



**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS
AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL**

UNIDAD ZACATENCO
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA
SECCIÓN DE BIOELECTRÓNICA

Evaluación objetiva de la calidad vocal mediante la presión del aire durante la fonación.

Tesis que presenta

Perla Elizabeth Jimarez Rocha

**Para obtener el Grado de
Maestra en Ciencias
En la Especialidad de**

Ingeniería Eléctrica

Director de la Tesis

Dr. Arturo Minor Martínez

Ciudad de México

Mayo 2018

Dedico este trabajo a mis amados padres y a mis hermanos.

Gracias por estar ahí para mí.

AGRADECIMIENTOS

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca brindada durante estos dos años de la maestría así como al Centro De Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV- IPN).

Agradezco especialmente a mi asesor y director de tesis el Dr. Arturo Minor Martínez por haberme aceptado en su laboratorio, por su paciencia, las enseñanzas por los consejos para hacerme una mejor persona, gracias por todo el apoyo brindado durante todo este tiempo desde que llegue al laboratorio 5 así como la amistad, muchas gracias por todo.

La valiosa colaboración del Servicio de Foniatría del Instituto Nacional de Rehabilitación (INR) en especial al Dr. Víctor Valadez Jiménez por haberme aceptado en su laboratorio durante la investigación, por el apoyo brindado y enseñanza así como la Dra. Carolina Daniela Tejada Franco gracias.

El trato amable y la actitud servicial del personal administrativo y el grandioso trabajo que realizan sus secretarias en la sección de Bioelectrónica, agradezco a Georgina Olmos Villareal a la señora María Luisa Lorán Núñez en especial a Patricia Peña Rivero, gracias por su amistad y el cariño que recibí muchas gracias por todo Paty.

Gracias por la ayuda brindada en el desarrollo del proyecto y la amistad de nuestros experimentados auxiliares técnicos de investigación del laboratorio 5 al Ing. Irwin Rosales García y el Ing. José Rodolfo Rojas Ortiz.

Gracias a mis amigos de laboratorio 5, Chio gracias por ser mi amiga por apoyarme en la estancia y después de esta gracias por el cariño y amistad, Elba, Rafita, Emmanuel y Rey muchas gracias por su apoyo, conocimiento y en especial el cariño que me dieron durante mi estancia así como en la maestría, la verdad es que la pasamos genial y compartimos muchas alegrías, Dalia, Erick y Lenin los quiero a todos y muchas gracias por todo, a todas las personas que tuve el placer de conocer en laboratorio 5 unas personas maravillosas sigan así y nunca sean apáticos jaja La lista es enorme de las personas que quiero agradecer por el apoyo recibido de familiares y amigos que hicieron agradable y alegre mi experiencia en el Cinvestav los mejores recuerdos.

Resumen

La producción de la voz es el mecanismo fundamental de la comunicación oral. Esta comunicación se puede ver perturbada por diversas patologías que afectan la producción vocálica. Dichos problemas se pueden clasificar en mecánicos o estructurales (por ejemplo quistes, tumoraciones y laringitis) y neurológicas (por ejemplo parálisis cordales y enfermedades de Parkinson).

El trabajo propone el desarrollo de un sistema que captura y registra señales de voz y analiza la calidad vocal (frecuencia fundamental, cruce por cero y energía), así como el cálculo de la presión subglótica [cm H₂O] durante la fonación sostenida de la vocal /a/; se desarrolla un sistema de registro, así como un sistema interactivo que registre información de la presión subglótica de manera intuitiva. En la ciudad de México se propone un grupo control de pacientes que presentan una patología foniátrica, se realizaron pruebas experimentales de las patologías vocales más comunes (Nódulos, Pólipos, Laringitis Irritativa, Disfonía ventricular, Cáncer laríngeo, Disfonía y disfagia) de forma no invasiva.

Uno de los mayores problemas en cuanto al desarrollo de nuevas terapias para el manejo y tratamiento de los trastornos de la voz es la dificultad de evaluar objetivamente los resultados de cada tratamiento. Los métodos objetivos se basan fundamentalmente en la evaluación acústica de la señal eléctrica de la voz (frecuencia, ruido, intensidad perturbaciones de amplitud, distribución armónica, timbre y energía), mientras que los métodos subjetivos, se basan en la evaluación perceptual y pueden ser realizados por el médico tratante. El trabajo consiste en registrar las patologías foniátricas con métodos objetivos ejercida durante una vocalización.

Abstract

The production of the voice is the fundamental mechanism of oral communication. This communication can be disturbed by various pathologies that affect vowel production. These problems can be classified as mechanical or structural (for example cysts, tumors and laryngitis) and neurological (for example, chordal paralysis and Parkinson's disease).

The work proposes the development of a system that captures and records voice signals and analyzes the vocal quality (fundamental frequency, zero crossing and energy), as well as the calculation of the subglottic pressure [cm H₂O] during the sustained phonation of the vowel /to/; a registration system is developed, as well as an interactive system that registers subglottic pressure information intuitively. In Mexico City, a control group of patients with a phoniatrics pathology was proposed. Experimental tests of the most common vocal pathologies (nodules, polyps, Irritant Laryngitis, Ventricular Dysphonia, Laryngeal Cancer, Dysphonia and dysphagia) were performed in a non-invasive manner.

One of the biggest problems in the development of new therapies for the management and treatment of voice disorders is the difficulty of objectively evaluating the results of each treatment. The objective methods are fundamentally based on the acoustic evaluation of the voice's electrical signal (frequency, noise, amplitude disturbances, harmonic distribution, timbre and energy), while the subjective methods are based on the perceptual evaluation and can be made by the treating doctor. The work consists of registering the pathologies with objective methods exercised during a vocalization.

Contenido

Capítulo 1. Introducción.....	1
1.1 Planteamiento del problema	2
1.2 Objetivo general	3
1.3 Objetivos específicos.....	3
1.4 Estado del arte	4
Capítulo 2. Antecedentes.....	6
2.1 Sistema fonatorio anatomía y fisiología laríngea	6
2.2 Mecanismos de la fonación.....	9
2.2.1Cuerdas vocales.....	9
2.2.2 Estructura cartilaginosa	10
2.3 Estructura muscular.....	11
2.4 La voz Humana	13
2.4.1 Producción de la voz	13
2.5 Calidad de la voz.....	15
2.6 Vocales	17
2.6.1 Producción de las vocales	17
2.6.2 Formas de evaluación en la fonación.....	18
2.7 Tiempo máximo de fonación (MPT).....	19
2.8 Presión subglótica	20
Capítulo 3. Desarrollo	22
3.1 Descripción del sistema de voz.	22
3.2 Instrumentos empleados	23
3.2.1 Dispositivo de registro electrónico de voz.	23
3.3 Cadena de medida del sistema propuesto.....	25
3.3.1 Frecuencia fundamental	25
3.4 Energía de la voz.	27
3.5 Cruce por cero de la señal de voz.....	27
3.6 Caracterización de sensores de presión.....	28
3.6.1 Sensor de presión.	28

3.7 Interfaz gráfica del sistema del sistema de registro.....	32
3.8 Desarrollo del sistema interactivo de voz.....	33
3.8.1 Programa de control del sistema interactivo.	33
3.8.2 Algoritmo de control de la aplicación.	37
3.9 Desarrollo de la aplicación.	39
Capítulo 4. Pruebas	44
4.1 Protocolo de medición del grupo control.	44
4.1.1 Criterios de inclusión	45
4.1.2 Criterios de eliminación	46
Capítulo 5. Resultados	47
5.1 Resultados gráficos de la calidad de la voz.	47
5.1.1 Resultados gráficos No patológicos foniatricos	50
5.1.2 Resultados gráficos patológicos foniatricos.	57
5.1.3 Resultados gráficos patológicos foniatricos del sistema interactivo.	63
5.1.4 Resultados gráficos No patológicos foniatricos del sistema interactivo.	64
Capítulo 6. Conclusiones y perspectivas.....	65
Anexos. Especificación general de los dispositivos usados	72
Anexo 1 Especificaciones del sensor de presión Honeywell.	72
Anexo 2 Formato de consentimiento informado y aviso de privacidad.....	72

