



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Politécnica Superior de Gandia

Desarrollo de una aplicación de análisis acústico de voz  
para uso sanitario

Trabajo Fin de Grado

Grado en Ingeniería de Sistemas de Telecomunicación, Sonido e  
Imagen

AUTOR/A: Peral de León, Jorge

Tutor/a: Sansaloni Balaguer, Trinidad M<sup>a</sup>

CURSO ACADÉMICO: 2022/2023

## Resumen

En el siguiente proyecto se plantea diseñar una aplicación informática de análisis acústico de la voz mediante la interfaz Matlab para ser utilizada por los logopedas especialistas en voz. Estos programas permiten realizar y estudiar de forma objetiva los parámetros acústicos de la voz de los pacientes para poder realizar tratamientos efectivos y personalizados. Para analizar la muestra se recoge la voz del paciente y el programa realiza el procesado de la muestra, que arroja los resultados de forma gráfica (espectrograma, espectro FFT y cepstrum) y numérica (parámetros acústicos). Además, el programa permite guardar los datos del paciente, la anamnesis, adjuntar documentos y almacenar los resultados obtenidos en las distintas evaluaciones.

En conclusión, el desarrollo de estos programas puede resultar muy útil para ayudar a los profesionales de la salud a agilizar sus evaluaciones y sus tratamientos, facilitándoles el trabajo mediante la tecnología.

**Palabras clave:** Matlab, Análisis de voz, Disfonía, Espectrograma, Matlab App Designer.

## Abstract

In the following project, we propose to design a computer application for voice acoustic analysis using a Matlab interface to be used by speech therapists specializing in voice. These programs allow the study, objectively, of the acoustic parameters of the patient's voice to perform effective and personalized treatments. To analyze the sample, the patient's voice is collected, and the program performs the processing of the sample, which yields the results graphically (spectrogram, FFT spectrum, and cepstrum) and numerically (acoustic parameters). In addition, the program allows the user to save the patient's data, fulfill anamnesis, upload documents, and store the results obtained in the different evaluations.

In conclusion, the development of these programs can be very useful to help health professionals speed up their evaluations and treatments, facilitating their work through technology.

**Keywords:** Matlab, Voice analysis, Dysphonia, Spectrogram, Matlab App Designer.

## Índice

Resumen.....	1
Abstract.....	1
1. Justificación, objetivos y etapas del proyecto. ....	6
2. Herramientas utilizadas. ....	7
3. Biomecánica de la fonación. ....	8
3.1. Potencia o fuelle (aparato respiratorio).....	8
3.2. Fuente o vibrador (la laringe). ....	8
3.3. Filtro o resonadores (pabellón faringo bucal y cavidades anexas). ....	9
3.4. Formantes. ....	9
4. Análisis acústico de la voz. ....	10
4.1. Anamnesis y valoración perceptual voz. ....	11
4.1.1. Anamnesis. ....	11
4.1.2. Herramientas análisis perceptual. ....	11
4.2. Análisis paramétrico de la voz.....	11
4.2.1. Espectrograma. ....	12
4.2.2. Espectro FFT.....	14
4.2.3. Cepstrum (análisis cepstral).....	14
4.2.4. Parámetros acústicos ....	15
5. Diseño de la aplicación.....	17
5.1. Base de datos.....	17
5.1.1. Diseño de la base de datos.....	17
5.2. Repositorios.....	19
5.3. Gestión de archivos. ....	20
5.4. Parámetros acústicos.....	20
5.4.1. <i>Shimmer</i> . ....	21
5.4.2. <i>Jitter</i> .....	22
5.4.3. <i>Dysphonia severity index (DSI)</i> .....	22

5.4.4.	<i>Signal-to-noise ratio (SNR)</i> .	23
5.4.5.	<i>Normalized noise energy (NNE)</i> .	23
5.4.6.	<i>Harmonic-to-noise ratio (HNR)</i> .	24
5.4.7.	Estimación de los formantes.	25
5.5.	Gráficas.	27
5.5.1.	Espectrogramas.	27
5.5.2.	FFT.	27
5.5.3.	Cepstrum.	28
5.6.	Interfaz de usuario.	28
5.6.1.	Páginas estáticas.	29
5.6.2.	Páginas dinámicas.	30
6.	La aplicación a nivel de usuario.	30
6.1.	Requerimientos.	30
6.2.	Tutorial.	31
7.	Resultados.	38
8.	Conclusiones y futuras líneas de trabajo.	38
	Bibliografía.	40
	Anexo 1. Resultados de la encuesta a logopedas para evaluar la oferta actual de software de análisis acústico de la voz.	42
	Anexo 2. Descripción de los ficheros del proyecto.	44
	Anexo 3. Encuesta de satisfacción, para evaluar la aceptación del usuario.	45
	Anexo 4. Resultados obtenidos en la aplicación con un paciente real en una consulta de logopedia.	47

## *Índice de ilustraciones*

Ilustración 1. Temporalización y nivel de actividad en cada etapa.....	7
Ilustración 2. Visión de conjuntos de los órganos de la voz y del habla [3]......	9
Ilustración 3. Gráficas de cada formante dependiendo de la vocal emitida.....	10
Ilustración 4. Espectrograma de banda estrecha vocal /a/ sostenida. ....	13
Ilustración 5. Espectrograma de banda ancha vocal /a/ sostenida. ....	13
Ilustración 6. Gráfica del espectro FFT y la posición de los formantes en una muestra de vocal /a/ sostenida.....	14
Ilustración 7. Cepstrum de una vocal sostenida en hombres y mujeres.....	15
Ilustración 8. Diagrama de la base de datos.....	18
Ilustración 9 Archivos de repositorios .....	19
Ilustración 10. Esquema del sistema de archivos de la aplicación.....	20
Ilustración 11. Acceso a la aplicación.....	31
Ilustración 12. Ventana de creación nueva cuenta. ....	31
Ilustración 13. Pantalla de datos del usuario. ....	32
Ilustración 14. Pantalla General. ....	32
Ilustración 15. Datos del paciente y de la mutua. ....	33
Ilustración 16. Anamnesis: Datos personales y Antecedentes ORL. ....	33
Ilustración 17. Anamnesis: Antecedentes personales y familiares.....	34
Ilustración 18. Anamnesis: Datos médicos de relevancia. ....	34
Ilustración 19. Anamnesis: Aspectos relacionados con la vida diaria y descripción del problema y sus causas. ....	34
Ilustración 20. Anamnesis: Balance Vocal y GRBASI.....	35
Ilustración 21. Ventana de documentos adjuntos. ....	35
Ilustración 22. Ventana de registro de pruebas. ....	36
Ilustración 23. Ventana de nueva prueba. ....	36
Ilustración 24 Ventana de muestra de resultados de las pruebas.....	37

### *Índice de tablas*

Tabla 1. Especificaciones Micrófono EGM-O4 ELEGIANE.....	8
Tabla 2. Correlatos espectrográficos del análisis perceptual [7]. .....	12
Tabla 3. Correlaciones básicas del análisis acústico. ....	14
Tabla 4. Frecuencias normativas por edad y sexo. [10].....	16
Tabla 5. Implementación de la visualización del cepstrum. ....	28
Tabla 6. Ejemplo de definición de datos del componente HTML en MATLAB. ....	29
Tabla 7. Ejemplo de implementación de evento DataChanged en JavaScript. ....	29
Tabla 8. Ejemplo de implementación del envío de datos a MATLAB con JavaScript. ....	29
Tabla 9. Ejemplo de generación dinámica de página HTML con MATLAB. ....	30

### *Índice de ecuaciones*

Ecuación 1. Cálculo del Shimmer relativo. ....	21
Ecuación 2. Cálculo del Jitter relativo. ....	22
Ecuación 3. Cálculo del DSI. ....	23
Ecuación 4. Cálculo del SNR en dB. ....	23
Ecuación 5. Cálculo del SNR en base a la media y la desviación estándar. ....	23
Ecuación 6. Cálculo del SNR en base a las potencias de señal de audio y señal de ruido .....	23
Ecuación 7. Cálculo del SNR para una única señal haciendo uso del NNE. ....	24
Ecuación 8. Cálculo del NNE en base al SNR de la señal. ....	24
Ecuación 9. Cálculo del HNR utilizando la función de autocorrelación. ....	24
Ecuación 10. Cálculo del HNR en dB. ....	24
Ecuación 11. Cálculo del número de coeficientes. ....	25
Ecuación 12. Cálculo del intervalo de aparición de formantes. ....	25
Ecuación 13. Cálculo del número de formantes. ....	25

# 1. Justificación, objetivos y etapas del proyecto.

Los programas de análisis de voz, como PRAAT, DrSpeech y VoceVista, se utilizan en clínicas de logopedia para realizar y estudiar de forma objetiva los parámetros acústicos de la voz de los pacientes y así realizar tratamientos efectivos y personalizados.

Hoy en día el número de programas que sirven para este propósito es escaso y en su gran mayoría, de pago, y los programas gratuitos como PRAAT que podrían servir para el mismo propósito quedan descartados por su complejo manejo.

Así pues, el objetivo principal del proyecto es desarrollar una aplicación informática con un entorno de sencillo manejo basado en pestañas, para su uso clínico por logopedas, fonoatras y otorrinolaringólogos; concretamente se utilizará en la evaluación, diagnóstico y tratamiento de patologías vocales.

La aplicación permitirá evaluar a los pacientes en diferentes etapas del tratamiento y permitirá observar su evolución de forma más objetiva.

El programa podrá:

- Gestionar, guardar y editar los datos personales y la anamnesis de los pacientes.
- Realizar los análisis acústicos (cálculo de parámetros) de voz hablada y cantada necesarios para la evaluación.
- Almacenar las muestras vocales y los análisis realizados en su fichero correspondiente.

Para conseguir el objetivo principal será necesario:

- Conocer la opinión de los logopedas acerca de las aplicaciones de análisis acústico de la voz, seleccionar los parámetros acústicos principales y las funcionalidades de la aplicación.
- Buscar y revisar la bibliografía actualizada sobre el tema.
- Definir los algoritmos para el cálculo de los parámetros acústicos seleccionados.
- Diseñar la interfaz de usuario de la aplicación.
  - Determinar las tecnologías a utilizar para la implementación de la aplicación.
  - Fijar la estructura de la base de datos. Establecer que el programa permita añadir y modificar pacientes.
  - Decretar que el programa recoja y analice una señal acústica.
- Testear la aplicación.

Para ello se ha organizado el proyecto en 5 fases: documentación, planificación, desarrollo, análisis y evaluación. Estas etapas se han representado en la Ilustración 1. En la propia figura se muestra la temporalización y el nivel de actividad de estas.

- Etapa 1: Documentación. En esta etapa, se decidió el tema del proyecto y se llevó a cabo una búsqueda bibliográfica exhaustiva para poder definir los objetivos de este y planificar los pasos a seguir. También se realizó una encuesta a una pequeña muestra de logopedas para conocer las necesidades reales que tienen en su práctica clínica con la finalidad de que la aplicación les sea útil.
- Etapa 2: Planificación. Tras esto se definieron los objetivos generales y los específicos necesarios para desarrollar la aplicación, se organizaron las diferentes tareas a realizar y se estimaron los tiempos requeridos para cada una de ellas. Además, se eligió el material técnico necesario para llevar a cabo la construcción de la aplicación y la grabación de los audios.
- Etapa 3: Desarrollo. En esta etapa nos centramos en conseguir los objetivos planteados para desarrollar la aplicación en los tiempos estimados. Durante el desarrollo del proyecto se fueron revisando los objetivos y las tareas para ir definiendo cada vez más el aspecto y funcionamiento de la aplicación, también se desecharon algunas de las ideas iniciales.

- Etapa 4: Análisis. Se realizó un seguimiento del proyecto, comprobando continuamente su desarrollo y su funcionamiento. Se fueron corrigiendo y subsanando los errores que iban apareciendo. También se modificaron y se ajustaron las medidas de los parámetros para que funcionaran correctamente y se finalizó toda la creación del *back end* y el *front end*.
- Etapa 5: Evaluación. Por último, se realizaron las pruebas de la aplicación para corroborar que todo funcionaba correctamente. Esto permitió observar los resultados y compararlos con los objetivos que nos habíamos planteado al principio para comprobar que se habían cumplido.

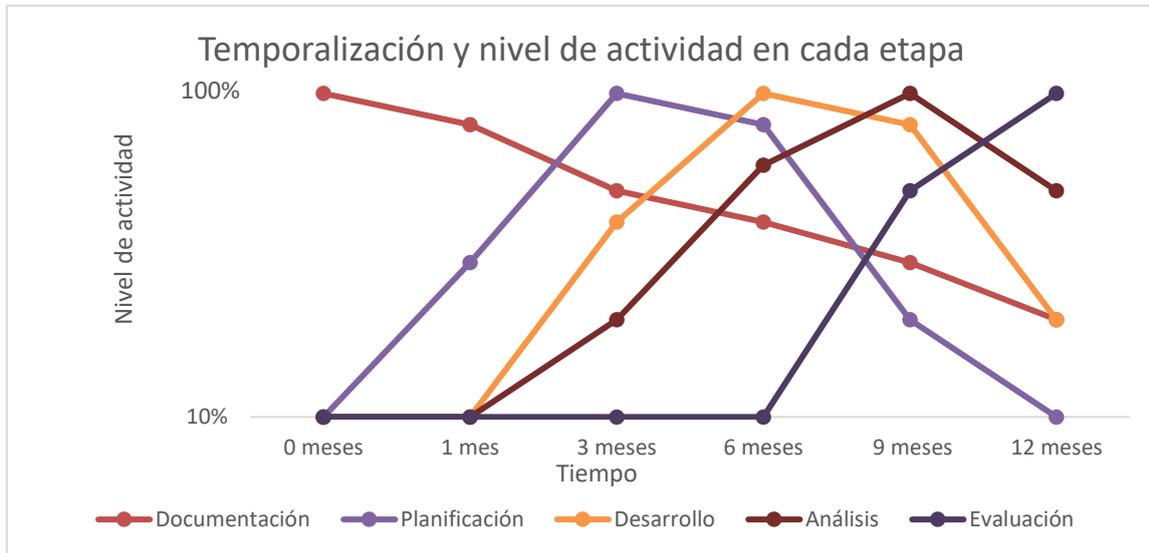


Ilustración 1. Temporalización y nivel de actividad en cada etapa.

## 2. Herramientas utilizadas.

Tanto las pruebas como el desarrollo de la aplicación han sido realizadas a la par en 2 equipos, siendo el más limitante de ellos un ordenador portátil HP Pavilion i7 de séptima generación con 16GB de memoria RAM.

Para el desarrollo de la aplicación se ha utilizado el programa Matlab en su versión de 2022b, debido a que es la versión más estable en el periodo de realización del proyecto. Dentro de esta se han descargado y utilizado los paquetes de *Signal Processing Toolbox*, debido a la amplia gama de funcionalidades disponibles para el análisis de señal, y el de *AppDesigner*, útil para el diseño de aplicaciones tanto web como de escritorio y para la integración fácil de una interfaz de usuario con las funciones de análisis implementadas con Matlab.

Se ha descargado el sistema de gestión de bases de datos SQLite en la versión 3.0 para la gestión de la base de datos necesaria para el funcionamiento de la aplicación.

Para el desarrollo de las páginas HTML que se muestran en la interfaz de usuario, se ha utilizado el editor de Visual Studio Code en la versión 1.71. Para la edición de estilo de las páginas HTML, se ha descargado y utilizado librería de CSS Bootstrap 5, se recomienda la descarga puesto que los archivos HTML visualizados en Matlab tienen limitado el acceso a enlaces externos.

Para la realización de las pruebas se ha seleccionado el micrófono que más se ajustaba, dentro de los posibles, a las especificaciones óptimas que ha de tener para el uso dado [1], un EGM-04 ELEGANT, cuyas especificaciones se definen en la Tabla 1.

Conexión	USB-A
Tipo	Condensador
Patrón	Omnidireccional
Rango frecuencias	60Hz - 18kHz

Tabla 1. Especificaciones Micrófono EGM-O4 ELEGIAN.T.

