4]=analogRead(26);

Bf1[0][CB][TB]=sensor[0];
Bf1[1][CB][TB]=sensor[1];
Bf1[2][CB][TB]=sensor[2];
Bf1[3][CB][TB]=sensor[3];
```

```

Bf1[4][CB][TB]=sensor[4];

TB++;
  if(TB>=1000){
    band[CB]=true;
    TB=0;
    CB++;
  }
  if(CB>=5){
    CB=0;
  }
portEXIT_CRITICAL_ISR(&timerMux);
}

void setup() {
  Serial.begin(115200);
  if(!SD.begin()){
    Serial.println("Card Mount Failed");
    return;
  }
  timer = timerBegin(0, 80, true);
  timerAttachInterrupt(timer, &onTimer, true);
  timerAlarmWrite(timer, 125, true);
  timerAlarmEnable(timer);
}

void loop() {
  Serial.available();
  char dato=Serial.read();

  if(dato=='0'){
    banderareg=true;
  }

  if(banderareg==true){
    //Serial.println("corriendo");
    File file = SD.open("/DATOS.txt", FILE_APPEND);
    if(dato=='1'){
      banderaini=true;
    }
    if(dato=='3'){
      banderares=true;
    }
    if(dato=='4'){
      banderainp=true;
    }
  }
}

```



```

if(banderaini==true) {
Serial.println("guardando");
    if(dato=='2'){
        banderafin=true;
        banderaini=false;
    }
    if(band[0]==true) {
        for(int a=0;a<=4;a++ ){
            for (int b=0;b<=999;b++ ){
                file.print(Bf1[a][0][b]);
                file.print(" ");
            }
        }
        band[0]=false;
    }
    if(band[1]==true) {
        for(int c=0;c<=4;c++ ){
            for (int d=0;d<=999;d++ ){
                file.print(Bf1[c][1][d]);
                file.print(" ");
            }
        }
        band[1]=false;
    }
    if(band[2]==true) {
        for(int e=0;e<=4;e++ ){
            for (int f=0;f<=999;f++ ){
                file.print(Bf1[e][2][f]);
                file.print(" ");
            }
        }
        band[2]=false;
    }
    if(band[3]==true) {
        for(int g=0;g<=4;g++ ){
            for (int h=0;h<=999;h++ ){
                file.print(Bf1[g][3][h]);
                file.print(" ");
            }
        }
        band[3]=false;
    }
    if(band[4]==true) {
        for(int i=0;i<=4;i++ ){
            for (int j=0;j<=999;j++ ){
                file.print(Bf1[i][4][j]);
            }
        }
    }
}

```

```

        file.print(" ");
    }
    }
    band[4]=false;
}

    if(banderafin==true){
        Serial.println("Termino el programa");
        timerEnd(timer);
        file.close();
        banderafin=false;

    }
    if(banderares==true){
        SD.remove("/DATOS.txt");
        banderares=false;
timer = timerBegin(0, 80, true);
timerAttachInterrupt(timer, &onTimer, true);
timerAlarmWrite(timer, 125, true);
timerAlarmEnable(timer);
        banderaini=true;
        File file = SD.open("/DATOS.txt",
FILE_APPEND);
        banderaini=true;
    }
}
}

```

A.4 Función para importar el archivo

```
function Importar_Callback(hObject, eventdata, handles)
global var1
[FileName Path]=uigetfile({'*.txt'}, 'Abrir Documento');
y=horzcat(Path, FileName);
load(y);
x=DATOS;
    T=(length(x));
    F=(T/5)/1000;
ini=1;
inc=1000;
global Sensor1
global Sensor2
global Sensor3
global Sensor4
global Sensor5
for i=1:F
    Sensor1=horzcat(Sensor1,x(:,ini:inc));
    ini=ini+1000;
    inc=inc+1000;
    Sensor2=horzcat(Sensor2,x(:,ini:inc));
    ini=ini+1000;
    inc=inc+1000;
    Sensor3=horzcat(Sensor3,x(:,ini:inc));
    ini=ini+1000;
    inc=inc+1000;
    Sensor4=horzcat(Sensor4,x(:,ini:inc));
    ini=ini+1000;
    inc=inc+1000;
    Sensor5=horzcat(Sensor5,x(:,ini:inc));
    ini=ini+1000;
    inc=inc+1000;
    i=i+1;
end
    F2=(T/5);
Voltaje=(4095/3.3);
Sensor1=Sensor1*Voltaje;
Sensor2=Sensor2*Voltaje;
Sensor4=Sensor4*Voltaje;
Sensor5=Sensor5*Voltaje;
global Tiempo
    Tim = F2/8000;
    Tiempo=linspace(0,Tim,F2);
Sensor3a=int16(Sensor3);
audiowrite('P1.wav',Sensor3a,8000);
[segmentos, fs] = detectVoiced ('P1.wav', 5);