Índice de figuras

Fig 1. Aparato respiratorio.	8
Fig 2. Aparato fonador	8
Fig 3. Cuerdas vocales.	9
Fig 4. Anatomía de la laringe.	10
Fig 5. Ciclo de vibración de los pliegues vocales	15
Fig 6. Fase de abertura	21
Fig 7. Fase de cierre	21
Fig 8. Diagrama a bloques de la solución propuesta del sistema de voz.....	22
Fig 9. Sistema de registro de la voz.	23
Fig 10. Sistema de acople con sensores de presión para el registro de la señal de voz.....	24
Fig 11. Instrumentación del sistema propuesto.	25
Fig 12. Sensor de presión marca Honeywell modelo (MPxx5004).....	28
Fig 13. Caracterización del sensor de presión.....	31
Fig 14. Interfaz gráfica para el registro de la voz.....	32
Fig 15. Diagrama de flujo general del sistema interactivo de voz	34
Fig 16. Inicialización del programa y lectura del puerto serial.	35
Fig 17. Diagrama de flujo inicial del juego interactivo.....	36
Fig 18. Algoritmo de ajuste de escala del grafico mediante la presión de fonación sostenida inicial. ..	37
Fig 19. Relación de posición[Px] vs presión[cm H ₂ O].....	39
Fig 20. Pantalla de inicio del sistema interactivo de registro del paciente	40
Fig 21. Ajuste de posición del control de mando (Flappy Bird).	41
Fig 22. Secuencia de las escenas de la aplicación interactiva	42
Fig 23. En la figura se muestra algunos de los usuarios que realizaron pruebas con el sistema correspondiente (registro e interactivo).....	45
Fig 24. Señal de voz de la vocal /a/.....	47
Fig 25. Ciclos de la señal de la vocal /a/.....	48
Fig 26. Energía de la vocal /a/.....	48
Fig 27. Cruce por cero de la vocal /a/.....	49
Fig 28. Frecuencia fundamental de la vocal /a/ "pitch".....	49
Fig 29. Presión subglótica (20.39 cm H ₂ O), paciente hombre de 46 años de edad.....	50
Fig 30. Presión subglótica (21.75 cm H ₂ O), paciente hombre de 42 años de edad.....	51
Fig 31. Presión subglótica (13.58 cm H ₂ O), paciente hombre de 52 años de edad.....	51
Fig 32. Presión subglótica (26.87 cm H ₂ O), paciente hombre de 29 años de edad.....	52
Fig 33. Presión subglótica (24.47 cm H ₂ O), paciente mujer de 37 años de edad.....	52
Fig 34. Presión subglótica (19.03 cm H ₂ O), paciente mujer de 46 años de edad.....	53
Fig 35. Presión subglótica (25.56 cm H ₂ O), paciente mujer de 29 años de edad.....	53
Fig 36. Presión subglótica (32.62 cm H ₂ O), paciente hombre de 32 años de edad.....	54
Fig 37. Presión subglótica (21.39 cm H ₂ O), paciente hombre de 37 años de edad.....	54

Fig 38. Presión subglótica (26.32 cm H ₂ O), paciente hombre de 34 años de edad.....	55
Fig 39. Presión subglótica (16.31 cm H ₂ O), paciente mujer de 29 años de edad.....	55
Fig 40. Presión subglótica (27.19 cm H ₂ O), paciente mujer de 33 años de edad.....	56
Fig 41. Presión subglótica (29.39 cm H ₂ O), paciente hombre de 28 años de edad.....	56
Fig 42. Presión subglótica (25.4 cm H ₂ O), paciente hombre de 51 años de edad que presenta una patología común (parálisis cordal del lado izquierdo cuerda vocal).	57
Fig 43. Presión subglótica (18 cm H ₂ O), paciente hombre de 47 años de edad que presenta una patología común (parálisis bilateral cordal).	57
Fig 44. Presión subglótica (35.69 cm H ₂ O), paciente mujer de 33 años de edad diagnosticada una patología común (Disfonía espasmódica).	58
Fig 45. Presión subglótica (5.37 cm H ₂ O), paciente hombre de 43 años de edad diagnosticado una patología común (Disfonía Ventricular).	58
Fig 46. Presión subglótica (7.61 cm H ₂ O), paciente hombre de 30 años de edad diagnosticado con una patología común (Distrofia muscular cordal).....	59
Fig 47. Presión subglótica (3.8 cm H ₂ O), paciente hombre de 36 años de edad presenta una patología común de (Distrofia muscular Miotónica).	59
Fig 48. Presión subglótica (9.70 cm H ₂ O), paciente mujer de 31 años de edad que presenta una patología común (Nódulos).....	60
Fig 49. Presión subglótica (36.5 cm H ₂ O), paciente hombre de 42 años de edad diagnosticado con una patología común (Nódulos).....	60
Fig 50. Presión subglótica (16.11 cm H ₂ O), paciente hombre de 32 años de edad que presenta una patología común (Hiperquinesia vocal).	61
Fig 51. Presión subglótica (11.05 cm H ₂ O), paciente hombre de 43 años de edad que presenta una patología común (Disartria).	61
Fig 52. Ajuste de presión máxima por medio de posición [pixel].....	63
Fig 53. Ajuste de posición de control de mando	64

Índice de tablas

Tabla 1. Músculos intrínsecos de la laringe.....	21
Tabla 2. Frecuencia fundamental en hombre y mujeres respecto a edades.....	23
Tabla 3. Tabla de conversión de presiones con relación a la salida de voltaje del sensor.....	38
Tabla 4. Cálculos para encontrar el voltaje real en relación a la presión en [cm H ₂ O].....	39

Capítulo 1. Introducción

Desde inicios de la civilización humana, la comunicación oral ha sido fundamental como medio de comunicación interpersonal. Esta característica fundamental y exclusiva de los seres humanos, se desarrolló partiendo de simples señas y sonidos hasta convertirse en sofisticado lenguaje capaz de transmitir información de manera rápida y eficiente.

Los tratamientos de fonación comúnmente utilizados son quirúrgicos y no quirúrgicos. Los métodos quirúrgicos suelen reservarse para los casos en los cuales las terapias de voz no han sido efectivas.

- Métodos quirúrgicos: tiroplastia de medialización, aducción de aritenoides y re inervación laríngea.

Los métodos no quirúrgicos están relacionados durante el tiempo de fonación, consiste en la cantidad de aire que una persona expulsa relacionando la respiración en cada proceso.

- Métodos no quirúrgico: técnicas de relajación muscular, técnica masticatoria, técnica de giro lingual, terapia de voz y bostezo fonatorio.

Uno de los mayores problemas en cuanto al desarrollo de nuevas terapias para el manejo y tratamiento de los trastornos de la voz, es la dificultad de evaluar objetivamente los resultados de cada tratamiento. La evaluación cualitativa describe el padecimiento inicial de acuerdo a los resultados que se obtienen a partir de los resultados de la calidad de voz. La evaluación cuantitativa es una serie de procesos empleados de acuerdo al método que se utiliza para la valoración de un diagnóstico [1]. El análisis comparativo entre distintos tratamientos es

importante porque permite diseñar equipos para aplicaciones específicas, que potencian el impacto en la sociedad. Al analizar un tratamiento se deben definir cuáles son sus alcances y limitaciones, que tan seguro es y qué grupo de pacientes pueden beneficiarse de este. El presente escrito trata del desarrollo de un dispositivo que permite evaluar la calidad de la voz mediante la producción de la presión del habla, en función de la eficiencia fonatorio a partir de análisis multivariados de la voz (frecuencia fundamental, espectro de amplitud, cruces por cero, energía y presión subglótica) durante ejercicios de fonación sostenida.

1.1 Planteamiento del problema

En México en la última década, la preocupación de los problemas relacionados con la voz y sus patologías se han considerado primordiales en el ámbito social de cada persona ya que las alteraciones de un parámetro vital es la incapacidad de adaptarse en el medio social y laboral. La disfonía tiene una prevalencia de incluso 47% en la población, es más frecuente en personas cuya voz es un instrumento de trabajo, tiene una estrecha relación en mujeres de 6 a 4 respecto a los hombres en quienes es más frecuente observar el abuso o mal uso vocal, lo cual es un factor de riesgos de lesiones principalmente estructurales de las cuerdas vocales. Afecta a cualquier grupo de edad, en niños es de 3.9% a 32.4% entre las edades de 8 a 14 años de edad. En este grupo el abuso vocal está relacionado con los gritos frecuentes por actividades recreativas [2], [3].

Los problemas referentes a la calidad de la voz en México, el inadecuado uso de la voz representa un tema importante para su abordaje. Es importante resaltar que solo se ha tomado en cuenta la calidad de la voz, además de las patologías que puede haber referente a la voz humana. Por otra parte, debe considerarse el porcentaje de niños que pueden padecer alguna alteración en la voz y como afectarían en un futuro el ámbito en el que se desarrollen, aunque actualmente, se encuentra en desarrollo técnicas más avanzadas de rehabilitación de la voz por métodos que no sean quirúrgicos para lograr una recuperación de la calidad de la voz.