### 3. Biomecánica de la fonación.

*“La fonación está determinada por características neuromusculares, biomecánicas, acústicas, y aerodinámicas que, a su vez determinan la duración, tono, sonoridad, cualidad, registro vocal y movimientos de los pliegues vocales a través del control en el largo, masa, tensión, contorno de los pliegues vocales, aducción de los aritenoides, presión subglótica y tamaño y forma del tracto vocal”* (Susanibar, Dioses, Marchesan, Guzmán, Leal, Guitar and Junqueira, 2016).

Con el objetivo de comprender mejor para qué se realiza el análisis acústico de la voz, es necesario conocer cómo se produce la fonación y que mecanismos influyen en los cambios de los parámetros que evaluaremos más adelante, ya que con la aplicación de este proyecto podremos observar y obtener parámetros que nos indicarán qué está fallando en el proceso fonatorio.

A continuación, se exponen las tres partes anatómicas imprescindibles en la producción vocal y que deben funcionar correctamente para que la voz se produzca con una calidad óptima. Si alguna de estas partes falla (por función o por estructura) la voz no se emitirá de manera adecuada [2].

#### 3.1. Potencia o fuelle (aparato respiratorio).

La potencia que se necesita para emitir voz comienza en el aparato respiratorio, que es el encargado de expulsar la cantidad de aire correcta para mover los pliegues vocales, generar vibración y producir sonido.

En el inicio de la producción vocal es necesario realizar un impulso respiratorio, con la ayuda del diafragma (principal músculo inspiratorio). Gracias a él, se introduce aire en los pulmones, pasando por la tráquea, los bronquios y bronquiolos acabando en los alveolos pulmonares. Durante la fonación, se realiza un soplo espiratorio que recorre el camino inverso accediendo a la laringe con una presión y velocidad que fluctúan según el tipo de voz que se va a producir.

#### 3.2. Fuente o vibrador (la laringe).

La laringe es el órgano principal de la voz. Está formada por el hueso hioides, por una serie de cartílagos (tiroides, cricoides y aritenoides) que se encuentran unidos por fascias y ligamentos rodeados de músculos (intrínsecos y extrínsecos) y se recubren por una mucosa.

El músculo más importante de la laringe es el músculo tiroaritenoso, que forma los pliegues o cuerdas vocales. Estos músculos pueden alejarse y aproximarse entre sí. Al aproximarse son capaces de vibrar, gracias a la salida del aire que proviene de los pulmones, y producir sonido.

En ocasiones los pliegues vocales no se mueven correctamente y se genera un hiato glótico (los pliegues vocales no se juntan del todo y vibran menos) esto produce un escape de aire al emitir sonido, que generará una voz débil, soplada o estrangulada (si aparece tensión en la musculatura extrínseca).

### 3.3. Filtro o resonadores (pabellón faringo bucal y cavidades anexas).

Los resonadores actúan como amplificador y modificador de la onda sonora. Gracias a ellos podemos variar la corriente de aire, frenándola, deteniéndola totalmente o vibrando por su acción. Algunos de los órganos que constituyen el aparato resonador son: la faringe, la boca, las fosas nasales y los senos. Es muy importante aprender a utilizar los resonadores, ya que, gracias a ellos y a su función de amplificación, los pacientes conseguirán emitir la voz sin realizar esfuerzos. En canto son imprescindibles para dar brillo a la voz, colocar las notas correctamente, realizar cambios de registro sin hacerse daño, amplificar sonidos o realizar técnicas específicas (como el *twang*, en el que es muy importante la activación del esfínter ariepiglótico).

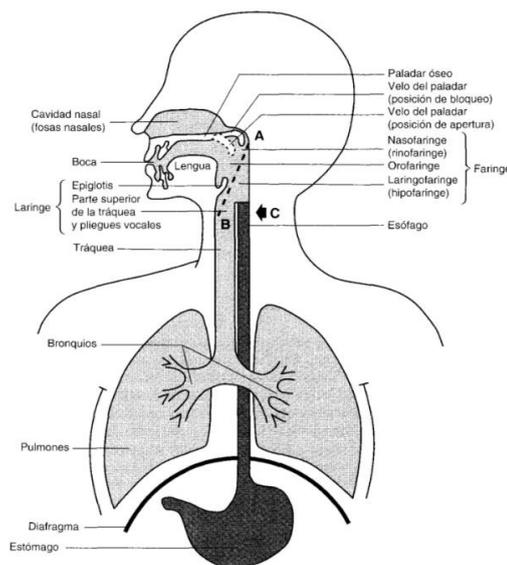


Ilustración 2. Visión de conjuntos de los órganos de la voz y del habla [3].

En conclusión, el sonido que se genera en los pliegues vocales (fuente) atraviesa el tracto vocal (filtro), y en él se transforma a su forma compleja. Las propiedades resonanciales que tienen las cavidades del tracto vocal pueden modificar el sonido. De esta forma, algunas cavidades potencian el sonido que se genera en la glotis y hay otras que lo atenúan. Las frecuencias de resonancia naturales que surgen del tracto vocal se denominan formantes.

### 3.4. Formantes.

Las frecuencias que provienen del tracto vocal dependen de su forma y de su configuración. Podemos identificarlas en el espectro de una vocal, cuando se descompone la señal acústica. Este espectro mostrará unos picos o regiones de frecuencias que tienen mayor amplitud y que sobresalen por encima de todas las demás frecuencias.

Los formantes son las frecuencias o regiones en las que aparecen estos picos, su análisis nos permite obtener datos acústicos significativos de la configuración del tracto vocal. Las frecuencias formánticas cambiarán dependiendo de las variaciones que se realicen en el tracto vocal, como el cambio de la posición de la lengua, el cierre o apertura de la mandíbula, los labios, el descenso o la elevación de la laringe, del velo del paladar, etc.

Por ello, la frecuencia de los formantes nos arroja información acerca de cómo es la articulación de una persona.

El trazado de estos formantes nos entrega también información sobre la articulación y además nos da información de la estabilidad de la voz. Cuando la voz tiene estabilidad podemos observar una buena definición de los formantes, en cambio, cuando la voz es inestable, los formantes son difíciles de identificar.

Si la energía y las bandas de los formantes están bajas, quiere decir que la energía está disminuida y por lo tanto sabremos que la voz emitida está perdiendo intensidad (volumen)

Por otra parte, si observamos disminución de energía en las bandas y formantes altos, sabremos que esa voz está perdiendo brillo.

Cada formante está determinado por una estructura anatómica determinada, tal y como se expone a continuación y se muestra en la Ilustración 3:

- La f0 está determinada por los pliegues vocales, los cambios de frecuencia en esta nos indicarán la cantidad de ciclos que se están produciendo por segundo, es decir, si la voz está grave o aguda.
- La f1 está determinada por la apertura bucal, lo que quiere decir que si la frecuencia es menor habrá menor apertura bucal y si hay una frecuencia alta habrá mayor apertura bucal.
- La f2 está determinada por el cuerpo de la lengua y el redondeamiento o protrusión labial. Si la frecuencia es alta la lengua estará anteriorizada y habrá menos protrusión bucal. Por el contrario, si la frecuencia es baja, nos indicará que la lengua está en una posición más posterior y hay mayor protrusión labial.
- Por último, la f3 se ve influenciada por la colocación de la punta de la lengua y el velo del paladar. Si la frecuencia es baja la punta de la lengua se encontrará en una posición anterior y baja y habrá descenso velar. Si la frecuencia es alta la punta de la lengua estará en posición elevada y posterior y el velo del paladar estará elevado. [4]

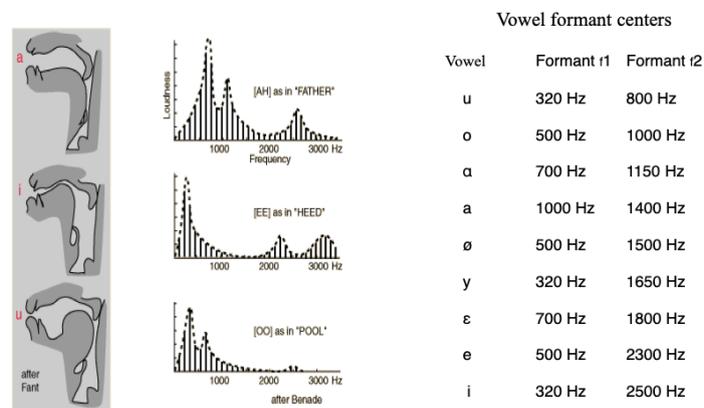


Ilustración 3. Gráficas de cada formante dependiendo de la vocal emitida.

#### 4. Análisis acústico de la voz.

Para determinar si una voz es sana o patológica en las clínicas de logopedia se realizan análisis perceptuales de la voz y análisis objetivos con programas de procesado de voz. Este segundo análisis no sustituye, sino que complementa la evaluación perceptual que realizan los foniatras y logopedas.

## **4.1. Anamnesis y valoración perceptual voz.**

### **4.1.1. Anamnesis.**

La anamnesis es un proceso que se realiza en la primera consulta, mediante el cual un profesional sanitario recopila información clínica a partir de un diálogo con el paciente en el que se le realizan preguntas acerca de sus datos personales, el motivo de consulta, sus antecedentes médicos, personales y familiares, medicamentos, hábitos diarios, informes o pruebas de otros profesionales y explicación de la problemática actual.

Esta información, junto con los datos que aporta el análisis acústico de la voz y las herramientas perceptuales, permite al terapeuta de la voz a realizar una buena evaluación, diagnóstico e intervención. [5]

### **4.1.2. Herramientas análisis perceptual.**

Hay diversas herramientas de para realizar valoraciones perceptivas de la voz, pero en este proyecto hemos seleccionado una de las escalas más fiables y una de las más utilizadas por los logopedas, la escala GRBASI (Grade, Roughness, Breathiness, Asteny, Strain, Inestability) [6] [7].

Está escala es subjetiva, de fácil aplicación y se cumplimenta por el evaluador. Consiste en escuchar la voz del paciente y desde la perspectiva del oído del clínico hay que indicar con números la aparición y el grado de los diferentes aspectos acústicos.

Las características que se deben observar en la voz son:

- G = Grado de disfonía.
- R = Aspereza.
- B = Soplo
- A = Asténica.
- S = Tensión.
- I = Inestabilidad.

Se indicará su gravedad mediante un número del 0 al 3. En el que 0 es normal, 1 es leve, 2 es moderado y 3 es severo.

## **4.2. Análisis paramétrico de la voz.**

El análisis acústico de la voz es una manera objetiva de estudiar y analizar la voz humana. Al ser objetivo es capaz de eliminar la subjetividad de la evaluación realizada mediante el oído y criterio del evaluador.

El análisis comienza con la grabación de la voz del paciente empleando una cadena de grabación adecuada. El paciente debe realizar distintos ejercicios fonatorios siguiendo las instrucciones que se le indican, por ejemplo, emisión de vocal sostenida, glissando, habla espontánea, voz cantada, etc. Estas señales se incorporan a un software capaz de extraer las características de la onda sonora, analizarlas de manera cualitativa y cuantitativa y devolver resultados gráficos y parámetros numéricos que deberán ser interpretados por un evaluador.

Tal y como indica Faustino Núñez – Batalla et al (2011) [8] la evaluación de la voz se debe realizar a través de herramientas perceptuales, como el GRBASI (incluido en la aplicación de este proyecto), y contrastar los resultados perceptivos con los parámetros acústicos y espectrográficos.

Por ello, los profesionales que utilizan estas herramientas deben estar formados para poder interpretar los resultados correctamente.

Entre las utilidades del análisis acústico de la voz se incluyen:

- Transmitir mayor comprensión acústica sobre el output vocal y poder realizar asociaciones entre este análisis y el análisis perceptivo de la voz (GRBASI).
- Entregar datos normativos para realidades vocales diferentes.
- Otorgar la posibilidad de trazar la línea base de la voz de un paciente mediante los resultados obtenidos.
- Monitorear y dar *feedback* sobre el avance de la terapia o el entrenamiento vocal de los pacientes.
- Es un buen instrumento que sirve para poder detectar precozmente posibles patologías vocales.

A pesar de la gran aportación que ofrece el análisis acústico de la voz, la evaluación mediante herramientas perceptuales realizadas por el oído entrenado de los profesionales sanitarios nunca se debe eliminar. El análisis acústico es una herramienta complementaria de evaluación vocal que sirve para que los clínicos obtengan de manera objetiva y no invasiva (no se aborda directamente a la persona) los datos que les indiquen el estado vocal de los pacientes. Tal y como se muestra en la Tabla 2, en el espectrograma, aparecen unas imágenes en las que se puede conocer el correlato espectrográfico, pero el encargado de realizar la correlación perceptual debe ser el logopeda.

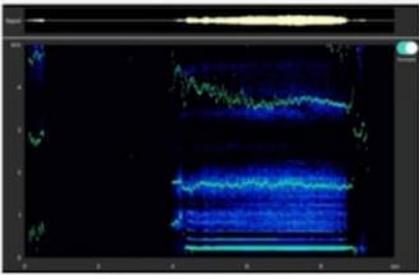
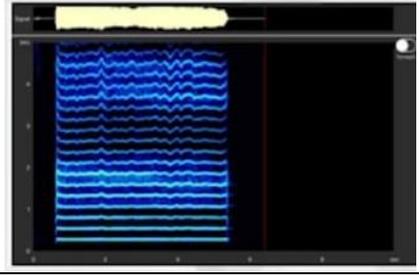
PERCEPTUAL	CORRELACIÓN ESPECTROGRÁFICA	IMAGEN ESPECTROGRAMA
El cierre glótico es insuficiente y aparece ruido por el paso del aire.	Presencia de ruido en el espectrograma.	
Debilidad en fonación, típica en voz hipocinética o hipofuncional (p. ej.: parálisis de cuerda vocal).	No aparecen ni ruido ni armónicos en el espectrograma.	
Impresión auditiva de sobreesfuerzo y tensión asociada a la fonación.	Señal acústica con mucha intensidad, hay presencia de ruidos, armónicos y subarmónicos.	

Tabla 2. Correlatos espectrográficos del análisis perceptual [9].

#### 4.2.1. Espectrograma.

El espectrograma es una representación gráfica de la frecuencia y la intensidad de la forma de onda en función del tiempo.

El ancho de banda del filtro aplicado en la creación del espectrograma determina la claridad de los elementos que podemos observar en él.

Un filtro de banda estrecha presenta una buena resolución frecuencial o buena resolución de los armónicos, como se puede observar en la Ilustración 4.

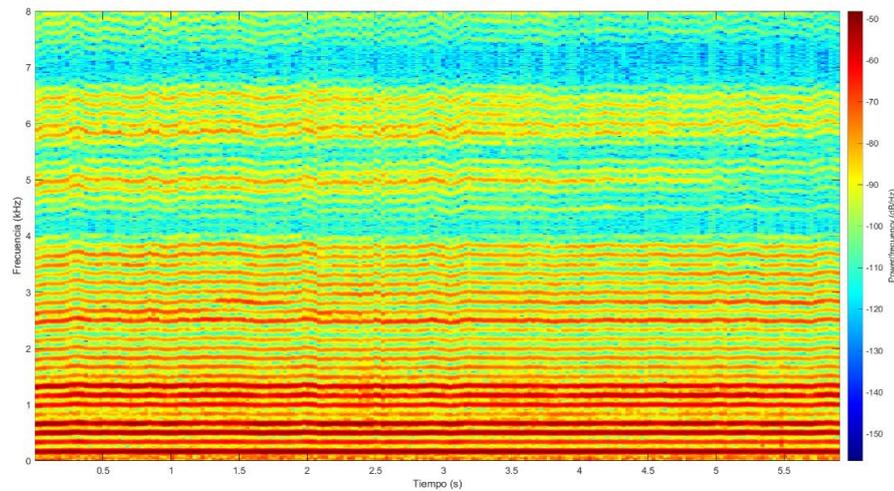


Ilustración 4. Espectrograma de banda estrecha vocal /a/ sostenida.