% Programa de llamadas a funciones.
salida = [];
for i = length(segmentos)
salida = [salida; segmentos {i}];
```

```

i=i+1;
end
audio=salida{1,1};
audio=audio';
audiowrite('P1.wav',audio,8000);
system('Praat.exe --run "Parametros"');
load 'parametros.csv';
%Extraer los parametros requeridos
global aud
aud=parametros(2,1:17);
p=parametros(2,1:15);
%%%%%%%%%Operaciones para normalizar%%%%%%%%%
p(1,1)=(p(1,1)-65)/(400-65);
p(1,2)=(p(1,2)-0)/(150-0);
p(1,3)=(p(1,3)-0)/(30-0);
p(1,4)=(p(1,4)-0)/(30-0);
p(1,5)=(p(1,5)-0)/(10-0);
p(1,6)=(p(1,6)-0)/(3-0);
p(1,15)=(p(1,15)-0)/(0.001-0);
p=p';
%%%%%%%%%
global r
Sensor3=Sensor3*Voltaje;
r= myNeuralNetworkFunction(p);
r=r(1,1);
nam=char(var1);
s=char('.wav');
d=char('.csv');
nombreaudio=strcat(nam,s);
nombre analisis=strcat(nam,d);
direc=char('Pacientes\');
folder=strcat(direc,nam);
movefile ('P1.wav',folder);
movefile ('Parametros.csv',folder);
cd(folder)
movefile ('P1.wav',nombreaudio);
movefile ('Parametros.csv',nombre analisis);
global D
global V
global E
global P
D=importdata('Datos.txt');
V=importdata('VHI.txt');
E=importdata('Edad.txt');
P=importdata(nombre analisis);
cd ..
cd ..
global et
et=importdata('Etiquetas.txt');