1.2 Objetivo general

Desarrollar un sistema no invasivo que registre de manera objetiva la calidad vocal durante la fonación sostenida, usando variables específicas referentes a la voz (energía, densidad de cruces por cero, frecuencia fundamental, espectro de la señal de voz y presión subglótica) para el análisis clínico no invasivo de las patologías foniátricas más comunes (Nódulos, Pólipos, Laringitis Irritativa, Disfonía ventricular, Cáncer laríngeo, Disfonía y disfagia).

1.3 Objetivos específicos

Diseño y construcción de un dispositivo electrónico para el registro y análisis de la voz instrumentado con sensores de presión [kPa].

- Diseño y construcción de un dispositivo electrónico para registro y análisis de voz instrumentado con sensores de presión.
- Análisis de la señal en frecuencia y tiempo (frecuencia fundamental, número de cruces por cero, densidad de cruces por cero, espectro de amplitud, energía y presión subglótica) en pacientes (hombre y mujer) entre las edades de 27-50 años que presenten una patología común de la voz y no patológicos.
- Desarrollo de un sistema de rehabilitación interactivo que registre la presión subglótica [cm H₂O] durante la fonación sostenida.
- Desarrollo de una interfaz gráfica y un sistema interactivo no invasivo para registrar la presión subglótica en pacientes durante la fonación sostenida.
- Análisis y registro de las señales (voz) obtenidas en el sistema de voz interactivo durante una fonación constante.

1.4 Estado del arte

Existen sistemas de evaluación de la calidad de la voz, así como, la presión subglótica cuando se realizan ejercicios de fonación. A continuación se describen los sistemas más importantes para dicha evaluación hasta la fecha, así como referencias en las que se basaron esos sistemas para la fonación.

En 1966, Gunnar Fat presenta una forma de estudio en relación a la presión de la voz, grabó datos simultáneos de las presiones subglótica y supra glóticas y de la señal de voz ya que refiere a la frecuencia fundamental que es la que se calcula en un estado de intensidad e interpretación en variaciones aerodinámicas [4].

En 2006 se emplea una técnica no invasiva para estimar la presión subglótica sin interrumpir la fonación en donde el flujo del aire es parcialmente disperso introduciendo dos impedancias diferentes a través de una boquilla controlada con una válvula. Los cambios del flujo de aire se utilizaron para estimar la presión subglótica en usuarios sanos [5].

Speech Research Laboratory de la universidad de Syracuse fundada en 1966.- *Glottal Enterprises Syracuse 2015* es un sistema que conecta una interfaz de audio mediante una computadora, el procedimiento que se realiza a los usuarios es colocar electrodos en la zona del cuello en cada ala del cartílago tiroideo asegurando una fijación con una malla velcro alrededor del cuello, se le pide a los usuarios producir una vocal sostenida durante el tiempo de prueba estándar (15 segundos) observando así el comportamiento simultáneo de señales de voz, la señal se guarda en un archivo “.WAV “en un programa especializado en donde el usuario puede obtener ciertas características específicas de su voz [6].

En el 2015 se desarrolló el sistema *The Phonatory Aerodynamic System (PAS)* este mide flujo de aire, presión y otros parámetros relacionados con la producción de voz. El PAS utiliza un software fácil de usar, basado en protocolos, proporciona medida de presión incluyendo datos gráficos y cuantitativos para monitorear y reportar el desempeño del paciente en una vocalización, la medida de los parámetros aerodinámicos mejora la comprensión clínica en acústica incluyendo la capacidad vital, presión de aire, eficiencia de voicing (presión intraoral y resistencia laríngea) [7].

Capítulo 2. Antecedentes

En este capítulo se presentan los antecedentes que sirvieron para tener una mejor comprensión del método elegido en el desarrollo del este trabajo.

2.1 Sistema fonatorio anatomía y fisiología laríngea

La voz humana podemos estudiarla desde diferentes aspectos como medios de la comunicación y expresión o como una combinación de ambos. Podemos decir que la voz humana es producida por el aire espirado que después de una serie de modificaciones se convierte en palabras o canto. El aire que pasa entre las cuerdas vocales las hace vibrar produciendo un sonido, que será más agudo o más grave dependiendo la tensión a la cual están sometidas [8]. Este sonido se amplifica a su paso por las cavidades de resonancia, que están formadas por todas aquellas estructuras situadas por encima de las cuerdas vocales. Los resonadores principales son la boca o cavidad bucal, en la cual el sonido se articula por medio de la lengua y los labios convirtiéndose en lenguaje, en la cavidad nasal resonarán los sonidos, mediante una adecuación de estos mecanismos transforman este sonido en voz y principalmente en el lenguaje [9].

Desde el punto de vista anatómico tenemos que considerar que el hombre no posee ningún órgano que sea específicamente productor de sonidos. Para hablar intervienen elementos del aparato digestivo, respiratorio y una serie de grupos musculares [10]. La laringe no tiene como función principal la reproducción de la voz, sino que su misión es controlar la entrada de aire y evitar la entrada de cuerpos extraños en el aparato respiratorio del ser humano. La laringe es una estructura-cartilaginosa, hueca y móvil, con la forma de pirámide triangular invertida, que forma parte del conducto aerífero.

Se localiza en la parte media y anterior del cuello, por delante de la faringe, debajo del hueso hioides y detrás de los planos musculo-aponeuróticos de la región infra-hioidea. Lateralmente se encuentra limitada por los lóbulos del cuerpo o tiroideo y el paquete vaso-nervioso del

cuello. Su relación con la columna varía según un la edad y el sexo; es más alta en los niños que en los adultos y ligeramente más alta en las mujeres que en los varones [11]. En el adulto promedio mide aproximadamente 5cm de longitud, 3cm de diámetro y su extremidad inferior se corresponde con la sexta vértebra cervical. Posee tres funciones básicas que en orden de prioridad funcional son:

- Protectora o esfinteriana: Las estructuras móviles de la laringe (epiglotis, pliegues ariteno-epiglóticos y bandas ventriculares) desarrollan acciones reflejas de construcción que les permite actuar de manera conjunta como esfínter para impedir que los alimentos y cuerpos extraños pese a la tráquea y obstruyan la vía aérea. Su acción es especialmente importante durante la deglución.
- Respiratoria: La apertura glótica favorece el flujo de aire a los pulmones durante la inspiración, mientras que su cierre parcial permite el control del aire espirado y la regulación de la presión intra-alveolar.
- Fonatoria: Durante la fonación, los pliegues vocales (cuerdas vocales) vibran modificando la presión de aire aspirado y generando un tono, cuya frecuencia varía en función de la tensión y longitud de las mismas, del ancho de la hendidura glótica y de la intensidad del esfuerzo espiratorio [12].

El aparato de la reproducción del habla puede ser considerado para su estudio como parte del sistema respiratorio o puede ser visto de forma independiente, debido a que el apartado productor del habla depende de la función primordial del aparato respiratorio, la ventilación.

En la figura 1, se puede observar de forma general la integración de órganos que intervienen en la producción de voz, como son pulmones, bronquios, tráquea, laringe, faringe, nariz y boca.

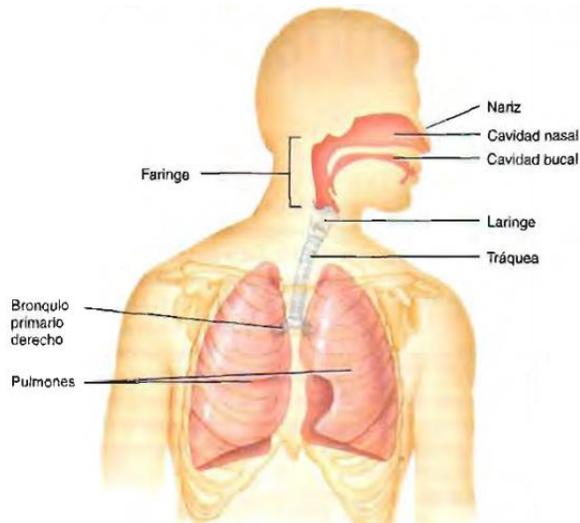


Fig 1. Aparato respiratorio.

La ventilación es el proceso de inhalación y expiración del aire. En el cual, los pulmones se expanden en la inhalación del aire y contraen en la expiración para la expulsión del aire, en el cual el aire viaja a través de los órganos que se mencionaron y que forman parte del aparato productor del habla.