El filtro de banda ancha muestra una buena resolución temporal que permite que se visualicen mejor los pulsos glóticos y que haya buena resolución para las resonancias del tracto vocal o formantes [2]. En la Ilustración 5 se muestra un ejemplo de como se vería un espectrograma de banda ancha.

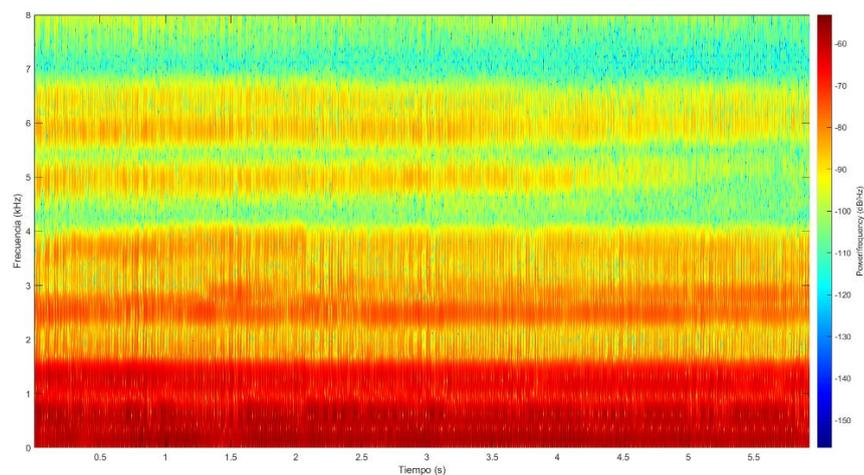


Ilustración 5. Espectrograma de banda ancha vocal /a/ sostenida.

Como explicábamos anteriormente, los espectrogramas deben ser interpretados por un profesional especializado en voz. El espectrograma por sí solo no arroja ningún resultado concreto, por lo que la aplicación aporta la imagen espectrográfica y los datos numéricos, pero no puede leer aquellas cosas que aparecen en la imagen y que varían de un paciente a otro.

La presencia de ruido, la nitidez con la que aparecen los formantes o su ausencia, si hay subarmónicos o no, etc. se puede relacionar con información perceptual de la voz, que es de gran utilidad para los logopedas siempre y cuando estén formados y lo sepan descifrar. En la Tabla 3 se puede observar cada uno de los parámetros físicos con la variación acústica que aparece en el espectrograma y el correlato psicoacústico que equivale a cada una de esas variaciones.

Parámetro físico	Variación acústica	Correlato psicoacústico
Energía total (Et)	Ascendida	Sonoridad fuerte
	Descendida	Sonoridad débil
	Irregular	Timbre áspero
	Con cortes repentinos	Emisión con quiebres
	Rápido crecimiento inicial	Ataque vocal duro
	Ausente en final de frase	Final de frase "abortado"
Frecuencia fundamental (f0)	Ausente	Cuchicheo
		Afonía
	Irregular	Ruido
	Ascendida	Tono agudizado
	Descendida	Tono agravado
	Sin variación	Tono monótono
	Con doble contorno	Diplofonía
	Rango disminuido	Extensión vocal limitada
Formantes (f1, f2, f3, fn)	Valor normal desplazado	Articulación incorrecta
	Formantes altos atenuados	Mordientes ausente- voz opaca
	Formantes bajos atenuados	Voz sin volumen
	Formantes nasal banda de f1 aumentada	Voz nasalizada
	Banda de f3-f4 aumentadas	Voz sin brillo
Ruido	Ruido en todo el espectro	Ruido acoplado
	En posición inicial de emisión	Ataque soplado

Tabla 3. Correlaciones básicas del análisis acústico.

#### 4.2.2. Espectro FFT.

En el campo del análisis de voces disfónicas, el espectro FFT (*Fast Fourier Transformation*) se representa mediante una imagen gráfica que se obtiene de la extracción de un punto determinado de tiempo del espectrograma. Nos enseña la amplitud relativa de los armónicos, refleja la función de transferencia, la combinación del espectro de la fuente y las propiedades de radiación. En este, no es posible determinar de forma precisa la frecuencia de los formantes ya que el pico de la frecuencia no suele coincidir con los armónicos [2], tal y como se muestra en la Ilustración 6.

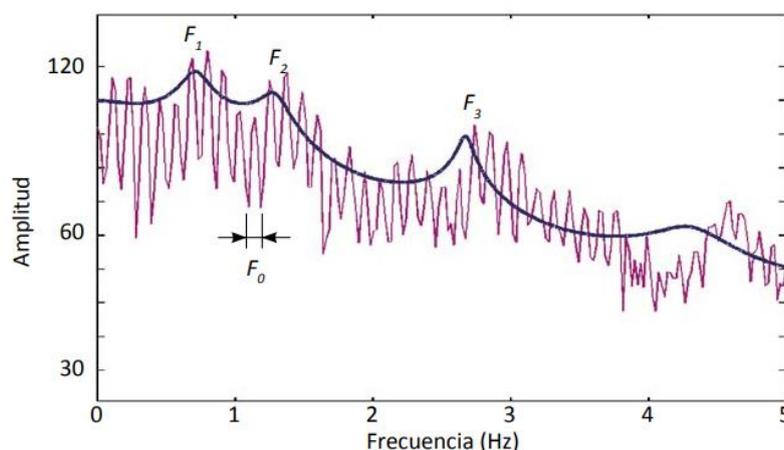


Ilustración 6. Gráfica del espectro FFT y la posición de los formantes en una muestra de vocal /a/ sostenida.

#### 4.2.3. Cepstrum (análisis cepstral)

En el análisis de la calidad vocal, lo que se observa en el análisis cepstral es la prominencia y estabilidad del pico cepstral. En la Ilustración 7 se muestran los picos cepstrales de una vocal sostenida /a/ en hombres y mujeres.

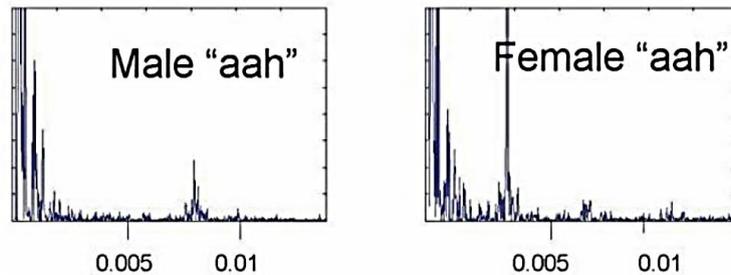


Ilustración 7. Cepstrum de una vocal sostenida en hombres y mujeres.

Un pico cepstral de valor alto nos indica que la voz se caracteriza por una estructura armónica bien definida y la anchura reducida nos indica que hay una gran estabilidad de la frecuencia, por lo que será una voz de buena calidad.

Por el contrario, cuando el pico cepstral presenta valores bajos e inestabilidad en la anchura es indicativo de alto grado de inestabilidad por la baja energía de la señal (pico bajo) una estructura armónica débilmente definida. Ambos rasgos suelen aparecer en voces de baja calidad, como las voces sopladadas roncas [2].

#### 4.2.4. Parámetros acústicos

Además de utilizar las técnicas espectrográficas o cepstrales, también disponemos de medidas o índices numéricos mediante los cuales obtenemos valores objetivos y que no requieren de la experiencia visual del clínico, como en la interpretación de los espectrogramas en las que sí que se requiere la experiencia del profesional especialista.

Entre estos índices podemos encontrar: las medidas de perturbación de la amplitud, las medidas de perturbación de la frecuencia y los índices de ruido glótico. [2]

##### 4.2.4.1. Jitter.

El jitter es una característica acústica que indica la variación de la frecuencia. Se genera por la vibración irregular de las cuerdas vocales y se percibe como aspereza o ronquera. Medir este parámetro sirve para detectar posibles patologías vocales. En el apartado 5.4.2 se indica cómo se obtiene este parámetro.

Dentro de los índices de perturbación de la frecuencia, encontramos el Jitter relativo, que es el índice más usado en la clínica vocal. Su valor normativo debe ser cercano a 1 y los valores mayores a 1 nos indican que hay una alteración. Se detecta mejor en las vocales largas y sostenidas [10].

##### 4.2.4.2. Shimmer.

El shimmer es una característica acústica causada por la vibración irregular de los pliegues vocales que nos indica una variación en la amplitud. Al igual que el jitter, se puede percibir como ronquera o aspereza en la voz y sirve para detectar patologías vocales. En el apartado 5.4.1 se indica cómo se obtiene este parámetro.

En los índices de perturbación de la amplitud, evaluaremos el shimmer relativo por ser el más utilizado en el análisis de voz. Su valor normativo es cercano a 3 y los valores por encima de 3 estarían fuera de la normalidad. Se analiza mejor en vocales largas y sostenidas [10].

##### 4.2.4.3. NHR / SNR / NNE.

Yumoto (1982) propuso una manera de cuantificar la calidad vocal. En el espectrograma de la voz aparecen componentes periódicas y componentes de ruido, uno de los rasgos más característicos de las disfonías es el reemplazo de los armónicos por ruido. De aquí se obtiene el mejor índice predictor de disfonía, que es la relación de amplitud armónico/ruido (HNR). La calidad vocal es mejor cuanto más alto es el valor. En disfonías, el valor del HNR será bajo.

La señal ruido/ armónico (SNR) también nos arroja resultados relacionados con la cantidad de ruido que aparece en la voz. En el apartado 5.4.4 se indica cómo se obtienen es valor.

La energía de ruido normalizada (NNE) es especialmente útil en la detección de cáncer de laringe, nódulos vocales y parálisis de cuerda vocal [2]. El procedimiento para su cálculo se puede observar en el apartado 5.4.5.

#### 4.2.4.4. *Frecuencia fundamental.*

La frecuencia fundamental (f0) es el número de ciclos por segundo que se realizan en el movimiento de la glotis en cada fonación. En una persona sana con voz normal, el movimiento de los pliegues vocales será similar de un ciclo a otro, por lo que los periodos glóticos también serán similares. La F0 es un parámetro que se puede modificar por la persona que habla y en función de los objetivos vocales que tenga. Estas modificaciones se producen gracias a la actividad de la musculatura que controla la tensión del cuerpo, de la cubierta de los pliegues vocales (cricoideo y tiroaritenoides) y de la presión subglótica. En la evaluación de las disfonías, los pacientes tienen que emitir una vocal sostenida en su tono normal de habla para poder comparar con los parámetros normativos [11]. En la Tabla 4 se muestran las frecuencias normativas de hombres y mujeres para las distintas franjas de edad.

SEXO MASCULINO		SEXO FEMENINO	
Edad	F0	Edad	F0
7	294	7	281
8	297	8	288
10	270	11	238
11	220	19	217
14	242	20 - 29	224
19	117	30 - 40	196
20 - 29	120	40 - 50	189
30 - 39	112	60 - 69	200
40 - 49	107	70 - 79	202
50 - 59	118	80 - 94	200
60 - 69	112		
70 - 79	132		
80 - 89	146		

Tabla 4. Frecuencias normativas por edad y sexo. [10]

#### 4.2.4.5. *Tiempo máximo espiración.*

El tiempo máximo de espiración sirve para medir la capacidad respiratoria. Se mide en segundos con un cronómetro. Se solicita al paciente que coja aire y mientras produce una /s/ sorda (sin sonido) tire el aire y aguante durante el mayor tiempo posible, intentado gestionar correctamente el flujo de aire, es decir, sin tirarlo muy bruscamente.

Cuando este índice es muy bajo (< 10 s) significa que la persona tiene una capacidad pulmonar deficiente. [3]

#### 4.2.4.6. *Tiempo máximo fonación.*

El tiempo máximo de fonación se emplea para medir la coordinación fonorespiratoria. Se mide en segundos con un cronómetro. Se solicita al paciente que coja aire y emita una /a/ o una /z/ sonora (con sonido) y aguante sin parar el mayor tiempo que le sea posible.

En disfonías es muy típico observar que el TMF está muy por debajo al TME, ya que los pacientes disfónicos suelen tener buena capacidad pulmonar si solo expulsan aire, pero al expulsar el sonido, aparecen descoordinación en los pliegues vocales, exceso de tensión en la musculatura y/o escape de aire por incompetencia en el cierre glótico, por lo que el TMF disminuye notablemente en comparación al TME. [3]

#### 4.2.4.7. *Rango vocal.*

El rango vocal es una gama de frecuencias que varía en función de sexo, la edad y con el entrenamiento vocal de cada individuo. Se puede obtener mediante la grabación de un glissando (realizar una voz desde el tono más bajo que la persona pueda emitir subiendo hasta el tono más alto posible, sin hacerse daño). Con él se puede calcular la frecuencia mínima y la frecuencia máxima a la que llega la voz del paciente.

Toda esa gama frecuencial que va desde la frecuencia mínima hasta la máxima es el rango vocal. Gracias a su evaluación, los terapeutas pueden adaptar los ejercicios con las notas y sonidos que el paciente puede realizar de manera cómoda sin realizar esfuerzos. Calculándolo a mitad de terapia se puede observar si el paciente ha aumentado este parámetro y con ello observar su evolución. Este valor es muy usado por los cantantes y profesores de canto. [3]

## 5. Diseño de la aplicación.

La aplicación consta de distintas partes: base de datos, repositorios, gestión de archivos, cálculo de parámetros acústicos y gráficas, e interfaz de usuario. En este apartado se describe el diseño de cada una de ellas. Para conocer los distintos archivos y su propósito dentro de la aplicación se puede consultar el Anexo 2.

### 5.1. Base de datos.

Para la implementación de la base de datos se decidió utilizar una base de datos relacional. Por ello se ha utilizado el motor de bases de datos SQLite, debido a que se trata de un proyecto de dominio público y a que su biblioteca en lenguaje C que implementa un motor de base de datos SQL pequeño, rápido, autónomo, de alta fiabilidad y con todas las funciones; aparte de ser el motor de base de datos más utilizado del mundo.

#### 5.1.1. Diseño de la base de datos.

En base a los requerimientos de la aplicación definidos tras el estudio de la encuesta realizada a un grupo reducido de logopedas españoles (Anexo 1. Resultados de la encuesta a logopedas para evaluar la oferta actual de software de análisis acústico de la voz. ), se diseñó el esquema de la B.D. que contiene las siguientes tablas:

- *Users*: Tabla que alberga los datos de los distintos usuarios que acceden a la aplicación. Como ID único (*PRIMARY KEY*) se decidió utilizar el correo electrónico, para evitar que un usuario generase múltiples cuentas con el mismo. Esta tabla también es la encargada de guardar la contraseña del usuario y por lo tanto los parámetros de inicio de sesión se comprobarán en ella.
- *Pacients*: Tabla que contiene los datos de identificación de los distintos pacientes que pueda tener el usuario. La tabla está relacionada por la *FOREIGN KEY ID\_USUARIO* que referencia la *PRIMARY KEY* de la tabla *Users* con una relación many-to-one.
- *Anamnesis*: Tabla que engloba todos los datos de la anamnesis de un paciente. Los datos necesarios para la anamnesis se obtuvieron basándose en el artículo de "Herramientas para la evaluación logopédica de la voz" [5]. La tabla está

relacionada por la *FOREIGN KEY IDUSER* que referencia la *PRIMARY KEY* de la tabla *Pacientes* con una relación *one-to-one*.

- *Documents*: Tabla que referencia los distintos archivos anexados a un paciente, como, por ejemplo, un informe del otorrinolaringólogo. La tabla está relacionada por la *FOREIGN KEY IDUSER* que referencia la *PRIMARY KEY* de la tabla *Pacientes* con una relación *many-to-one*.
- *Pruebas*: Tabla que incluye las distintas pruebas realizadas a un paciente, archivos a los que hace referencia y los parámetros acústicos obtenidos en dicha prueba. La tabla está relacionada por la *FOREIGN KEY IDUSER* que referencia la *PRIMARY KEY* de la tabla *Pacientes* con una relación *many-to-one*.
- *Notas*: Tabla que alberga el nombre de las notas (en formato anglosajón) y su valor frecuencial.