set(hObject,'visible','off');
set(handles.Analizar,'visible','on');
set(handles.Reiniciar,'visible','off');

```

A.5 Diseño del circuito impreso

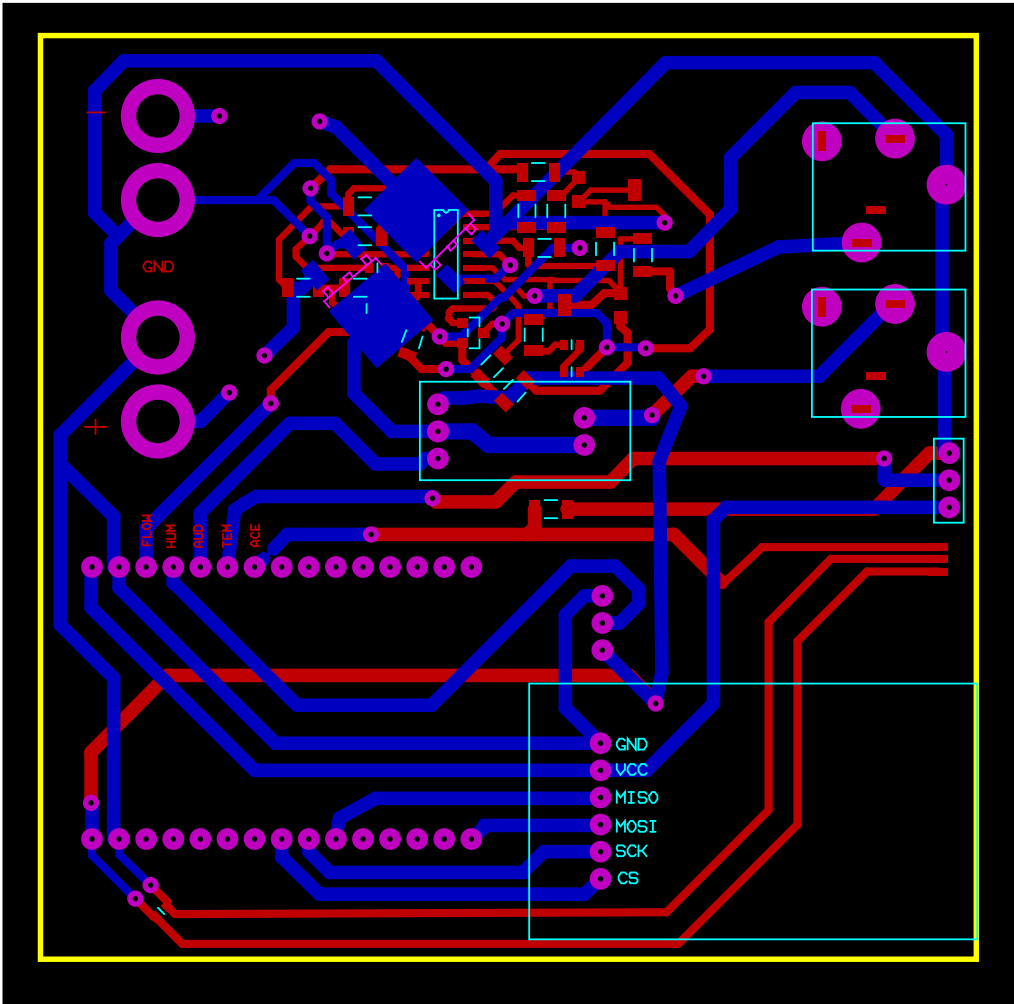


Fig. 67. Diseño PCB del sistema final.

A.6 Carta de consentimiento informado



CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL I.P.N.
Guía Operacional para la Evaluación Ética de los Proyectos de
Investigación en Seres Humanos

FOLIO _____

CARTA DE CONSENTIMIENTO INFORMADO

El que suscribe Sr/Sra/Srita. _____, de _____ años de edad, declaro que autorizo y consiento voluntariamente en participar en el proyecto de investigación: **“Sistema de evaluación foniatría para la detección temprana de anomalías en la producción de la voz”**, el cual se lleva a cabo bajo la responsabilidad del **Dr. Pablo Rogelio Hernández Rodríguez**. También declaro que he sido ampliamente informado en un lenguaje que entiendo claramente sobre el objetivo de este estudio: la medición de flujo de aire, audio y vibración en la superficie del cuello, durante una fonación, para evaluar de manera indirecta la calidad de la voz de manera no invasiva

Mi participación consiste en permitir el registro de las señales de flujo y audio a través de unos sensores que serán montados en una mascarilla y las vibraciones en la superficie del cuello por medio de un sensor colocado en la parte inferior izquierda del cuello.

También, he sido ampliamente informado sobre los beneficios y riesgos que representa participar en este estudio. Asimismo, recibiré, si lo solicito, reportes por escrito de los resultados que se me practiquen, los cuales pueden ser útiles para mi expediente clínico, como son: Medidas de frecuencia fundamental, *Jitter*, *Shimer*, y relación ruido-armónico obtenidas del análisis de mi señal de audio.

Se me ha informado y acepto los riesgos a que me someteré al participar en el estudio, los cuales son clasificados como mínimos. Se me ha informado que durante todo el curso del trabajo seré tratado con decoro, dignidad y discreción. Asimismo, todos los estudios correspondientes al trabajo de investigación serán gratuitos y en el remoto caso de sufrir algún daño resultado de los procedimientos empleados en el estudio, la atención profesional para su resolución será gestionada sin costo.

Se me ha notificado que la información que se obtenga será confidencial y que será empleada sólo para los fines de la investigación y que tengo el derecho de retirarme del estudio cuando así lo decida, sin ser sujeto a alguna medida de presión.

Ciudad de México, a _____ del mes de _____ de 2019.

Sr/Sra/Srita

Dr. Pablo Rogelio Hernández Rodríguez
Investigador Responsable