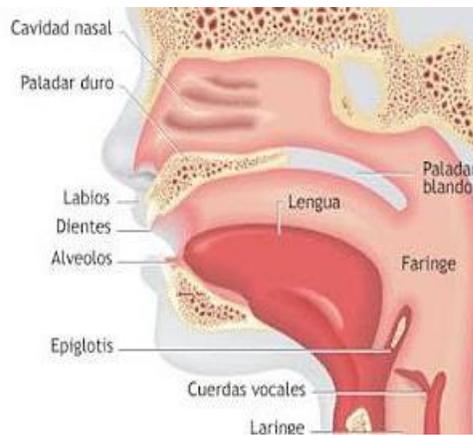


Fig 2. Aparato fonador. Los órganos de la fonación son los pulmones y la laringe; los órganos de la articulación son la cavidad del tracto vocal, nasal y senos paranasales, velo o paladar, lengua, labios y mandíbula inferior.

Para su comprensión el aparato de la producción de voz puede ser dividido en los órganos de fonación (producción del sonido) y los de articulación (órganos del habla) [13].

2.2 Mecanismos de la fonación

2.2.1 Cuerdas vocales

Se comienza con un flujo de aire contenido en los pulmones que es expulsado a su paso en dirección a la cavidad oral, produciendo una vibración de las cuerdas vocales; esta presión de aire hace que aparezca una abertura en estas cuerdas, que se conoce como glotis como se muestra en la figura 3.

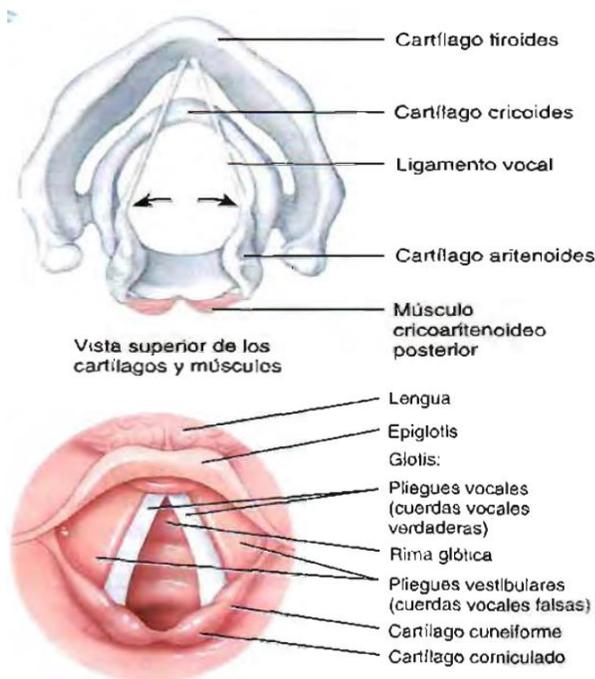


Fig 3. Cuerdas vocales.

Durante la fonación, las cuerdas vocales actúan como un transductor que convierte la energía aerodinámica, generada por el aparato respiratorio en energía acústica radiada al nivel de los labios que es escuchada como voz. La transducción de la energía se produce en la glotis por la vibración de las cuerdas vocales modulada por las variables subglótica y supra glóticas. El conocimiento de la composición estructural de las cuerdas vocales no solo es clave para comprender la vibración sino que también hace que se entienda que una leve alteración en dicha estructura implica la aparición de perturbaciones en su comportamiento vibratorio. La vibración de las cuerdas vocales tiene una apariencia visual a la de ondas que atraviesan la superficie mucosa de abajo hacia arriba en ciclos regulares [14]. La propagación del movimiento vibratorio de la mucosa se ordena en una secuencia de movimientos medial de cierre lateral y de apertura

a lo largo del borde libre de las cuerdas vocales, desde su parte inferior donde se forma un labio inferior hasta su parte más alta donde se forma un labio superior [15]. Dado que este movimiento se asemeja a las ondas que se propagan en la superficie de un líquido, se denomina onda mucosa cuya velocidad varía en función de las condiciones de la cuerda cuando se somete a un estiramiento existe mayor presión subglótica, mayor flujo aéreo espiratorio y en presencia de una contracción muscular laríngea asociada con la fonación de frecuencias altas, viaja a mayor velocidad [16].

2.2.2 Estructura cartilaginosa

Está constituida por tres cartílagos impares y tres pares. Los cartílagos impares y medios, de arriba abajo son: epiglótico, tiroideos y cricoides; los cartílagos pares y laterales son: los aritenoides, los corniculados de Santorini y los cuneiformes de Wrisberg (también conocidos como de Morgagni). Se presentan también los cartílagos sesamoideos anteriores y posteriores y el inter-aritenoides [17], [18].

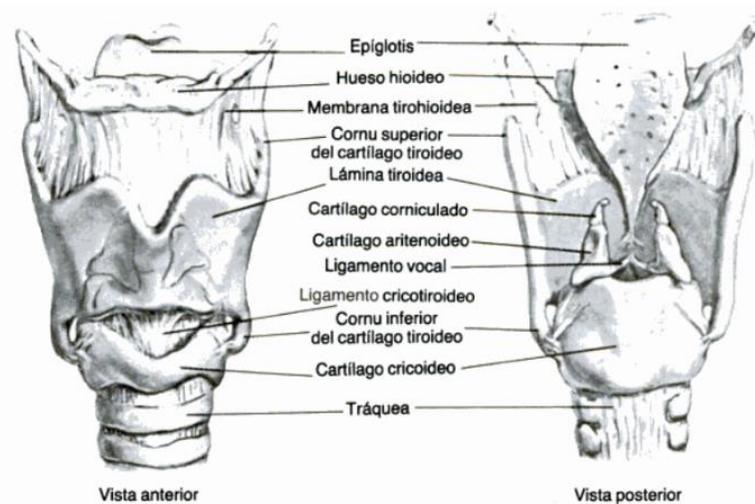


Fig 4. Anatomía de la laringe.

La laringe es un órgano que participa en funciones como la deglución, respiración y fonación; tiene un esqueleto y músculos intrínsecos y extrínsecos que, de acuerdo con las circunstancias, le permite cerrar y abrir tensando las cuerdas vocales, las desplaza verticalmente en el cuello,

según lo requiere cada función específica [19]. Los músculos intrínsecos tienen que ver con la actividad “glótica”. Pueden dividirse en dos grupos según actúen sobre la tensión de la cuerda vocal o el cierre y apertura de la glotis [20]. El músculo tiroaritenoso, también llamado tensor externo, cumple su función de incrementar la distancia entre los puntos de inserción de la cuerda vocal.

Cartílago Epiglótico: Es un cartílago fibro-elástico situado en la parte ántero-superior de la laringe, por detrás del cartílago tiroideo y de la membrana tiro-hioidea [21]. Tiene el aspecto de una lámina de forma oval con eje mayor vertical, más largo que ancho, su terminación inferior se denomina cola de la epiglotis y tiene forma de lámina afilada [22].

Cartílago Tiroideo: está situado por encima del cartílago cricoideo, al que se unen por medio de la membrana crico-tiroidea y por debajo del hueso tiroideo, del que se suspende mediante la membrana tiro-hioidea [23]. Está constituido por dos láminas cuadradas de cartílago hialino, denominadas placas o láminas laterales, fusionadas por delante en la línea media formando un ángulo diedro, que por fuera se denomina ángulo tiroideo y por dentro escotadura tiroidea.

Cartílago Aritenoides: son dos pequeños cartílagos hialinos, bilaterales y simétricos, dispuestos a ambos lados de la línea media. Está situado por encima de la lámina del cricoideo sobre la que se apoya y con la que se articula [24].

2.3 Estructura muscular

Los músculos son los responsables de la producción activa de la mayoría de los movimientos que ocurren entre las estructuras propias de la laringe y de adyacentes a ellas. Se clasifican en intrínsecos y extrínsecos al igual que los ligamentos y membranas. Los músculos intrínsecos tienen todas sus inserciones en los cartílagos laríngeos [25]. Al contraerse o relajarse producen

diversos efectos sobre la posición o tensión de los pliegues vocales o bien sobre la hendidura glótica para dar lugar a su función esfinteriana. La contribución muscular a la producción de la voz y particularmente a la regulación de la frecuencia fundamental sigue siendo un tema de controversia ya que no se comprende por completo. Todos los músculos intrínsecos de la laringe están inervados por la rama laríngea recurrente del nervio vago, a excepción de los músculos crico-tiroideo, que se inerva por una rama externa del nervio laríngeo superior. La tabla 1, resume las principales características y funciones de los músculos intrínsecos de la laringe [26].

Tabla 1. Músculos intrínsecos de la laringe.