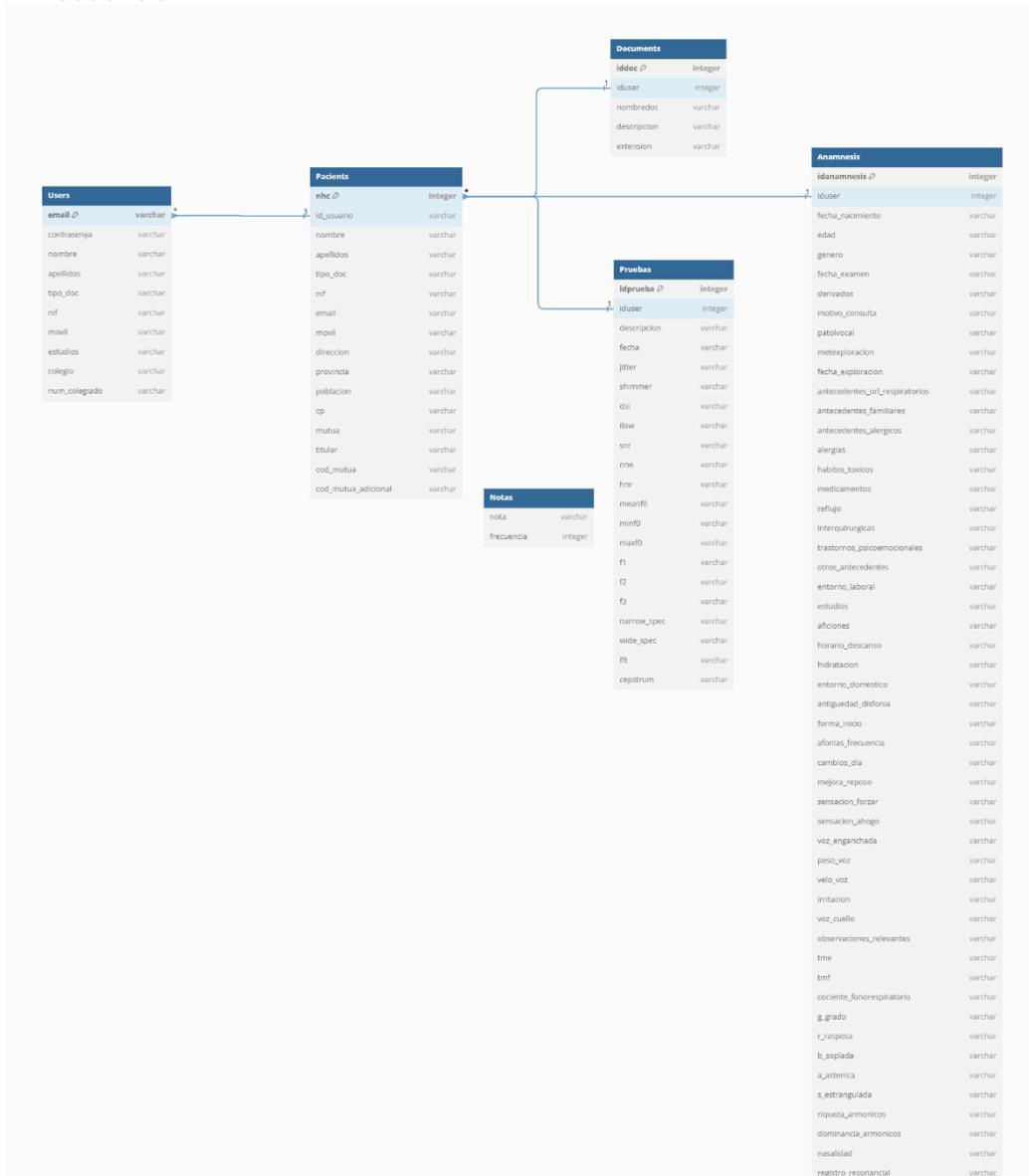


Ilustración 8. Diagrama de la base de datos.

En la Ilustración 8 se muestra el esquema completo de las tablas de la B.D., los datos que albergan, y de las relaciones entre ellas.

## 5.2. Repositorios.

Para facilitar las consultas a las distintas tablas de la B.D. se implementaron múltiples repositorios con los comandos SQL para la inserción, recuperación, edición y borrado de los datos, y que pudieran ser llamados desde las distintas funciones y eventos implementados en la aplicación.

Cada uno de estos repositorios ha sido implementado como una clase de Matlab independiente, generando uno para cada tabla de las incluidas en la B.D. En la Ilustración 9 se muestra una imagen con las distintas clases de repositorios implementadas.



Ilustración 9 Archivos de repositorios

Los repositorios han sido implementados utilizando las siguientes funciones de Matlab:

- *sqlite*: que establece la conexión a la base de datos, dada la cadena de la ubicación del archivo.
- *Insert*: En base a la conexión generada en la llamada al método *sqlite*, el nombre de la tabla (*[nombre\_tabla]*), una lista de celdas con los nombres de las columnas a insertar (*[nombres\_columnas]*) y otra lista de celdas con los valores a insertar en dichas columnas (*[valores\_columnas]*); ejecuta un comando de inserción de la nueva fila en la tabla. Lo que en SQL se traduciría en el siguiente comando.  
“*INSERT INTO [nombre\_tabla] ([nombres\_columnas]) values ([valores\_columnas]);*”
- *fetch*: Ejecuta una QUERY de recuperación de datos en base a una conexión y a una cadena de comando SQL del tipo *SELECT*.  
“*SELECT [nombres\_columnas] FROM [nombre\_tabla] WHERE [condiciones]*”
- *execute*: Ejecuta un comando de SQL en base a una conexión y una cadena de comando SQL del tipo *DELETE* o *UPDATE*.
- *close*: Función que cierra la conexión a la base de datos.

La estructura de los repositorios sería la siguiente:

```
classdef [nombre_tabla]Repository
    properties (Access = public)
    end
    methods(Static)
        function [nombre_objeto] =
get[nombre_tabla]By[nombre_parametro]([nombre_parametro])
            connSql = sqlite('[directorio_bd]');
            sql = "select [nombres_columnas] from [nombre_tabla] where
[condiciones]";
            [nombre_objeto] = fetch(connSql, sql);
            close(connSql);
        end

        function create[nombre_objeto]([valores_columnas])
            colnames = {'[nombres_columnas]'};
            connSql = sqlite('[directorio_bd]');
            insert(connSql, "[nombre_tabla]", colnames, ...
                {[valores_columnas]});
            close(connSql);
        end

        function update[nombre_objeto] ([valores_columnas])
            connSql = sqlite('[directorio_bd]');
```

```

sql = "update [nombre_tabla] set [nombres_columnas] =
[valores_columnas] where [condiciones]";
execute(connSql, sql);
close(connSql);
end

function delete[nombre_objeto] ([id_objeto])
connSql = sqlite('[directorio_bd]');
execute(connSql, "delete from [nombre_tabla] where [id_objeto]");
close(connSql);
end
end
end

```

### 5.3. Gestión de archivos.

Teniendo en cuenta que parte de la funcionalidad es la subida de documentos de pruebas de los pacientes y la revisión de los resultados obtenidos en las pruebas realizadas en la misma, se hace necesaria la gestión de dichos archivos. Con la intención de obtener un mejor rendimiento en la aplicación, ya que almacenar archivos en la B.D. puede generar sobrecarga [12], se ha decidido almacenar los distintos ficheros en el sistema de archivos del ordenador del usuario dentro de la carpeta que alberga el proyecto. Se ha definido un sistema de archivos en árbol como se puede observar en la Ilustración 10.



Ilustración 10. Esquema del sistema de archivos de la aplicación.

### 5.4. Parámetros acústicos.

Basándonos en el estudio de la encuesta realizada (Anexo 1) y diversos artículos científicos [3] [13] [11], se seleccionaron para su implementación los parámetros acústicos que se consideraron como los más utilizados o útiles.

El cálculo de los distintos parámetros se implementó en la clase *SignalAnalysis*, para facilitar la llamada a las distintas funciones desde la aplicación.

### 5.4.1. Shimmer.

El Shimmer relativo expresa la media de la diferencia absoluta entre las amplitudes de los periodos consecutivos dividida por la amplitud media. Calculándose este utilizando la Ecuación 1 del Shimmer relativo [14].

$$Shim = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A_i - A_{i+1}|}{\frac{1}{N} \sum_{i=1}^N A_i} * 100 \quad (1)$$

*Ecuación 1. Cálculo del Shimmer relativo.*

Para su implementación se analizaron los periodos glóticos de la señal de audio, siendo la duración de un periodo glótico 8.3 ms. Tras ello se almacenaron aquellos valores absolutos que se encontraran por encima del 70% de la media, con la intención de eliminar los picos no deseados (*outliers*). Una vez obtenidas las amplitudes, se procede a calcular el sumatorio de la norma de las amplitudes adyacentes y el sumatorio de todas las amplitudes. Después de esto, se calcula el Shimmer (%) obtenido en base a los sumatorios de las amplitudes [14].

A continuación, se muestra la implementación del método de cálculo del Shimmer, *calcular\_Shimmer*.

```
function shimmer = calcular_Shimmer(audio, fs, gp)
    arguments
        audio (:,1) double
        % frecuencia de muestreo
        fs (1,1) double {mustBeFinite} = 96e3;
        % duracion de un ciclo glotico
        gp (1,1) double {mustBeFinite} = 8.3e-3;
    end
    % Calculamos solo con aquellos que sean superiores al 70% de la
    % media
    meanAmplitude = mean(abs(audio), "all", "omitnan");
    minAmplitude = 0.7 * meanAmplitude;
    sampleLength = round(gp * fs);
    L = length(audio);
    if(mod(L/sampleLength, 1) ~= 0)
        x = ceil(L/sampleLength);
        valoresDeRelleno = (x * sampleLength) - L;
        audio (end + 1 : end + valoresDeRelleno) = NaN;
    end
    R = reshape(audio, sampleLength, []);
    R = max(R, [], 1);
    R = R(R >= minAmplitude);
    N = length(R);
    s = 0;
    a = 0;
    % calculamos los sumatorios de la norma y de las amplitudes
    for i = 1 : (N-1)
        s = s + norm(R(i) - R(i + 1));
        a = a + R(i);
    end
    % sumamos el ultimo valor de las amplitudes
    a = a + R(N);
    % calculamos el shimmer
    shimmer = ((1 / (N - 1)) * s) / ((1/N) * a) * 100;
end
```

### 5.4.2. Jitter.

El Jitter relativo expresa la media de la diferencia absoluta entre los periodos fundamentales consecutivos dividida por la media de los periodos fundamentales. Calculándose este utilizando la Ecuación 2 para Jitter relativo [14].

$$Jitt = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i+1}|}{\frac{1}{N} \sum_{i=1}^N T_i} * 100 \quad (2)$$

*Ecuación 2. Cálculo del Jitter relativo.*

Para su implementación se calculan en primera instancia las frecuencias fundamentales del audio utilizando ventanas del periodo de un ciclo glótico. Tras ello, y al igual que en el cálculo del Shimmer, se almacenaron sólo aquellos que se encontraran por encima del 70% de la media, con la intención de eliminar los picos de frecuencias no deseadas. Una vez obtenidas las amplitudes, se procede a calcular el sumatorio de la norma de los periodos fundamentales adyacentes y el sumatorio de los periodos. Después de esto, se calcula el Jitter relativo (%) obtenido en base a los sumatorios de los periodos [14].

A continuación, se muestra la implementación del método de cálculo del Jitter, *calcular\_Jitter*.

```
function jitter = calcular_Jitter(audio, fs, gp)
    arguments
        audio(:,1) double
        % frecuencia de muestreo
        fs(1,1) double {mustBeFinite} = 96e3;
        % duración de un ciclo glótico
        gp(1,1) double {mustBeFinite} = 8.3e-3;
    end
    % Calculamos el tamaño de la ventana en función del tiempo de un
    % ciclo glótico y la frecuencia de muestreo del audio
    L = length(audio);
    wLength = round(L / (gp * fs));
    ovLength = round(wLength * 0.807);
    % Calculamos el pitch del audio
    f0 = pitch(audio,fs,Method="SRH", ...
        WindowLength=wLength, ...
        OverlapLength=ovLength);
    % Calculamos solo con aquellos que sean superiores al 70% de la
    % media
    meanf0 = mean(f0, "all", "omitnan");
    f0 = f0(f0 >= 0.7*meanf0);
    N = length(f0);
    j = 0;
    t = 0;
    for a = 1 : (N-1)
        j = j + norm((1/f0(a)) - (1/f0(a + 1)));
        t = t + 1/f0(a);
    end
    t = t + f0(N);
    jitter = ((1 / (N - 1) * j)/((1/N)*t))*100;
end
```

### 5.4.3. Dysphonia severity index (DSI).

Se trata de una combinación ponderada de cuatro valores, que permiten una mejor predicción de la percepción auditiva de la gravedad de la disfonía [13]. Se calcula según la Ecuación 3.

$$DSI = 0.13 \cdot MPT (s) + 0.0053 \cdot F_0High(Hz) - 0.26 \cdot I_{low}(dB) - 1.18 \cdot Jitter (\%) + 12.4 \quad (3)$$

Ecuación 3. Cálculo del DSI.

Siendo:

- $MPT (s)$ : Tiempo máximo de fonación del paciente. Este parámetro se obtiene automáticamente de la anamnesis del paciente en caso de haberse rellenado ese campo, en caso contrario se obtendrá de la longitud del propio audio.
- $F_0High(Hz)$ : Frecuencia fundamental más alta.
- $I_{low}(dB)$ : Intensidad de la voz más baja dentro de la muestra de audio analizada.
- $Jitter(\%)$ : Porcentaje de *Jitter* obtenido en el análisis de la muestra de voz.

#### 5.4.4. Signal-to-noise ratio (SNR).

El SNR se trata de una medida de la cantidad de ruido presente en una señal de audio, definida como la razón entre intensidad de ruido y señal de voz, según la Ecuación 4.

$$SNR_{dB} = 10 \log \left( \frac{I_{voz}}{I_{ruido}} \right) \quad (4)$$

Ecuación 4. Cálculo del SNR en dB.

Teniendo en cuenta que la aplicación solo consta de una señal de entrada de audio y que por lo tanto no es posible una separación de ambas señales, por lo que nos atenemos a la definición de la SNR basada en la relación entre la media aritmética y la desviación estándar (Ecuación 5) [15].

$$SNR = \frac{\mu}{\sigma} = \frac{media}{desviación\ estandar} \quad (5)$$

Ecuación 5. Cálculo del SNR en base a la media y la desviación estándar.

Para ello obtenemos la media y la desviación estándar de los valores de intensidad del audio en dB y calculamos la razón entre ellos.