Músculo	Origen	Inserción	Función
Crico-tiroideo	Cara anterior-lateral del arco del cartílago cricoides.	Parte oblicua: hasta menor del cartílago hioides, parte recta: borde inferior del cartílago tiroides.	Rotación adelante y abajo del cartílago. Tiroides sobre la articulación crico-tiroidea.
Crico-aritenoides	Superficie posterior del cartílago cricoides.	Superior posterior de la apófisis muscular del cartílago.	Rotación externa y abducción del cartílago aritenoides.
Aritenoides transversos	Borde lateral posterior del cartílago aritenoides	Borde lateral posterior del cartílago aritenoides opuesto	Aducción del cartílago aritenoides.
Vocales	Superficie lateral de la apófisis del cartílago aritenoides.	Ligamento vocal y ángulo tiroideo	Regula la tensión de los pliegues vocales.

2.4 La voz Humana

2.4.1 Producción de la voz

La voz es un sonido producido por el cuerpo humano, en específico órganos de fonación y articulación. Para que se produzca dicho sonido se necesitan acciones complejas del cuerpo, las cuales suceden a través de unidades funcionales, las cuales son:

- Generación de la presión del aire.
- Regulación de la vibración.
- Control de resonadores.

La generación de la presión de aire se genera en los pulmones y cuando se está hablando se tiene una fase corta de expiración por una prolongación de fase de inhalación. La regulación de la vibración vocal es una acción entre el flujo de aire de los pulmones y la acción de las cuerdas vocales debido a un conjunto de músculos laríngeos [27].

Al oscilar las cuerdas convierten el aire espiratorio en intermitentes pulsos de flujo de aire, producidos en una frecuencia fundamental, este sonido se considera imperceptible a los rangos audibles. La frecuencia de este sonido depende de varios factores, que pueden ser el tamaño y la masa de las cuerdas vocales, la tensión que se aplique y la velocidad del flujo proveniente de los pulmones. A mayor tamaño, menor frecuencia de vibración. Por ello los hombres tienen una voz más grave a una frecuencia menor en comparación con las mujeres las cuales tienen una frecuencia mayor y por consecuencia más aguda, la voz presenta cambios principalmente en la adolescencia, también a mayor tensión en las cuerdas vocales la frecuencia aumenta, por lo que para emitir sonidos agudos es necesario un mayor esfuerzo vocal tabla 2 [28].

Tabla 2. Frecuencia fundamental en hombre y mujeres respecto a edades.

Frecuencia fundamental (Hz)								
Edad	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90
Hombre	112-120	112	107-113	188	112	124-132	127-138	-
Mujer	227-246	200-214	214	212	209	206	197	192

Para transformar el sonido de la voz, que es un sonido turbulento y con funciones lingüísticas especiales, se tiene el control de los resonadores; estos se localizan en el tracto respiratorio superior, por la laringe, los cuales son la cavidad oral y nasal. Durante la fonación los pulmones actúan como la fuente principal de energía del sistema, generando durante la espiración un flujo de aire continuo con presión controlada, produce un tono sonoro el cual también se le conoce como sonido laríngeo [29].

Las características principales de ese tono (intensidad, frecuencia fundamental y timbre) dependen del volumen de aire espirado, tensión y longitud de los pliegues en aducción. Los principales articuladores son: la lengua, mandíbula, labios y el velo [30]. Dentro del lenguaje se consideran de forma general dos tipos de fonemas: las vocales y consonantes, su diferencia principal radica en la forma de vibración de las cuerdas vocales. La lengua y los labios también participan en la producción de sonidos no vocálicos por medio de los cuales se articulan las palabras, estos no dependen del sonido laríngeo producido por la vibración de los pliegues vocales, son producto de un obstáculo o una interrupción total momentánea del paso del aire a través del canal vocal, generando turbulencia [31]. En la tabla 2 se muestran las frecuencias de vibración de los pliegues vocales de una determinada persona puede variar dentro de un rango (en promedio de 300 Hz o dos octavas) al cual se le denomina registro vocálico o tesitura que a su vez puede estar comprendido entre 50 y 1000 Hz, siendo más agudo en las mujeres y niños y más grave en los hombres adultos [32].

En cada ciclo de vibración (ciclo vocálico) pueden distinguirse seis fases principales; tres de aducción (oclusión) y tres de abducción (apertura). Cada ciclo inicia con la aproximación de los pliegues vocales y termina con su separación máxima como se muestra en la (figura 5). Se ha estudiado un componente de movimiento vertical de los pliegues y un ciclo de movimiento ondulatorio muy particular de sus bordes libres.

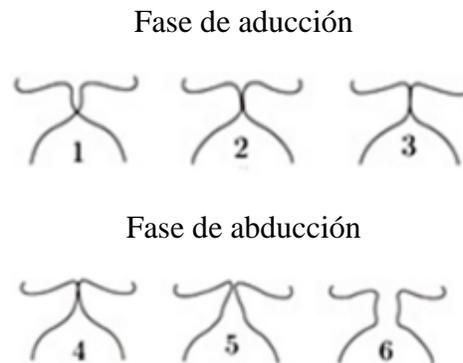


Fig 5. Ciclo de vibración de los pliegues vocales. La consecuente disminución de la presión subglótica y las propiedades elásticas de los pliegues hacen que estos vuelvan a la línea media en donde se inicia el ciclo; al acercarse a la línea sufren una succión debido al efecto Venturi que obedece al principio de Bernoulli, dicha succión produce una aducción abrupta que interrumpe el paso de aire y da origen a un nuevo ciclo de vibración.

En la teoría mio-elástica-aerodinámica, las vibraciones de los pliegues vocales dependen de sus propiedades elásticas, de la acción muscular y de fenómenos aerodinámicos. La producción de sonido es precedida por una fase prefonatoria de aproximadamente 30 ms, durante la cual, por efecto muscular los pliegues vocales, se aducen ocluyendo la glotis. La presión subglótica vence la resistencia que ofrecen los pliegues vocales, iniciando la abducción de los mismos y permitiendo el escape de aire [33].

2.5 Calidad de la voz

Unos de los mayores desafíos para la investigación en foniatría han consistido en la búsqueda de un método que posibilite la valoración confiable y precisa de la calidad de la voz, lo cual es indispensable para determinar la eficacia de los tratamientos y en la realización de análisis comparativos entre ellos. La tendencia actual en la evaluación de tratamientos es recurrir al uso

de análisis multidimensionales basados en evaluaciones tanto objetivas como subjetivas de la calidad de la voz, diferentes autores coinciden en que la valoración subjetiva de los trastornos de la voz deben referenciarse al efecto que estos producen en la calidad de vida de los pacientes [34].

En algunos de los casos la valoración de la calidad vocal se efectúa de manera perceptual por parte de médicos foniatras con entrenamiento en la detección, por simple audición, ciertos indicadores de su calidad vocal tales como la ronquera, ruido etc. A pesar de ser ampliamente utilizados, el análisis subjetivo de la voz carece de valor científico requerido y es poco útil para la comparación de tratamientos debido a la dificultad que establece respecto al juicio auditivo y sus correlaciones fisiológicas. La valoración objetiva se basa en la medición de diversos parámetros acústicos de registro de voz, tales como la frecuencia fundamental “*pitch*”, las diferencias o perturbaciones entre ciclo y ciclo, los niveles de ruido y el tiempo máximo de fonación “MTP” [35].

Entre las opciones para el análisis y valoración de la calidad de la voz se encuentran las siguientes:

- Escala GRBAS (*Grade, Roughness, Breathiness, Asthenia, strain*).
- Índice de severidad de disfonías.
- Índice de discapacidad vocal (*VHI-Voice Handicap Index*).
- Escala sintomatológica de la voz (*VoiSS-Voice Symptoms Scale*).
- Tiempo máximo de fonación (*MTP- Maximum Phonation Time*).
- Análisis acústico de la voz.

El análisis acústico también presenta algunas desventajas debido a la variabilidad de la voz, es difícil determinar la relación entre cada parámetro y la calidad perceptual de la misma, además, no se han logrado establecer valores normativos que definan una voz normal o sin disfonía y que pueden ser empleados universalmente [36].

2.6 Vocales

La palabra “vocal” proviene del latín “*vocalis*” que significa “con la voz”; este término se usa comúnmente para definir tanto a los sonidos vocálicos como los signos escritos que los representan, los cuales son importantes en la mayoría del lenguaje tanto en palabras como en el discurso. Las vocales en el alfabeto español, de acuerdo a la academia mexicana de la lengua, son cinco: /a/, /e/, /i/, /o/ y /u/ y su posición se debe a la progresividad del cierre de la abertura oral. Para las frecuencias fundamentales podemos decir que son los armónicos que tienen una armónica diferente de las demás en donde se puede apreciar la forma de onda con algunos detalles de las formantes [37].

2.6.1 Producción de las vocales

Para originar un sonido vocálico los mecanismos de fonación y articulación trabajan juntos. Un flujo de aire expulsado de los pulmones, el cual viaja a través de bronquiolos y la tráquea, llega a la laringe en donde se encuentra con las cuerdas vocales, este flujo de aire y la adaptación fisiológica de las cuerdas vocales, propicia que estas se abran y cierren, creando así pulsos cuasi periódicos de aire con una frecuencia fundamental llamada “*pitch*”; sin embargo se crea un sonido sordo por el movimiento de las cuerdas, el cual es modificado por el tracto vocal. La configuración que tengan los órganos de articulación determinará la frecuencia de resonancia y por tanto determinará el sonido producido, la calidad de las vocales orales es determinada por los formantes [38].

La lengua, velo y mandíbula inferior influyen de gran manera en la formación del sonido de una vocal. El velo contribuye a la articulación de las vocales abiertas por la disminución del área del tracto vocal haciendo estrecha la rama de la cavidad nasal. La posición de la lengua, así como la forma que tome su cuerpo son indispensables para el tipo de pronunciación [39]. La apertura de la mandíbula así como la forma que tomen los labios, redondeada, alargada, protrusión y retracción son parámetros importantes para la producción de una vocal. En la teoría “filtro-fuente” se modela la producción de las vocales, en donde la función de la laringe es como una

fuelle generadora y el tracto vocal es un filtro acústico para modular los sonidos de la fuente, que salen de la apertura de los labios.

En la producción de vocales, la forma del tracto vocal es como un tubo cerrado con la parte terminal y cerrada en el fin de la glotis, la parte abierta final como la apertura de la boca. Múltiples reflexiones de ondas de sonido entre las dos terminales del tracto vocal dan las formantes. Cuando se realiza un análisis sobre las señales producidas en la pronunciación de vocales, se observa que son generalmente largas en duración en comparación con las señales producidas por las consonantes, contienen alta energía y espectralmente se encuentran bien definidas y tiene periodicidad en un tiempo corto [40].

2.6.2 Formas de evaluación en la fonación.

Existen formas de evaluar la calidad de voz, estas no siempre se obtiene mediante espirómetro, frecuentemente se hacen otros procedimientos más especializados como son: los estudios radiológicos (tórax y vías respiratorias altas), tomografía (computarizada y angiografía pulmonar), radioisótopos, endoscopios (laringoscopia y broncoscopia), sin embargo, existe otro tipo de método en donde medimos el tiempo máximo de fonación. Con el fin de producir una buena calidad de voz, una persona debe ser capaz de llevar sus pliegues vocálicos entre sí con facilidad a lo largo de toda su longitud. Durante la fonación de sonidos de voz, las cuerdas vocales vibran alrededor de 100 veces por segundo para los hombres adultos y alrededor de 200 veces por segundo para las mujeres adultas. La capacidad de cerrar la glotis también es importante en la protección de la vía aérea en la laringe, el cierre de la laringe permite toser fuertemente [41].

Una persona que no es capaz de tener una deglución adecuada relacionada al paso del aire en las cuerdas vocales puede experimentar una mala calidad de voz, dificultad para proyectar la voz y fatiga vocal, ya que tienen que aumentar el esfuerzo requerido para la fonación.

2.7 Tiempo máximo de fonación (MPT)

Una persona tiene una cierta cantidad de aire y ser capaz de sostener una vocal durante un tiempo suficiente para lograr un buen control de su voz en la comunicación. La producción de la vocal sostenida es un método para obtener el MPT por sus siglas en inglés “*Maximum phonation time*” con una sola respiración. Se puede analizar tres tipos de vocales: aguda /i/, grave /u/ y media /a/ un cronometro inicia el tiempo de cada una hasta que termina la fonación de la vocal. El tiempo máximo de fonación se emplea para estudiar la eficiencia del cierre glótico y el sistema respiratorio general, este estudio se mide en segundos, se puede estimar MPT usando la capacidad vital (CV), mediante las siguientes formulas indicadas en un hombre adulto ecuación (1) y una mujer adulta ecuación (2).

$$MPT = \frac{CV}{1/0} \times 0.67 \quad (1)$$

$$MPT = \frac{CV}{1/0} \times 0.59 \quad (2)$$

Generalmente, los adultos pueden sostener una vocal de 15 a 20 segundos [42]. La duración mínima normal aceptada es de 15 segundos para los hombres y 14.3 segundos para las mujeres. Los niños generalmente pueden sostener una vocal por 10 segundos, la relación entre la duración máxima de la vocal sostenida dividida por el valor esperado de 0.70 puede indicar una o más anormalidades como son:

- Insuficiencia del control glótico.
- Insuficiencia respiratoria.
- Problemas neurológicos.
- Patrones erróneos del uso de la voz.

2.8 Presión subglótica

La presión subglótica representa la energética que crea la señal acústica en la palabra, es muy importante que los niveles de presión sean adecuados. Los niveles inadecuados de presión subglótica pueden producir problemas en el habla con cambios de intensidad y de frecuencia fundamental “*pitch*”, la variación de la presión subglótica puede deberse a diversos problemas de tipo neuromuscular (parálisis de cuerdas vocales o enfisemas).

La medida de la presión subglótica es importante en el diagnóstico y en la terapéutica de los problemas de la voz y el habla, requiere conocimientos en el sistema respiratorio y de cómo evaluarlo en función del problema. En la mayoría de los casos durante la fonación la presión subglótica es de 5 a 10 cm H₂O, la presión varía directamente con la intensidad de una vocal la relación no es lineal y varía dependiendo de la persona sea hombre, mujer o un niño. Se conocen diferentes técnicas para medir la presión subglótica en forma directa o indirecta [43].

Forma directa: punción traqueal, catéter transglotal y el uso de transductores diminutos.

Forma indirecta: medida de la presión a través de la traqueotomía, globo esofágico y máscara bucal.

La presión durante la fonación está dada en medidas menores, medianas y altas estas tiene un grado de presión correspondiente que son:

- Presión menor: 9.5 cm H₂O
- Presión media: 13.7 cm H₂O
- Presión alta: 20.1 cm H₂O

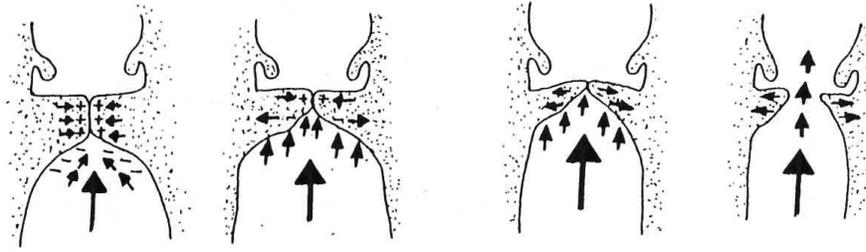


Fig 6. Fase de abertura, la presión subglótica abre los pliegues vocálicos.

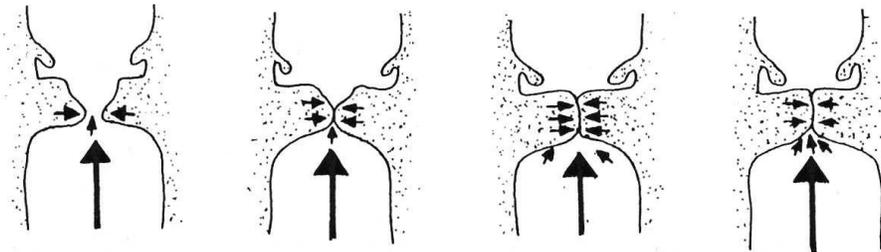


Fig 7. Fase de cierre, los pliegues se juntan por efecto de la elasticidad y de la disminución de la presión.

Los pliegues vocales se aproximan como resultado de las contracciones musculares que actúan sobre los aritenoides, la velocidad del aire aumenta por efecto de la disminución en la abertura de la glotis y la presión en la glotis disminuye además de que se aproximan por efecto de Bernoulli más la elasticidad muscular, la glotis se cierra y la presión subglótica aumenta al final los pliegues vocales se separan [44].