A continuación, se muestra la implementación del método de cálculo del SNR.

```

% Obtenemos la media
u = mean(audiodB, "omitnan");

% obtenemos la desviacion estandar
sd = std(audiodB, "omitnan");

% calculamos el SNR
snr = u / sd;

```

#### 5.4.5. Normalized noise energy (NNE).

Para el cálculo del NNE, una vez obtenido el SNR, sabemos que el NNE es igual a la potencia del ruido, tal y como se muestra en la Ecuación 6.

$$SNR = \frac{P_S}{P_N} = \frac{P_S}{NNE} \quad (6)$$

Ecuación 6. Cálculo del SNR en base a las potencias de señal de audio y señal de ruido

Teniendo en cuenta que la aplicación solo consta de una señal de entrada de audio y que, por lo tanto,  $P_S$  se trata de la suma de la señal de voz y la señal de ruido ( $P_{SN}$ ), el SNR se calcularía según la Ecuación 7.

$$SNR = \frac{P_{SN} - NNE}{NNE} \quad (7)$$

*Ecuación 7. Cálculo del SNR para una única señal haciendo uso del NNE.*

Podemos despejar NNE, obteniendo la Ecuación 8.

$$NNE = \frac{P_{SN}}{SNR - 1} \quad (8)$$

*Ecuación 8. Cálculo del NNE en base al SNR de la señal.*

Lo que calculamos de la siguiente manera en Matlab:

```
% Obtenemos la potencia del audio
pRMS = rms(audio)^2;
powbp = bandpower(audio, fs, [0 fs/2]);

% Calculamos el SNR en funcion de los dB
Psnr = 10^(snr/10);

nne = powbp/ (Psnr - 1);
```

#### 5.4.6. Harmonic-to-noise ratio (HNR).

Para calcular el HNR convencionalmente es necesario disponer tanto de la señal de voz como de la señal de ruido. Puesto que no disponemos de la señal de ruido aislada, podemos calcular el HNR haciendo una estimación del ruido. Para ello, usando la función de autocorrelación [16], obtenemos el máximo pico en la función, que corresponde al período fundamental de la señal ( $T_0$ ) y siendo este el inverso de la frecuencia fundamental, y calculamos la razón entre el valor en el pico y la media sin el pico (Ecuación 9).

$$HNR = \frac{R_{xx}[T_0]}{R_{xx}[0] - R_{xx}[T_0]} \quad (9)$$

*Ecuación 9. Cálculo del HNR utilizando la función de autocorrelación.*

Definiéndose HNR en decibelios según la Ecuación 10.

$$HNR_{dB} = 10\log(HNR) \quad (10)$$

*Ecuación 10. Cálculo del HNR en dB.*

A continuación, se muestra la implementación del método de cálculo del HNR.

```
% calculamos el máximo
audio = audio - mean(audio);
maximo = max(abs(audio));

% normalizamos entre -1 y 1
if maximo <= 0
    hnr = 0;
    return;
end
```

```

audio = audio / maximo;

% Calculamos la autocorrelacion y mantenemos los indices
% positivos
corr = xcorr(audio, audio);
corr = corr((length(corr)/2) + 1 : end);

% Buscamos los picos
[~, locs] = findpeaks(corr);

if isempty(locs)
    hnr = 0;
    return;
end

% localizamos el primer pico
primer_pico = locs(1);

% obtenemos el valor correspondiente a T0
[~, T0] = max(corr(primer_pico:end));
T0 = T0 + primer_pico;

%calculamos el HNR
hnr = 10.0 * log10(corr(T0) / (max(corr) - corr(T0)));

```

## 5.4.7. Estimación de los formantes.

### 5.4.7.1. Orden de predicción.

Para calcular la estimación de los formantes hemos de tener primero en cuenta el orden de predicción, siendo este el número de coeficientes de la codificación predictiva lineal (LPC). Este ha de ser el doble que el número de formantes ( $N_F$ ) o picos que se desea analizar. Por lo tanto, podemos estimar que el número de coeficientes ( $N_C$ ) según la Ecuación 11.

$$N_C = 2N_F + 2 \quad (11)$$

*Ecuación 11. Cálculo del número de coeficientes.*

Según los estudios de Gunnar Fant [17] los intervalos a los que ocurren los formantes están determinados por la longitud del tracto vocal (L) y por la velocidad del sonido (c). Por lo que podemos calcular dicho intervalo haciendo uso de la Ecuación 12.

$$\text{Intervalo (Hz)} = \frac{c}{2L} \quad (12)$$

*Ecuación 12. Cálculo del intervalo de aparición de formantes.*

Esto nos da que  $N_F$  es la razón de la frecuencia de Nyquist ( $f_{Nyquist}$ ) y el intervalo de aparición de los formantes, tal y como se observa en la Ecuación 13.

$$N_F = \frac{f_{Nyquist}}{\text{Intervalo}} \quad (13)$$

*Ecuación 13. Cálculo del número de formantes.*

Siendo la media del tracto vocal en hombres de 17.5cm y de mujeres 12.75cm, calculamos el número de coeficientes en Matlab de la siguiente forma.

```
if (gender == "M")
    L = 17.5;
else
    L = 12.75;
end

interHz = C / (2 * L);

%calculamos el número de formantes
nForm = round((fs/2) / interHz);

%calculamos el número de coeficientes que se esperan
nCoef = nForm * 2 + 2;
```

El valor del campo de género ha de ser especificado por el usuario en la anamnesis, en caso contrario el valor por defecto será el masculino.

#### 5.4.7.2. *Estimación de formantes con coeficientes LPC.*

Una vez obtenido el orden de predicción podemos calcular la estimación de las frecuencias de los formantes utilizando la LPC. Esto se hace mediante el cálculo de las raíces del polinomio de predicción [18].

Para ello, en primer lugar, tenemos que enventanar el audio utilizando una ventana de *Hamming*.

Tras ello, aplicamos un filtro de preénfasis paso alto de todos los polos, sobre la señal de audio enventanada. Y haciendo uso del orden de predicción de los coeficientes obtenemos los coeficientes LPC mediante el uso de la función *lpc* de Matlab.

Habiendo obtenido los coeficientes LPC, calculamos las raíces del polinomio, que tendrán lugar en pares complejos conjugados, por lo que mantenemos únicamente las que tienen parte imaginaria. A continuación, calcularemos los ángulos correspondientes a dichas raíces, y convirtiendo las frecuencias angulares a hercios, calculamos los anchos de banda de los formantes; siendo estos, representados mediante la distancia de los ceros desde el círculo unidad.

Finalmente, y ya obtenidos los anchos de banda de los distintos formantes, filtramos aquellos en los que su valor no se encuentre entre 90 y 400Hz.

A continuación, se muestra la implementación del método de estimación de los formantes.

```
% enventanamos el audio utilizando una ventana de hamming
L = length(audio);
xn = audio.*hamming(L);

%aplicamos filtro de preénfasis
preemph = [1 0.63];
xn = filter(1,preemph,xn);

% calculamos el número de coeficientes que se estima que se van
% a encontrar
nCoef = SignalAnalysis.calculate_number_coefficients(gender, fs);

%calculamos las raíces del polinomio que devuelve lpc
A = lpc(xn, nCoef);
rts = roots(A);

% determinamos los ángulos de las raíces imaginarias
rts = rts(imag(rts)>=0);
```

```

angz = atan2(imag(rts),real(rts));
%convertimos a Hz los ángulos
[freccs,locs] = sort(angz.*(fs/(2*pi)));
%Calculamos los anchos de banda de los formantes
bw = -1/2*(fs/(2*pi))*log(abs(rts(locs)));

n = 1;
for i = 1:length(freccs)
    if (freccs(i) > 90 && bw(i) < 400)
        formantes(n) = frqs(i);
        n = n+1;
    end
end
end

```

Tras su cálculo nos quedaremos con los 3 primeros formantes estimados, puesto que son los de mayor utilidad.

## 5.5. Gráficas.

### 5.5.1. Espectrogramas.

Para la generación de los formantes se ha utilizado la función de Matlab *spectrogram* haciendo uso de sus parámetros opcionales con la intención de obtener y definir mejor el espectrograma para el uso clínico para especialistas en voz. Por ello se ha definido el método *plotAudioSpectrogram*, mostrado más adelante, con la intención de eliminar código redundante. En este método se han tenido en cuenta dos tipos de espectrogramas útiles en el análisis acústico de la voz: espectrograma de banda ancha y el espectrograma de banda estrecha.

La diferencia entre el espectrograma de banda ancha y el de banda estrecha, es el tamaño de la ventana de muestreo que se utiliza, lo que hará que el espectrograma de banda estrecha tenga una resolución frecuencial mejor y que el de banda ancha tenga mejor resolución temporal [19]. Definiendo las frecuencias de las ventanas como:

- 45Hz: Frecuencia de la ventana de banda estrecha (0.3s).
- 250Hz: Frecuencia de la ventana de banda ancha ( $4 \cdot 10^{-3}$  s).

```

if(bandSize == "Narrow")
    % 45Hz
    tSize = 0.03;
else
    % 250Hz
    tSize = 4e-3;
end

windowSize = round(tSize * fs);
% definimos el tamaño del overlap como el 8% de la ventana
overlap = round(windowSize * 0.08);
% generamos el espectrograma
f1 = figure;
spectrogram(audio, windowSize, overlap, fs, fs, 'yaxis')
colormap jet
ylim([0 8])
xlabel('Tiempo (s)')
ylabel('Frecuencia (kHz)')

```

### 5.5.2. FFT.

La gráfica de la FFT se calcula haciendo uso de la función *fft* de Matlab. Para ello, primero se ha de obtener el tamaño de la FFT (*nfft*) para evitar *aliasing* por ello se define el *nfft* como el doble de  $2^L$  siendo L la longitud del audio en el próximo valor de la potencia de 2.

A continuación, se obtiene el rango de frecuencias que se van a mostrar en el gráfico. Por último, se calcula la FFT y se llama a la función *plot* de Matlab para mostrar únicamente las frecuencias positivas.

A continuación, se muestra la implementación del método de visualización de la FFT.

```
% Obtenemos la longitud del audio
lenX = length(audio);

% calculamos el tamaño de la fft para evitar aliasing
nfft = 2^nextpow2(lenX);

% calculamos las frecuencias de la fft
freqFft = (0 : (nfft/2 - 1)) * (fs/nfft);
Xs = fft(audio, nfft);

% plot de las frecuencias positivas
f2 = figure;
plot(freqFft, abs(Xs(1:nfft/2)))
xlabel('f(Hz)')
ylabel('|X(f)|')
xlim([0 5000])
ylim([0 2000])
```

### 5.5.3. Cepstrum.

Para la visualización del Cepstrum se ha hecho uso de la función *rcpes*, que devuelve los valores reales del Cepstrum de un audio. Una vez obtenido este, se calcula el tiempo de duración del audio usando el método *ObtenerTiempoMaximoFonacion*, para definir así los tiempos de la gráfica.

A continuación, se llama a la función *plot* de Matlab para mostrar la gráfica, y calcular un intervalo temporal en el que poder visualizar los picos.

La implementación en código se puede observar en la Tabla 5.

```
% calculamos los valores reales del cepstrum
c = rceps(audio);
% calculamos la longitud del audio en segundos
mpt = SignalAnalysis.ObtenerTiempoMaximoFonacion(audio, fs);
ts = mpt/length(c);
% definimos los tiempos de la gráfica
t = ts: ts : mpt;
f3 = figure;
% plot del cepstrum
plot(t,c)
xlabel('Tiempo (s)')
ylabel('Amplitud')
% visualizamos intervalo donde habrá picos
xlim([mpt /2 ((mpt /2) + 0.03)])
ylim([-4e-3 4e-3])
```

Tabla 5. Implementación de la visualización del cepstrum.

## 5.6. Interfaz de usuario.

Puesto que, actualmente, el catálogo de componentes de la librería del *App Designer* de Matlab es reducido y limitado a nivel de posicionamiento y modificación de apariencia, se decidió utilizar, casi en su totalidad, archivos HTML embebidos en el componente de muestra de HTML de Matlab.

Para ello se han definido dos tipos de fuentes de HTML en la aplicación: páginas estáticas y páginas dinámicas.

### 5.6.1. Páginas estáticas.

Para las pestañas de la aplicación de “Datos del paciente”, “Anamnesis” y “Nueva prueba”, y las páginas de “Inicio de sesión” y “Creación de nueva cuenta”, se decidió definir unas páginas HTML estáticas que se mostrarían a través del componente de muestra de HTML de Matlab. Estas se comunican con la aplicación de Matlab a través del intercambio de ficheros JSON en los eventos de *setup* y *DataChanged* tanto en el apartado de scripts de JavaScript de los ficheros HTML, como en el archivo *mlapp* de Matlab.

La comunicación se realiza, por lo general, de la siguiente forma:

- Primero, al entrar en la página que muestra el HTML, ya sea a través del cambio de pestaña o de la interacción con otro componente, se le indica al componente de muestra de HTML el *HTMLSource*, lo que sería una cadena con el directorio en el que se encuentra dicho archivo.
- Tras ello, se procedería a recuperar los datos a mostrar en la pantalla de la B.D. o de los objetos de la sesión, para pasárselos con formato de *struct* o tabla al componente de MATLAB en el parámetro *Data*, como se muestra en la Tabla 6.

```
app.[Nombre de la pantalla].Data = app.[objeto que almacena los
                               datos];
```

Tabla 6. Ejemplo de definición de datos del componente HTML en MATLAB.

- A continuación, se dispara el evento de “*DataChanged*”, implementado en el *setup* de la página HTML, que recibiría los datos aportados al parámetro *Data* y se introducirían los datos, si los hubiese, dentro de los valores de los controles HTML para su visualización. Un ejemplo de implementación se puede observar en la Tabla 7.

```
function setup(htmlComponent) {
    htmlComponent.addEventListener("DataChanged", function(event) {
        document.getElementById("[Id del control]").value =
            htmlComponent.Data.[nombre del dato];
    })
}
```

Tabla 7. Ejemplo de implementación de evento *DataChanged* en JavaScript.

- En el caso de ser necesario que la página HTML tuviera comunicación bidireccional con el componente de Matlab, como se observa en la Tabla 8, se pueden definir eventos en los controles HTML que modifican el parámetro *Data* del componente de visualización disparando el *callback DataChanged* en Matlab. Esto permitiría recuperar los datos, definidos en JSON en la página HTML, en Matlab como *structs*.

```
function setup(htmlComponent) {
    const btn = document.getElementById("[Id del control]");
    btn.addEventListener('click', evento);
    function evento(e) {
        htmlComponent.Data = {
            '[nombre del dato]':
                document.getElementById("[Id del control]").value,
            ...
        };
    }
}
```

Tabla 8. Ejemplo de implementación del envío de datos a MATLAB con JavaScript.

### 5.6.2. Páginas dinámicas.

Debido a las limitaciones de los componentes de visualización de HTML de Matlab, algunas funcionalidades, como modificar dinámicamente el archivo fuente de las imágenes/audios o la generación automática o borrado de controles, se decidió generar de manera dinámica algunas de las páginas de la aplicación. Se puede observar un ejemplo de lo descrito en la Tabla 9.

Para ello, se decidió utilizar los comandos *fopen*, *fclose*, *fprintf* y *drawnow* de la librería de funciones de Matlab, lo que permite sobrescribir el archivo al que apunta el parámetro *HTMLSource* de los componentes de Matlab y recargarlo para mostrar los nuevos datos.

```
function montarPagina(app)
    % Borramos el campo HTMLSource
    app.HTML.HTMLSource = '';
    % Abrimos el archivo HTML en modo escritura para sobrescribirlo
    fid = fopen('pagina.html','w');
    % Definimos nuestra pagina HTML con cadenas de string
    html = "<!DOCTYPE html><html lang='es'>" + ...
    ...
    "</html>";
    % Escribimos la cadena en el archivo
    fprintf(fid, html);
    % Cerramos la edición del archivo
    fclose(fid);
    % Recargamos el archivo en MATLAB
    Drawnow
    % Definimos de nuevo el campo HTMLSource para que apunte a la
    % página creada
    app.HTML.HTMLSource = 'pagina.html';
end
```

Tabla 9. Ejemplo de generación dinámica de página HTML con MATLAB.

La generación dinámica de archivos HTML se ha utilizado, mayoritariamente, para la carga de componentes ligados a listas dinámicas, como los documentos o las pruebas, que pueden ser eliminadas y añadidas por el usuario desde la misma página en la que se visualizan.

## 6. La aplicación a nivel de usuario.

### 6.1. Requerimientos.

Para instalar la aplicación es necesario disponer de un ordenador con, al menos, 16GB de memoria RAM. Esto se debe a las limitaciones que impone Matlab y la memoria que reserva para los datos.

Para realizar de las pruebas y grabar con la máxima calidad se elegirá una sala aislada y con buena acústica. Se recomienda el uso de un micrófono de condensador de respuesta plana, con patrón polar omnidireccional. A la hora de realizar las grabaciones, se recomienda posicionar el micrófono a 30 cm de la boca [1].

Se recomienda la instalación únicamente en dispositivos con sistema operativo Windows, puesto que la visualización de documentos está implementada con el método *winopen*, lo que selecciona el programa por defecto al abrir los archivos en ordenadores Windows.

También se requerirá la instalación de Matlab en su versión 2022b.

## 6.2. Tutorial.

Al abrir la aplicación se nos mostrará una ventana de inicio de sesión, como la que se aprecia en la Ilustración 11. Para poder iniciar sesión es necesario estar registrado.

Para registrarse se deberá acceder a la pantalla de registro pulsando sobre el botón “Crear nueva cuenta”. Una vez clicado se nos redirigirá automáticamente.



Ilustración 11. Acceso a la aplicación.

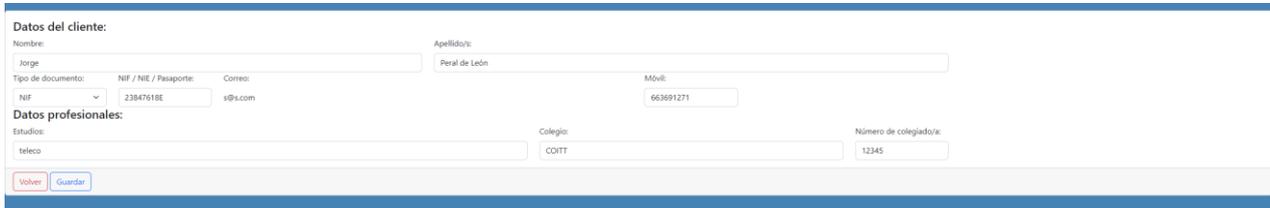
A continuación, se nos pedirán los datos de registro los cuales se conforman de un correo electrónico, una contraseña y un token de acceso, que limitará el acceso a la aplicación a las personas autorizadas por parte de la organización, como se observa en la Ilustración 12. Se requiere una confirmación de la contraseña para verificar que no se han introducido caracteres erróneos ni se ha producido algún fallo al escribirla, en caso de no pasar la verificación se nos instará a ponerla de nuevo. Únicamente se permite la creación de una cuenta por correo electrónico.



Ilustración 12. Ventana de creación nueva cuenta.

Una vez generada la cuenta, se permitirá iniciar sesión utilizando los datos de acceso personales que se han indicado en el paso anterior. Si es la primera vez que se inicia sesión, se mostrará la pantalla de datos del cliente que se observa en la Ilustración 13, y no se permitirá el

uso de la aplicación hasta rellenar los datos básicos (el nombre). En esta pantalla se podrán rellenar los datos del usuario principal de la cuenta. Se puede acceder a ella en cualquier momento a través del icono de usuario de la esquina superior izquierda que se muestra en la Ilustración 15. El botón de “Guardar” guardará los datos del usuario y el botón “Volver” nos redirigirá a la ventana general de la Ilustración 14.



*Ilustración 13. Pantalla de datos del usuario.*

Una vez añadidos los datos de usuario, se mostrará la pantalla general, en ella solo se mostrará el nombre de usuario, el acceso a los datos del usuario, un botón para añadir pacientes y la lista de pacientes actuales, como se muestra en la Ilustración 14. Al seleccionar un paciente se visualizarán las pestañas de datos del paciente seleccionado que se van a poder observar y modificar (Ilustración 15).



*Ilustración 14. Pantalla General.*

En la pestaña de datos del paciente se podrán indicar tanto los datos del cliente (nombre, apellidos, NIF, etc.) como los datos de la mutua aseguradora en caso de tenerlos. Para guardar los datos introducidos es necesario confirmarlo clicando sobre el botón de “Guardar”.

Ilustración 15. Datos del paciente y de la mutua.

Al seleccionar la segunda pestaña se mostrará la anamnesis del paciente. Este apartado permitirá realizar una anamnesis completa del paciente, rellenando los datos necesarios de su historial como observaciones subjetivas del profesional como la escala GRBASI. Algunos de los parámetros vienen acompañados de explicaciones, a modo de *tooltip*, que se mostrarán al colocar el ratón sobre los iconos de interrogante (Ilustración 20). Al igual que en las demás ventanas, será necesario clicar sobre el botón de “Guardar” para que se guarden los datos de la anamnesis. Los apartados campos de la anamnesis se han separado en los siguientes apartados: Datos personales, antecedentes ORL (Ilustración 16), antecedentes personales y familiares (Ilustración 17), datos médicos de relevancia (Ilustración 18), aspectos relacionados con la vida diaria, descripción del problema y sus causas (Ilustración 19), balance vocal y GRBASI (Ilustración 20).

Si se desea que se tenga en cuenta el tiempo máximo de fonación (TMF) en el cálculo del DSI, ha de ser especificado en esta pestaña, en caso contrario se tomara como valor por defecto la duración del audio de la prueba, lo que puede invalidar el cálculo de este parámetro. Al indicar también el tiempo máximo de espiración (TME) se calculará automáticamente el cociente fonorespiratorio.

Ilustración 16. Anamnesis: Datos personales y Antecedentes ORL.

### Antecedentes personales y familiares:

Antecedentes ORL y respiratorios:

Afonías de repetición, Rinitis alérgica, tiroidectomía total.

Antecedentes familiares con problemas de voz:

Abuela operada de quiste pseudoseroso de cuerda vocal

Antecedentes alérgicos:

Madre con alergia a los ácaros del polvo y al pelo de gato.

Ilustración 17. Anamnesis: Antecedentes personales y familiares.

### Datos médicos de relevancia:

Alergias:

No, presenta rinitis no alérgica inespecífica.

Hábitos tóxicos (tabaco, alcohol, drogas, etc.):

NO.

Medicamentos:

Eutirox 100mcgr, Avamys spray, Xazal, Famotidina 40mg, Fluoxetina y degrass 100mg (ocasionalmente para dormir)

Reflujo/ácidez:

Si, realiza dieta específica para prevención de reflujo y toma medicación (famotidina)

Intervenciones quirúrgicas (con intubación o sin intubación):

tiroidectomía con intubación (intervenida en dos ocasiones) y cordectomía con intubación (4 cordales)

Trastornos psico-emocionales:

Ansiedad y estrés (toma medicación)

Otros antecedentes de relevancia:

Actualmente presenta problemas gastrointestinales que en ocasiones le afectan a todas las mucosas, incluida la mucosa laríngea, provocando deshidratación y aumentando las molestias al hablar.

Ilustración 18. Anamnesis: Datos médicos de relevancia.

### Aspectos relacionados con la vida diaria:

Entorno laboral: ☺

Trabaja con niños y en ocasiones grita inadecuadamente o fuerza la voz al hablar.

Entorno doméstico:

Vive con sus padres y todos tienen una voz fuerte.

Aficiones: ☺

Realiza clases de canto y va al gimnasio a hacer pesas (activación de la válvula laríngea)

Relación horaria y descanso:

descanso inadecuado y horarios alterados

Estudios:

Ya finalizados

Hidratación:

Escasa. Nula hidratación directa en pliegues vocales

### Descripción del problema y sus causas:

Antigüedad de la disfonía:

Desde hace más o menos 1 año

Forma de inicio: ☺

Gradual/progresivo

¿Afonías?: ☺

Sí, cuando fuerza la voz, grita o sale de fiesta.

Cambios durante el día (mejor voz: mañana, mediodía o noche):

La voz está mejor por la mañana y va empeorando a lo largo del día (aparece fatiga vocal por la noche)

¿Mejora con reposo vocal?:

Sí

¿Sensación de forzar? ¿cuándo?:

Sí, siempre que habla

¿Sensación de ahogo / falta de aire?:

No.

¿Voz enganchada?:

No.

¿Peso en la voz?:

No.

¿Velo en la voz?:

No.

¿Rasca, pica o hay irritación?:

Sí, necesidad de carraspear a menudo.

¿Nota que la voz sale del cuello?:

Sí.

Observaciones relevantes: ☺

Siente que su voz es más grave que antes y que la gente de su alrededor lo nota. Se cansa mucho al hablar y se queda afónica con frecuencia. Le cuesta alzar la voz sin realizar esfuerzo (probablemente debido a la incompetencia en el cierre glótico y al escape de aire)

Ilustración 19. Anamnesis: Aspectos relacionados con la vida diaria y descripción del problema y sus causas.

### Balance vocal:

TME:  TMF:  Cociente fonorespiratorio:

### GRBAS:

Grado de disfonía (G):  Aspereza (R):  Soplo (B):  Asténica (A):  Tensión (S):  Inestabilidad (I):

### Valoración perceptiva del timbre:

Riqueza de armónicos:  Dominancia de armónicos:  Nasalidad:  Registro resonancial:

Ilustración 20. Anamnesis: Balance Vocal y GRBASI.

Si se selecciona sobre la pestaña de documentos adjuntos se podrá observar la lista de los documentos que hayan sido adjuntados en la ficha del paciente, cada uno de ellos con el nombre del documento adjuntado, un botón para visualizar los documentos (botón con icono de ojo) y un botón para eliminar el documento (botón con icono de papelera) (Ilustración 21). Para añadir un nuevo documento adjunto se deberá seleccionar el que se desee subir tras clicar sobre el botón de “Subir nuevo documento”. Puesto que la aplicación actualmente está implementada para su uso local, no existe un límite de formato o de tamaño para los ficheros que se deseen adjuntar.

Para la visualización de cualquier documento adjuntado, la aplicación abrirá el programa establecido por defecto por el sistema operativo.

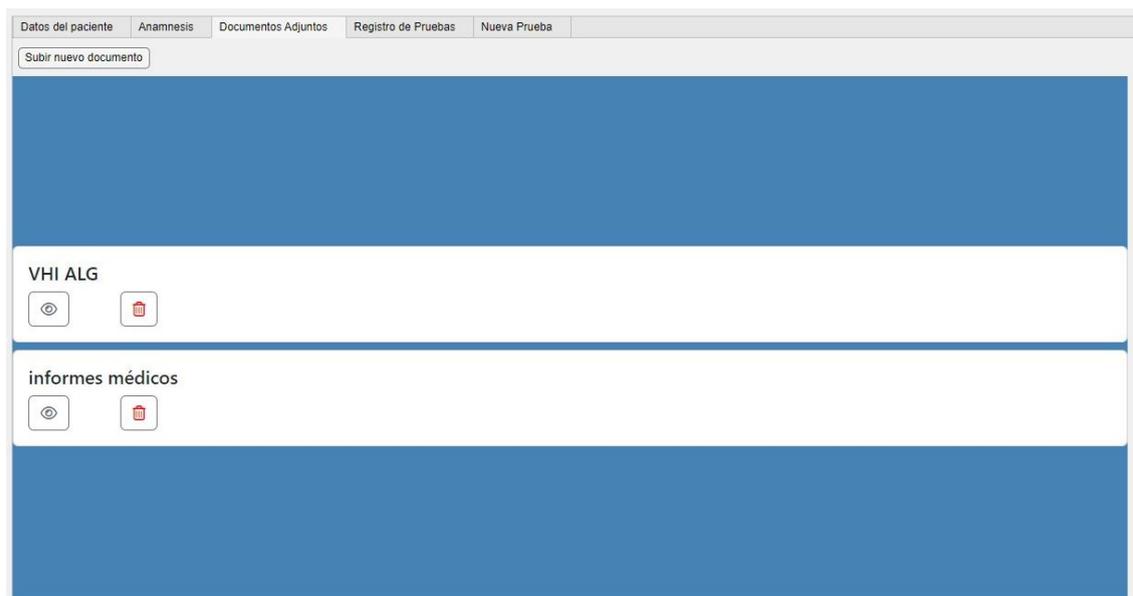


Ilustración 21. Ventana de documentos adjuntos.

Al navegar a la cuarta pestaña se mostrarán las pruebas acústicas realizadas al paciente en la aplicación con la descripción proporcionada por el usuario antes de realizar dicha prueba, la fecha de realización de esta y un reproductor que permitirá escuchar el audio analizado (Ilustración 22). Al igual que en la pantalla de documentos, se permite la eliminación y la visualización de las pruebas. La visualización de las pruebas mostrará la ventana con los datos obtenidos al igual que en la Ilustración 24.

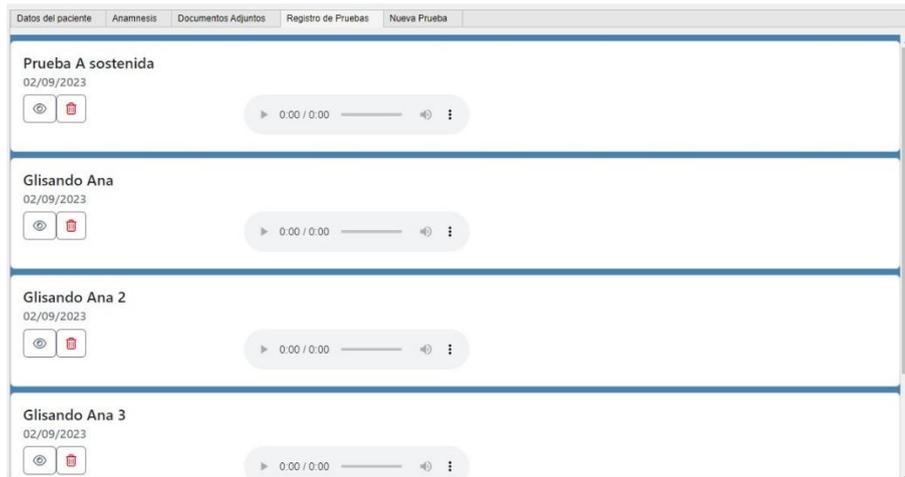


Ilustración 22. Ventana de registro de pruebas.

En la última pestaña, se muestra la pantalla de “Nueva prueba”, en esta se grabarán los audios para su posterior análisis acústico (Ilustración 23). Se puede definir la descripción de la prueba, si no se hace antes de realizar la grabación se asignará una por defecto (p.ej. “Prueba 1”) en función del número de pruebas realizadas anteriormente. También se aportarán recomendaciones de uso en función del objetivo que tenga la prueba.

Para realizar una prueba, se ha de clicar sobre el botón de grabar (icono círculo rojo), grabar el audio durante el tiempo deseado, y clicar sobre el botón de detener (icono cuadrado rojo), lo que automáticamente comenzará el análisis de la señal grabada.

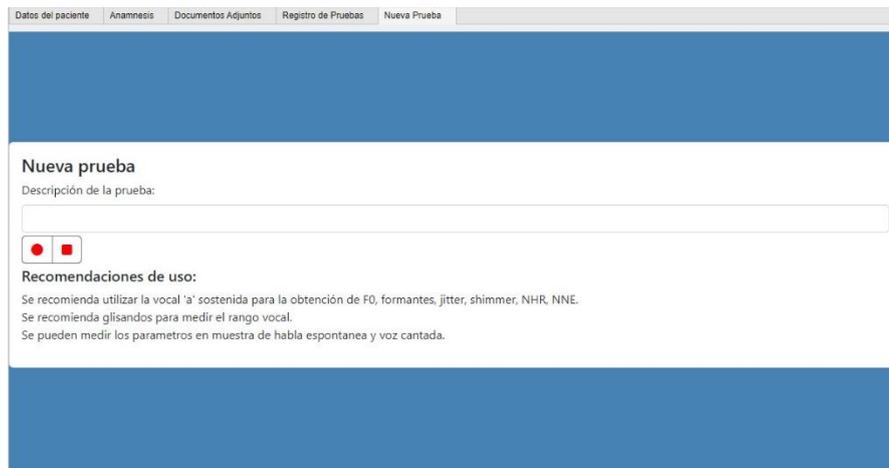


Ilustración 23. Ventana de nueva prueba.

Una vez finalizados los cálculos de los datos necesarios para la evaluación, se mostrarán en la página de resultados como se puede observar en la Ilustración 24. También se abrirán de forma simultánea las distintas gráficas en ventanas independientes, para que el usuario pueda, si lo desea, hacer una evaluación exhaustiva sobre las mismas, permitiéndose en estas ventanas mover la imagen, hacer *zoom* y seleccionar puntos del gráfico para observar los valores que hay.