Capítulo 3. Desarrollo

3.1 Descripción del registro de voz.

La voz nos permite la comunicación entre los seres humanos, contiene información que permite identificar ciertos parámetros como la frecuencia fundamental, energía, intensidad, tono y timbre. El procesamiento digital de señales de voz tiene una gran variedad de aplicaciones, podemos implementarla en valorar la calidad de voz de una persona independientemente del sexo y edad teniendo una base de datos y un sistema que realice un registro de la voz con ciertas características. Para la resolución del problema se desarrolló un sistema que registra la voz, utilizando un micrófono multimedia para capturar la señal de voz, por medio de una adquisición de datos se almacenara en un archivo .WAV para luego ser mostrada dependiendo de ciertos parámetros (frecuencia fundamental, energía de la señal, cruce por cero y una señal de presión subglótica). Así como un sistema interactivo que nos permitirá realizar una evaluación del paciente que se esté tratando dependiendo de la patología foniátrica común que presente. En la (figura 8) se muestra un diagrama a bloques del sistema de registro de voz propuesto.

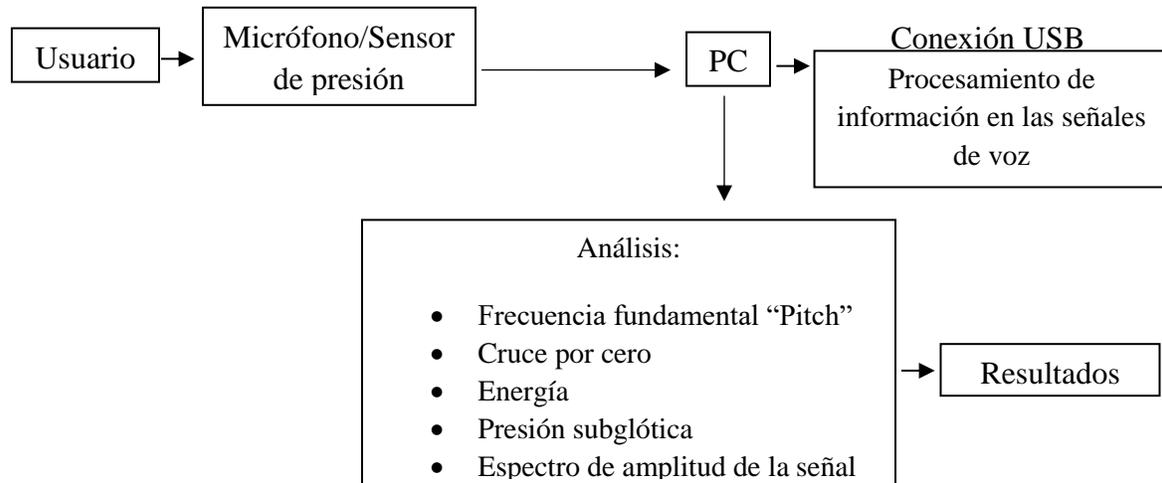


Fig 8. Diagrama a bloques de la solución propuesta del sistema de voz.

3.2 Instrumentos empleados

3.2.1 Dispositivo de registro electrónico de voz.

El sistema de registro de voz (figura 9) nos permite registrar la calidad de la voz, además de una señal de presión subglótica [cm H₂O].

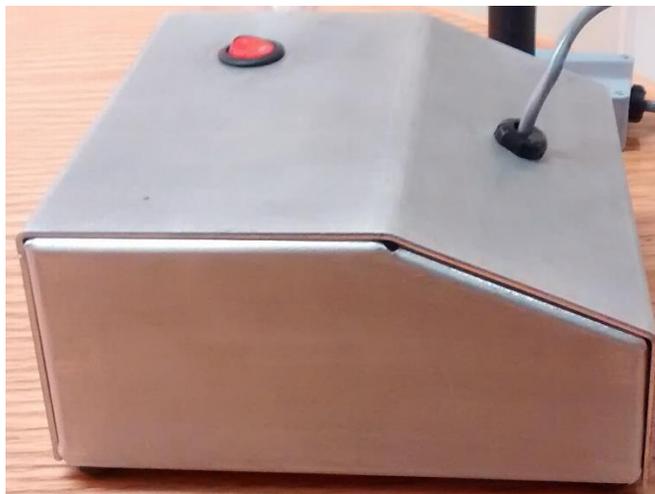


Fig 9. Sistema de registro de la voz.

En la figura 10 se muestra el sistema de captura de la voz por medio de un micrófono electret comercial, además de registrar la presión glótica por medio de un sensor de presión, al final de la mascarilla se tiene un conducto que responde en forma de Venturi, esto con el fin de tomar una muestra del aire cuando se realiza la fonación, esta muestra de aire pasa por medio de un sensor de presión en la parte inferior de la boquilla en forma de T.



Fig 10. Sistema de acople con sensores de presión para el registro de la señal de voz.

Las grabaciones de audio y presión se almacenan en un rango entre el máximo, mediano y mínimos producidos en una fonación mínima y máxima esto registros se capturaron en relación al tiempo e intensidad en que el paciente realiza una vocalización, independientemente del género y edad. El audio de la señal fue grabado usando un micrófono electret cubierto por una esponja para minimizar la turbulencia del aire cuando el paciente realice la vocalización, el micrófono se colocó en la mascarilla de plástico cerca de la boca del paciente, al mismo tiempo que actúa como cavidad resonante, contribuyendo a la obtención de una señal con menor energía de ruido armónico. La señal de presión se captura a través de una tarjeta de adquisición de datos multicanal llamada Arduino.

3.3 Cadena de medida del sistema propuesto.

En la figura 11 se muestra el sistema general del sistema propuesto para el registro de voz.

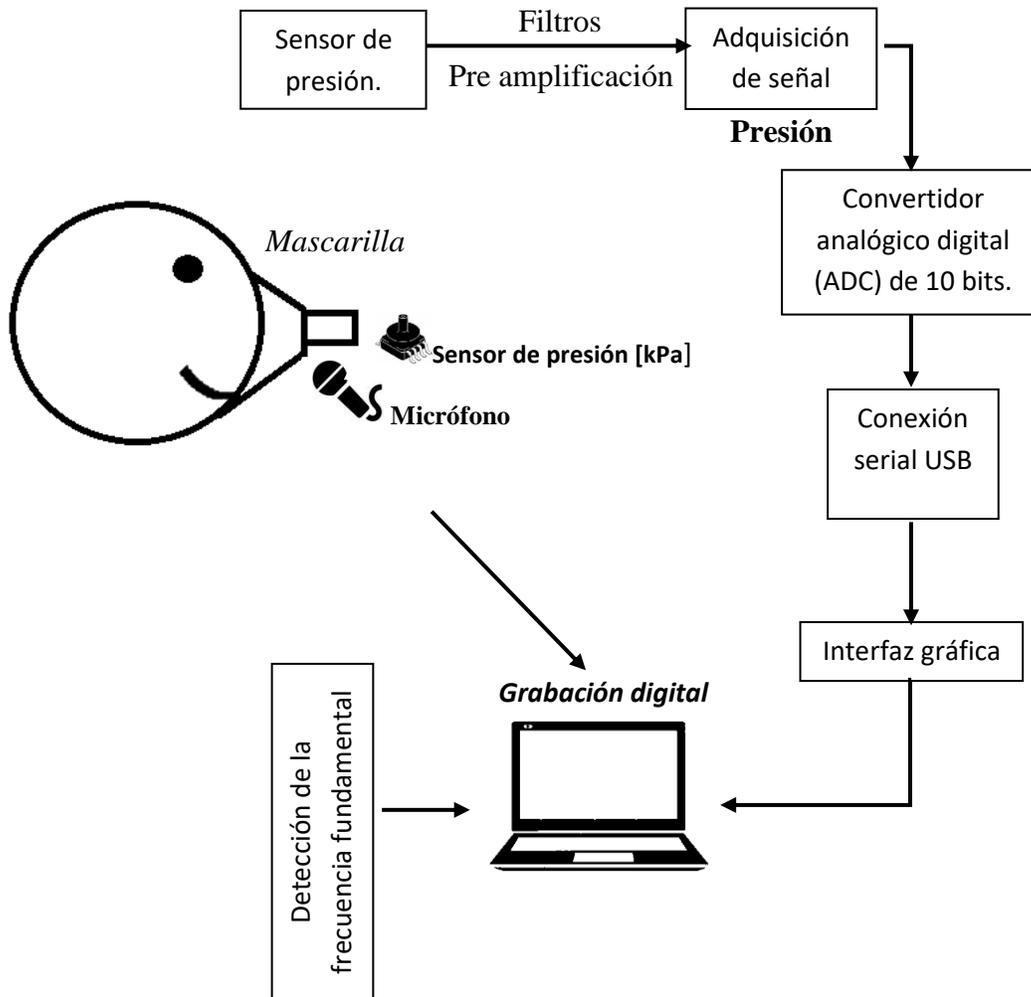


Fig 11. Instrumentación del sistema propuesto.