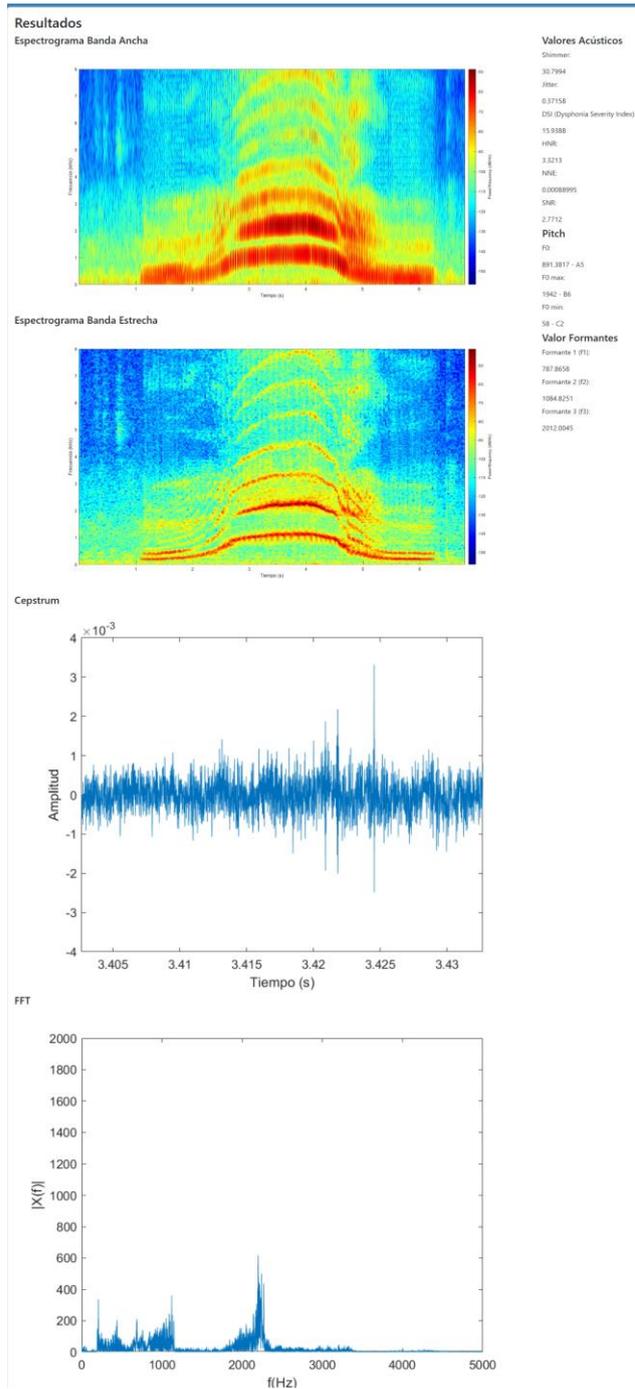


Ilustración 24 Ventana de muestra de resultados de las pruebas

## 7. Resultados.

Tras la finalización de la aplicación, era necesario comprobar el correcto desempeño de todas las funcionalidades y corroborar que la integración de todos los componentes que la conforman se hubiese llevado a cabo adecuadamente. Con este objetivo en mente, se realizaron pruebas de humo para verificar que, en el conjunto de la aplicación, no hubiera defectos evidentes que interrumpiesen su funcionamiento. Tras esto, se ejecutaron múltiples pruebas para comprobar el correcto funcionamiento del cálculo de parámetros acústicos, y se solventaron los problemas encontrados.

Por último, y con el afán de verificar su verdadera utilidad, se contactó con una logopeda especialista en voz, para que realizara las pruebas de aceptación del usuario, utilizando la aplicación con alguno de sus pacientes. Tras instalarle el software necesario para ejecutar las pruebas, dispuso de varios días para testar la aplicación. Al finalizar el plazo de pruebas, se le pasó una encuesta de satisfacción (Anexo 3. Encuesta de satisfacción, para evaluar la aceptación del usuario.) y se le pidió que compartiera algunos de los datos obtenidos (Anexo 4. Resultados obtenidos en la aplicación con un paciente real en una consulta de logopedia.).

## 8. Conclusiones y futuras líneas de trabajo.

En primer lugar, se puede considerar que se ha alcanzado el objetivo principal de este proyecto, ya que se ha conseguido desarrollar una aplicación informática que permitirá, a los logopedas especialistas en voz, la realización de pruebas acústicas mediante las muestras de audio de los pacientes para llevar a cabo evaluaciones objetivas.

La herramienta responde a las necesidades expresadas por un grupo de 20 logopedas a través de una encuesta inicial. Se han analizado los resultados de dicha encuesta, y se ha observado la necesidad de tener un programa que les permitiese recoger la información de los parámetros vocales más importantes para la práctica clínica en pacientes que presentan patologías vocales. Por ello, se ha implementado el cálculo de los parámetros más significativos obtenidos de la encuesta, lo que permitirá a los sanitarios especialistas en voz el uso de una única aplicación en lugar de tener que utilizar varias de funcionalidad reducida.

El programa les facilitará también la gestión de los datos de sus clientes proporcionándoles una plataforma donde poder guardar todas las pruebas realizadas, implementar la anamnesis inicial y adjuntar documentos que aporte el paciente.

Cabe destacar, teniendo en cuenta la repercusión económica que podría tener este proyecto, que si estos profesionales pudieran adquirir una sola aplicación que realizase todas las funciones necesarias, ahorrarían mucho en costes, ya que actualmente las aplicaciones que se encuentran en el mercado son caras y la mayoría no realizan todas estas funciones. Además, la aplicación podría ampliarse y actualizarse realizando mejoras continuas.

Por otra parte, el proyecto me ha permitido ampliar mis conocimientos de procesado digital de la señal, ya que para su realización ha sido necesario revisar las distintas formas de cálculo de los parámetros acústicos relevantes para la práctica sanitaria. También he desarrollado mis conocimientos de programación, ya que he buscado la mejor forma de llevarlo a cabo.

El proyecto es una primera versión de lo que podría llegar a ser la herramienta y en un futuro podría ampliarse para introducir nuevas funcionalidades y mejoras. Durante la realización de este proyecto se ha observado que, debido a la alta carga computacional que genera el análisis acústico de audios enteros y con la finalidad de mejorar la velocidad de procesamiento y la calidad de las gráficas ofrecidas, sería interesante ampliar la funcionalidad de la aplicación, permitiendo al usuario seleccionar los segmentos del audio que desea analizar. Esto permitiría realizar análisis de muestras más pequeñas pero obtenidas con una mayor frecuencia de muestreo, lo que daría una mayor definición a los espectrogramas y evitaría la interferencia de los *outliers* y del ruido en

el cálculo de los distintos parámetros, además de permitir el uso del programa en sistemas con menos recursos.

Por otra parte, con respecto a los modelos matemáticos y funcionalidades implementadas, sería interesante revisar el código para implementar mejoras de rendimiento. Además de plantear el desarrollo de la aplicación en otras tecnologías como .NET y Angular, que permitirían enlazar con Matlab para seguir realizando los cálculos, pero abrirían un mayor abanico de posibilidades de desarrollo de funcionalidades, como la descarga de informes en formato PDF.

La creciente utilización de los sistemas basados en inteligencia artificial pone de manifiesto su gran utilidad en una gran multitud de campos. Como futura línea de ampliación se plantea la generación de modelos *Deep Learning* o la utilización de modelos preentrenados como GPT-4, que permitan en base a los datos recogidos en esta aplicación como *datasets* como el *VOICED Database PhysioNet* [20] para sugerir a los especialistas posibles diagnósticos en base a los resultados obtenidos en la prueba.

## Bibliografía.

- [1] Sebastián Castro-Tighe , Gonzalo Inostroza-Moreno, «Variabilidad de micrófonos utilizados para,» *Revista de Investigación e Innovación*, pp. 93-101, 19 10 2020.
- [2] F. D. A. M. I. M. L. G. G. B. A. J. B. Susanibar, Trastornos del habla de los fundamentos a la evaluación., Madrid: EOS, 2016.
- [3] I. Bustos Sánchez, La voz : técnica vocal para la rehabilitación. 2a. edición revisada y ampliada, Barcelona: Editorial Paidotribo, 2012.
- [4] A. Quilis, «Tratado de fonología y fonética española,» Madrid, Editorial Gredos, 1997, pp. 13-550.
- [5] Graciela Oliva, Andreu Sauca, Josep M. Vila-Rovira, «Herramientas para la evaluación logopédica de la voz,» 2013.
- [6] F. D.-M. J. P. G.-L. I. M.-M. A. C.-M. M. M.-G. C. & M.-C. P. Núñez-Batalla, «El espectrograma de banda estrecha como ayuda para el aprendizaje del método GRABS de análisis perceptual de la disfonía.,» de *Acta Otorrinolaringológica Española*, 2012.
- [7] E. J. I. S. M. H. & H. T. Yamauchi, «Perceptual evaluation of pathological voice quality: A comparative análisis between the RASATI and GRBASI scales,» *Logopedics Phoniatrics Vocology*, pp. 121-128, 2010.
- [8] F. Nuñez-Batalla, J. P. Díaz- Molina, I. García-López, A. Moreno-Méndez, M. Costales-Marcos, C. Moreno-Galindo y P. & Martínez Camblor, «El espectrograma de banda estrecha como ayuda para el aprendizaje del método GRABS de análisis perceptual ed la disfonía.,» de *Acta Otorrinolaringología Española*, 2012.
- [9] Batalla, Faustino Núñez, Carlos Suárez Nieto, «Espectrografía clínica de la voz,» Universidad de Oviedo, 1999.
- [10] O. R. A. Z. P. P. Z. L. K. S. D. E. G. M. M. B. M. D. R. Tom Bäckström, «speechprocessingbook.aalto.fi,» [En línea]. Available: [https://speechprocessingbook.aalto.fi/Representations/Jitter\\_and\\_shimmer.html](https://speechprocessingbook.aalto.fi/Representations/Jitter_and_shimmer.html). [Último acceso: 17 04 2023].
- [11] C. J. H. M. Colton RH., Understanding Voice problems: A physiological perspective of diagnosis and treatment., Baltimore: Lippincott Williams & Wilkins, 2ºed, 1996.
- [12] D. Armstrong, Pro ASP.NET 2.0 Website Programming. Capítulo: Uploading Files, Appress, 2005.
- [13] F. L. e. a. Wuyts, «The Dysphonia Severity Index: An Objective Measure of Vocal Quality Based on a Multiparameter Approach,» *Journal of Speech, Language, and Hearing Research*, pp. 796-809, 2000.

- [14] João Paulo Teixeira, André Gonçalves, «Algorithm for Jitter and Shimmer Measurement in Pathologic Voices,» *Procedia Computer Science*, pp. 271-279, 2016.
- [15] Naim Mansour, Marton Marschall, Tobias May, Adam Westermann, Torsten Dau, «A method for realistic, conversational signal-to-noise ratio estimation,» *J Acoust Soc Am*, p. 1559–1566, 2021.
- [16] P. Boersma, «Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound,» *IFA Proceedings*, p. 1993, 97-110.
- [17] G. Fant, *Speech acoustics and phonetics: Selected writings*, Springer Science & Business Media, 2004.
- [18] R. C. y. F. M. Snell, «Formant location from LPC analysis data,» *IEEE® Transactions on Speech and Audio Processing*, pp. 129-134, 1993.
- [19] Ronald J. Baken, Robert F. Orlikoff, *Clinical Measurement of Speech and Voice*, Singular Thomson Learning, 2000.
- [20] U. Cesari, G. De Pietro, E. Marciano, C. Niri, G. Sannino, and L. Verde, «A new database of healthy and pathological voices,» *Computers & Electrical Engineering*, pp. 310-321, 5 2018.

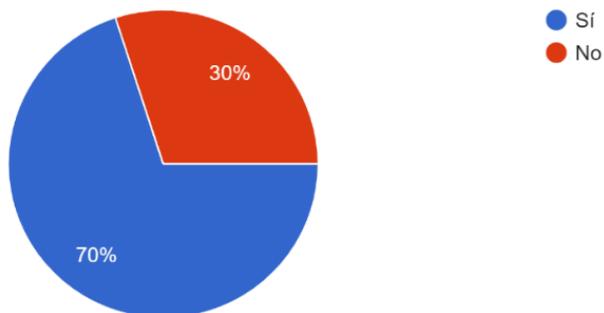
## Anexo 1. Resultados de la encuesta a logopedas para evaluar la oferta actual de software de análisis acústico de la voz.

Esta encuesta tiene como objetivo conocer la opinión general de los logopedas sobre la oferta actual de software clínico para el análisis de voz.

A continuación, se muestran los resultados de esta.

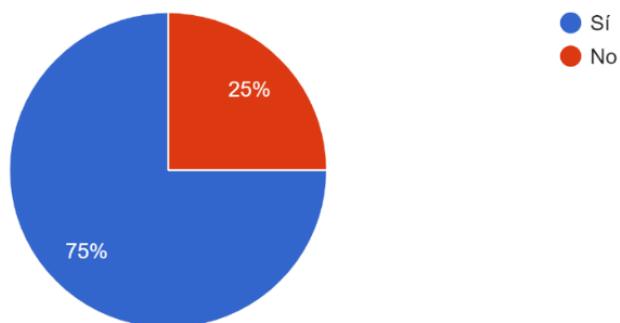
¿Eres logopeda especialista en voz?

20 responses



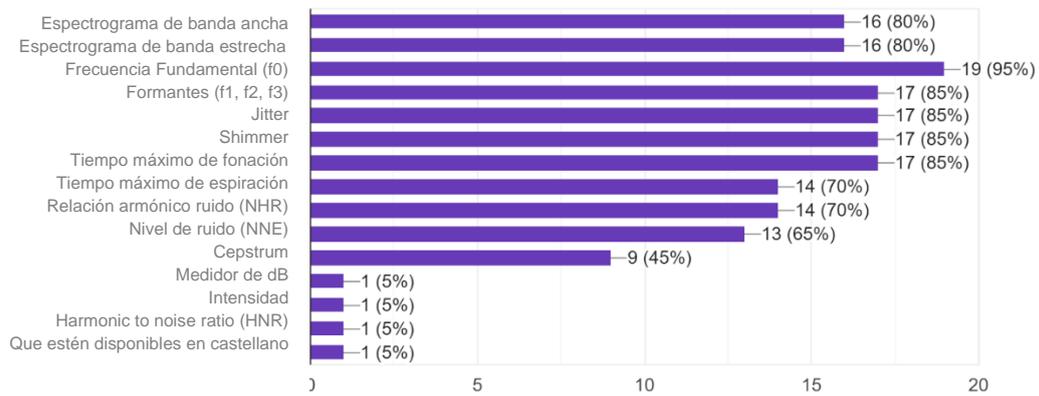
¿Utilizas software de análisis acústico en tu práctica clínica?

20 responses



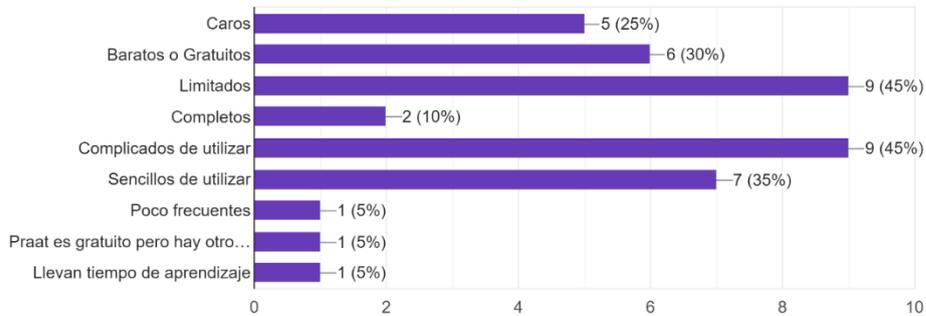
¿Qué parámetros consideras que deberían aparecer en estos programas?

20 respuestas



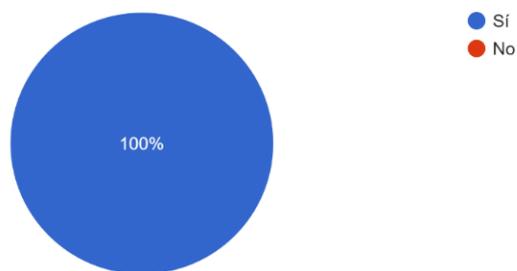
¿Consideras que estos programas son...?

20 respuestas



¿Te parecería útil que la propia aplicación tuviera un apartado para poder guardar las historias clínicas de tus pacientes?

20 respuestas



## Anexo 2. Descripción de los ficheros del proyecto.

A continuación, se describen los archivos implementados en la realización de este proyecto. Estos se han clasificado en cuatro grupos: Ficheros de diseño de la interfaz de usuario, ficheros para la generación y funcionamiento de la base de datos, ficheros de procesado de audio, y ficheros de clases repositorio.

Ficheros de diseño de la interfaz de usuario:

- *anamnesisPaciente.html*: Fichero HTML que conformará la pantalla de anamnesis del paciente.
- *datosPaciente.html*: Fichero HTML que conformará la pantalla de datos del paciente.
- *datosPrueba.html*: Fichero HTML sobre el que se escribirá para generar, de forma dinámica, la pantalla de resultado de las pruebas acústicas.
- *datosUsuario.html*: Fichero HTML que conformará la pantalla de datos del usuario.
- *docsPaciente.html*: Fichero HTML sobre el que se escribirá para generar, de forma dinámica, la pantalla de documentos adjuntados del paciente.
- *newUser.html*: Fichero HTML que conformará la pantalla de nuevo usuario, para la creación de un nuevo perfil.
- *nuevaPrueba.html*: Fichero HTML que conformará la pantalla de nuevas pruebas, que permitirá al usuario realizar grabaciones para su posterior análisis acústico.
- *pruebasPaciente.html*: Fichero HTML sobre el que se escribirá para generar, de forma dinámica, la pantalla de pruebas realizadas al paciente.
- *userLogin.html*: Fichero HTML que conformará la pantalla de inicio de sesión de la aplicación.
- *Css*: Directorio que alberga los paquetes descargados para el uso de la librería de estilos de Bootstrap.
- *fontawesome-free-6.2.0-web*: Directorio que alberga los paquetes descargados para el uso de la librería de iconos de FontAwesome en su versión 6.2.
- *js*: Directorio que alberga los paquetes descargados para el uso de la librería JQuery.
- *iconsuser.png*: Imagen utilizada como botón de acceso para el perfil de usuario.
- *LogoVaap.png*: Logo de la aplicación que se mostrará en la pantalla de inicio de sesión.
- *botoncircularplus.png*: Icono de "+" que se muestra en el botón de añadir pacientes.
- *ResultadosPrueba.mlapp*: Aplicación de Matlab que permite la visualización de resultados de las pruebas acústicas en una ventana *popup*.
- *vaap.mlapp*: Este fichero es el principal del proyecto, y es el que se ejecutará para operar con la interfaz de usuario.

Ficheros para la generación y funcionamiento de la base de datos:

- *createTableAnamnesis.sql*: Fichero con el comando SQL para la creación de la tabla *Anamnesis*.
- *createTableDocuments.sql*: Fichero con el comando SQL para la creación de la tabla *Documents*, donde se almacenarán los datos de los documentos adjuntos de los pacientes.
- *createTableNotas.sql*: Fichero con el comando SQL para la creación de la tabla *Notas*. En esta se encontrarán las notas del rango vocal con las frecuencias correspondientes.
- *createTablePatients.sql*: Fichero con el comando SQL para la creación de la tabla *Patients*, donde se guardarán todos los datos principales del paciente.
- *createTablePruebas.sql*: Fichero con el comando SQL para la creación de la tabla *Pruebas*. Esta albergará los datos de las pruebas acústicas realizadas a los pacientes.
- *createTableUsers.sql*: Fichero con el comando SQL para la creación de la tabla *Users*. En ella se incluirán los datos de inicio de sesión y de identificación del usuario.
- *test.db*: Fichero principal de la base de datos, donde se escriben todos los datos introducidos en ella.

Ficheros de procesado de audio:

- *AudioRecorder.m*: Fichero clase de Matlab, en el que están implementados los métodos que permiten la grabación de audios.
- *SignalAnalysis.m*: Fichero principal de análisis de parámetros acústicos, en el que se encuentran todos los métodos necesarios para realizar las pruebas de los audios seleccionados.

Ficheros de clases repositorio:

- *AnamnesisRepository.m*: Repositorio donde se implementan los métodos de consulta y edición de datos para la tabla de base de datos *Anamnesis*.
- *DocumentsRepository.m*: Repositorio donde se implementan los métodos de consulta y edición de datos para la tabla de base de datos *Documents*.
- *PatientRepository.m*: Repositorio donde se implementan los métodos de consulta y edición de datos para la tabla de base de datos *Pacients*.
- *PruebasRepository.m*: Repositorio donde se implementan los métodos de consulta y edición de datos para la tabla de base de datos *Pruebas*.
- *UserRepository.m*: Repositorio donde se implementan los métodos de consulta y edición de datos para la tabla de base de datos *Users*.
- *Util.m*: Repositorio donde se implementan los métodos genéricos que pudieran ser de utilidad en cualquier proceso de la aplicación.

### **Anexo 3. Encuesta de satisfacción, para evaluar la aceptación del usuario.**

A continuación, se muestran las respuestas de la encuesta de satisfacción con la aplicación que se le realizó a la logopeda que accedió a utilizarla en algunas de sus consultas.

#### **¿En cuantos pacientes lo has utilizado?**

*“Con 8 pacientes”*

#### **¿Qué patologías tenían?**

*“Disfonía hiperfuncional, disfonía espasmódica, quiste pseudoseroso de cuerda vocal, nódulos vocales y pólipo de cuerda vocal.”*

#### **¿Te ha resultado útil el apartado de la anamnesis? ¿añadirías alguna pregunta más?**

*“Me ha parecido súper útil, las preguntas que aparecen están muy bien escogidas y son ideales para los pacientes de voz. Me ha encantado que se incluya el GRBAS y que se pueda poner información perceptiva de la voz. También me ha gustado mucho que se incluyan el TME, el TMF y el cociente fonorrespiratorio, son muy útiles en la práctica clínica y no se suelen incluir en ningún programa.”*

#### **¿Te ha parecido fácil de utilizar el programa?**

*“Sí, la verdad es que el acceso y la creación de la cuenta es muy sencillo, los campos son fáciles de cumplimentar y guardar, la grabación de audio y la obtención del espectrograma y demás parámetros se realiza de forma rápida y sencilla, en otros programas de análisis acústico que he utilizado, es complicado grabar y guardar los datos, en algunos hay que realizar las grabaciones con otros programas para después adjuntar el audio. Todo ese proceso es largo y lioso y poder tenerlo todo guardado en un solo programa y que además grabe los audios y los*

analice es fantástico. Al final los profesionales necesitamos herramientas efectivas, rápidas y completas.”

**¿Te ha parecido útil el poder adjuntar los documentos en la propia aplicación?**

*“Es genial, porque permite tenerlo todo clasificado y ordenador, agiliza la búsqueda de archivos y evita tener que guardar todo por separado y perder tiempo teniendo que buscarlo.”*

**¿Qué opinas de tener todas las fichas de los pacientes en un único programa?**

*“Me parece muy útil, es mucho más fácil organizar la información, poder acceder a ella rápidamente y tenerlo todo a mano. Y me ha gustado mucho que dentro de cada paciente se pueda acceder a los análisis acústicos y poder grabar audios diferentes dependiendo de lo que tengamos que evaluar, además de poder añadir audios nuevos a lo largo de la terapia para realizar diferentes análisis para observar los progresos de los pacientes.”*

**¿Qué cosas positivas le encuentras a la aplicación?**

*“Todo lo mencionado anteriormente. Creo que es útil, rápida y completa. Fácil de usar y concisa. Las preguntas son adecuadas para los pacientes de voz, añade evaluación perceptual y los análisis acústicos que arroja son muy completos, justo lo que se necesita para hacer una buena evaluación y diagnóstico vocal. Me ha resultado muy funcional que el espectrograma se muestre en color, en vez de en escala de grises, ya que es mucho más sencillo de interpretar.”*

**¿Qué cosas cambiarías o añadirías?**

*“Añadiría algunas funcionalidades como poder seleccionar y recortar trozos de audio en el espectrograma. Poder descargar los datos y guardarlos en pdf y generar informes automáticos con la información de la anamnesis, o en su defecto, añadir una pestaña en la que aparezcan los apartados que incluye un informe para poderlo rellenar directamente y sacar un archivo aparte para poderlo imprimir o enviar al paciente y a otros profesionales. También se podría realizar con un diseño más atractivo.”*

**¿Qué opinas sobre los parámetros que se evalúan?**

*“Están muy bien seleccionados, aparecen todos los que suelo usar e incluso alguno más, como el cepstrum, que es muy útil pero que no se calcula en casi ninguna aplicación de las que hay en el mercado. Me ha encantado que se incluya la nota del piano equivalente a la frecuencia fundamental, ahorra mucho trabajo y lo realiza de manera objetiva, porque muchos logopedas buscamos la F0 de oído con el piano mientras escuchamos la voz del paciente, cosa que a veces es imprecisa, sobre todo en voces que están muy afectadas.”*

**¿Recomendarías esta aplicación a otros profesionales?**

*“Por supuesto, me ha parecido genial y en los días en los que la he utilizado me ha agilizado mucho las evaluaciones de los pacientes, ya que siempre tengo que evaluar las cosas por separado usando varias aplicaciones o programas, puesto que la mayoría de los programas no evalúan todos los parámetros que necesitamos ni incluyen nada para guardar información del paciente. Por un buen precio, incluso me plantearía adquirirla para mi clínica.”*

## Anexo 4. Resultados obtenidos en la aplicación con un paciente real en una consulta de logopedia.

A continuación, se adjunta la anamnesis y resultados del análisis acústico de un paciente real.

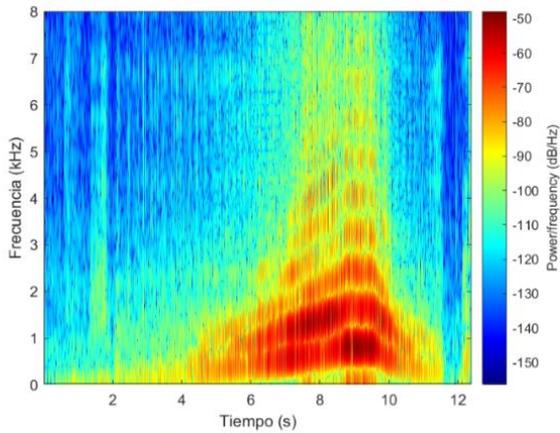
Datos del paciente	Anamnesis	Documentos Adjuntos	Registro de Pruebas	Nueva Prueba
<b>Datos personales:</b>				
Fecha de nacimiento:	Edad:	Género:	Fecha del examen:	Derivado por:
15/11/2002	20	Muje	09/06/2023	Profesor de canto y otorrino
Motivo de consulta:				
M.C acude a nuestra consulta debido a que hace unos meses comenzó a detectar una pérdida de agudos en su registro vocal. Es soprano y está finalizando sus estudios en el conservatorio. Cuando acababa las clases se sentía muy mal de la voz y poco a poco cada vez lograba realizar menos notas agudas, notó un agravamiento de la voz, pérdida en la resonancia vocal y pérdida de brillo de la voz. Acudió al otorrino y le detectaron nódulos vocales bilaterales. La derivan a logopedia para rehabilitar los nódulos vocales y aprender técnica vocal para evitar una recidiva en el futuro.				
<b>Antecedentes ORL:</b>				
Patología vocal:				
nódulos vocales bilaterales asociados a hipertensión, cierre de musculatura antero-posterior e hiato en forma de reloj de arena.				
Método de exploración:				
laríngeoscopia				
Fecha de exploración:				
17/05/2023				
<b>Antecedentes personales y familiares:</b>				
Antecedentes ORL y respiratorios:				
Otitis de repetición y dislocación de ATM.				

Datos del paciente	Anamnesis	Documentos Adjuntos	Registro de Pruebas	Nueva Prueba
No.				
Antecedentes alérgicos:				
No.				
<b>Datos médicos de relevancia:</b>				
Alergias:				
Ninguna.				
Hábitos tóxicos (tabaco, alcohol, drogas, etc.):				
No.				
Medicamentos:				
Aerius spray, omeprazol 40mg				
Reflujo/ácidez:				
Si.				
Intervenciones quirúrgicas (con intubación o sin intubación):				
Si, le realizaron una cirugía del hombro con intubación				
Trastornos psico-emocionales:				
Síndrome de Asperger con ansiedad asociada.				
Otros antecedentes de relevancia:				
Tiene una enfermedad rara llamada edhler danlos que afecta a la síntesis del colágeno y por lo tanto, a la elasticidad de los ligamentos. No podemos trabajar con ella ejercicios de movilidad corporal, ni de cambios posturales bruscos. La terapia manual no le beneficia y las técnicas invasivas de punción con electroestimulación tampoco.				

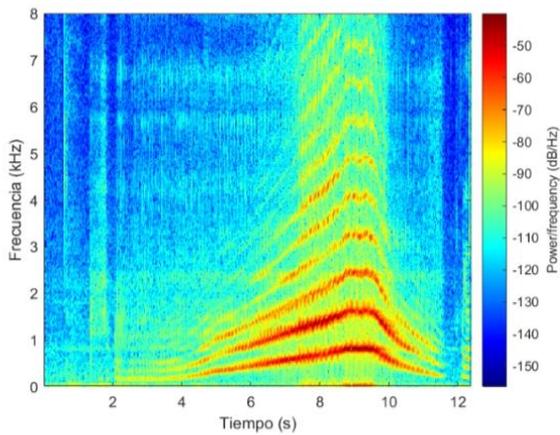


## Resultados

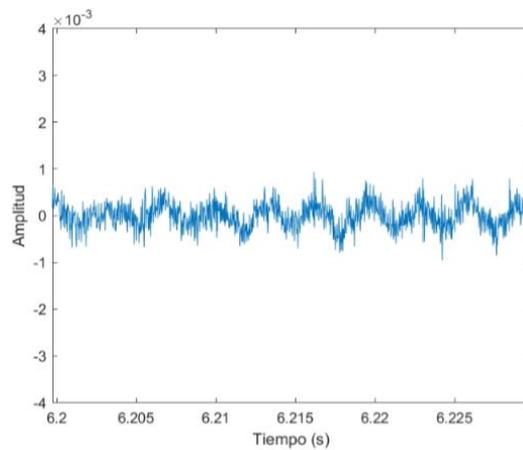
### Espectrograma Banda Ancha



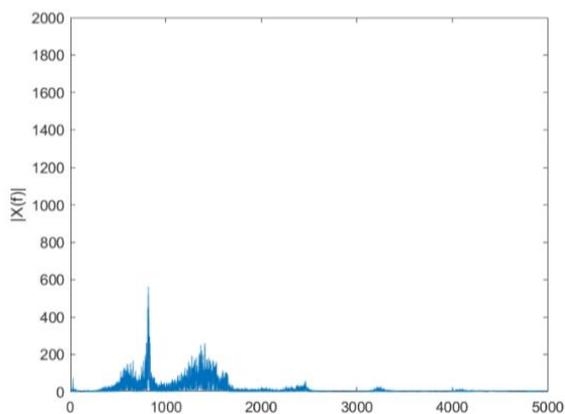
### Espectrograma Banda Estrecha



### Cepstrum



### FFT



### Valores Acústicos

Shimmer:

12.3263

Jitter:

0.23566

DSI  
(Dysphonia  
Severity  
Index):

27.6861

HNR:

0.51091

NNE:

0.0015176

SNR:

1.78

### Pitch

F0:

1032.2867 - C

F0 max:

1981 - B6

F0 min:

64 - C2

### Valor Formantes

Formante 1  
(f1):

1384.9313

Formante 2  
(f2):

2386.6351

Formante 3  
(f3):

3458.0346