3.3.1 Frecuencia fundamental

La “*pitch*” o frecuencia fundamental segmenta la señal y la divide en sub señales más pequeñas, la vibración glotal puede presentar una periodicidad debido a los cambios en la amplitud, la “*pitch*” se extrae del primer armónico así como el periodo fundamental que corresponde el inicio y el final en el que se cierra la glotis, este tipo de procesamiento nos facilita analizar la señal para eliminar el ruido y algunos factores que se presenten, además de obtener el tiempo contra

frecuencia. Es una medición que indica el número de ciclos por segundo del patrón vibratorio este parámetro define la voz de cada persona así como la intensidad, aunque existen cambios a lo largo de la adolescencia debido a los cambios tisulares que se tiene con la edad.

La primera etapa de desarrollo del proyecto consistió en grabar la voz de mujeres y hombres entre las edades de 27 a 50 años de edad con patología común vocálica y sanos, se realizó un registro electrónico en una base de datos de los pacientes la cual contiene nombre completo, edad y diagnóstico, una vez realizado el registro en la base de datos se le pidió a los pacientes que se colocaran la mascarilla de plástico a la altura de la boca ya que se está contiene el micrófono que realizara la grabación de la voz en un archivo WAV, estas grabaciones se guardan para aplicar el algoritmo que corresponde al análisis (*Pitch*, espectro de amplitud, cruce por cero y energía) en el software de MATLAB. Al inicio de la grabación se les explica a los pacientes que debe de realiza una fonación de la vocal /a/ durante 5 segundos en el que dura el registro de la voz.

La frecuencia fundamental se determina por la ecuación (3) en donde el periodo de muestreo (periodo fundamental) tiene un tiempo de muestreo estándar de $T_m=11025$.

$$f_0 = \frac{1}{s * T_m} \quad (3)$$

Las diferencias de magnitudes promedios de la señal se le realizo una ventana de 256 datos, sin embargo, se puede utilizar un ventaneo cuadrado de amplitud 1 respecto a la longitud de la señal. Las formantes las podemos identificar por los picos siguientes independientemente al primer pico que encontramos como frecuencia fundamental o "*pitch*".

3.4 Energía de la voz.

En la obtención de la señal de energía de la vocal, se determinó por las ecuaciones en relación al muestreo y el periodo en forma de ventanas “*hamming*”, la energía en la señal se debe a la variación de la presión subglótica y de la forma del tracto vocal, esto es útil para distinguir segmentos de fonemas no vocálicos y vocálicos; los primeros (no vocálicos) representan una alta energía, los fonemas vocálicos muestran que son de baja energía, el algoritmo que se utilizó para obtener la energía de la voz está determinado por la ecuación 4.

$$E = \sum_{m=0}^{N-1} [x]^2 \quad (4)$$

Los valores de x indican los segmentos “*frames*” de la señal (256 datos), realiza una secuencia de ventanas en toda la longitud, teniendo diferentes factores de ponderación, la ventana “*hamming*” que seleccionamos para “ N ” muestras del rango de la ventana que se calcula $n-N+1$ todos los valores de n son 256 datos que se pueden escoger dentro de los 10-20 ms de muestras en una frecuencia de muestreo de 11025 Hz, se incrementó el rango dinámico de la función para calcular la energía en periodos cortos durante la grabación.

3.5 Cruce por cero de la señal de voz.

Para determinar los cruces por cero (ecuación 5) de la señal de voz se obtuvieron todas las 256 muestras requeridas en un periodo de tiempo de 0.00025 s por cada “*frame*” (segmento) esto para que proporcione una aproximación de la frecuencia de muestreo de 11025 Hz, la señal tiende a cruzar el eje dos veces e indica el número de repeticiones que toma el valor de cero.

$$Z(n) = \frac{1}{L} \sum_{m=-M}^M W(m) |sgn[s(m)] - sgn[s(m-1)]| h(n-m) \quad (5)$$

M resulta el número de muestras de la señal y $sgn[s(n)]=1$, para $s(n)=0$, $s(n)=-1$ para $s(n)=0$ indican el número de veces que una señal toma un valor de cero. Dado que las frecuencias son

altas implica un alto valor en los cruces por cero, es por eso que a frecuencias bajas implica un cruce por cero muy bajo, esto quiere decir que hay una correlación y distribución de la energía de la voz con la frecuencia, si el cruce por cero es alto la señal de voz es fricativa, mientras que si la tasa de cruces por cero es bajo, la señal es perteneciente a una señal de voz (vocal), proporciona una medida aceptable en la distinción de tres tipos de fonemas que son: fricativos, nasales y silbantes. La manera de entender esta técnica radica en que cada vez que tiende a cero el segmento de la señal es aproximado al tiempo de duración de la función ventana, este algoritmo está relacionado con la frecuencia fundamental.

3.6 Caracterización de sensor de presión.

La primera etapa del sistema fue la caracterización de un sensor de presión marca Honeywell (figura 12), a continuación se explica el funcionamiento del sensor así como las ecuaciones que se utilizaron en la relación voltaje/presión.

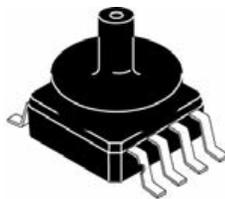


Fig 12. Sensor de presión marca Honeywell modelo (MPxx5004). Encapsulado de plástico con dos piezas mecánicas diferentes, este sensor combina un medidor de tensión implantado altamente sensible con técnicas avanzadas de micro mecanizado, metalización de película delgada y procesamiento bipolar para proporcionar una señal de salida analógica de alto nivel proporcional a la presión aplicada.

3.6.1 Sensor de presión.

La caracterización del sensor de presión Honeywell se realizó en conjunto con un sistema de medición de presión comercial (Baumanometro [mm Hg]) marca Mod Econom. Se midió el voltaje de salida del sensor comparando la medida del sistema de presión comercial (Baumanometro [mm Hg]), las lecturas iniciales son 0 mm Hg presión del Baumanometro no se le aplico ninguna presión así que la salida en voltaje del sensor Honeywell en una presión 0.

A continuación la tabla 3 muestra las lecturas de voltaje de salida del sensor en relación a la presión aplicada en el Baumanómetro en diferentes conversiones de presión [cm H₂O], [mm H₂O] y [kPa].

Tabla 3. Tabla de conversión de presiones con relación a la salida de voltaje del sensor.

Voltaje salida sensor[V] eléctrico	Presión [mm Hg] Baumanómetro (salida)	Presión [cm H ₂ O]	Presión [mm H ₂ O]	Presión [kPa]
1.198	0	0.000	0.000	0.000
1.403	2	2.719	27.191	0.267
1.650	4	5.438	54.382	0.533
1.880	6	8.157	81.573	0.800
2.200	8	10.876	108.763	1.067
2.510	10	13.595	135.954	1.333
2.750	12	16.315	163.145	1.600
2.970	14	19.034	190.336	1.867
3.300	16	21.753	217.527	2.133
3.510	18	24.472	244.718	2.400
3.910	20	27.191	271.909	2.666
4.120	22	29.910	299.099	2.933
4.440	24	32.629	326.290	3.200
4.630	26	35.340	353.400	3.466
4.890	28	38.067	380.672	3.733
4.910	30	39.974	339.740	3.920

Consideramos la conversión de diferentes presiones a la que se sometió el sensor de acuerdo a las especificaciones del fabricante (Anexo1), a continuación se muestran las ecuaciones de calibración del sistema general.

$$V_{out} = V_s * [(0.2 * P) + 0.2] \pm 6.25\%VF_{ss} \quad (6)$$

Para conocer si el voltaje de salida que arroja el sistema es el ideal o preciso consideramos las gráficas de la hoja de datos que el fabricante especifica en la caracterización habiendo realizado la conversión de presiones. La ecuación 6 utiliza la presión en kPa para realizar el cálculo de

error del sistema. Se consideró una presión extra que es la presión atmosférica, en la siguiente grafica (figura 13) de caracterización se visualiza que el sistema no comienza la medición en cero sino en 1.198v; el sistema calcula la presión absoluta esto quiere decir que nuestra presión de referencia inicial, como especifica el fabricante es 1v en la salida del sensor en un estado inicial de 0 cm H₂O presión, la presión absoluta existe porque varia con la altitud de la presión atmosférica, en la tabla 4 se muestran los resultados obtenidos utilizando la ecuación 6.

Tabla 4. Cálculos para encontrar el voltaje real en relación a la presión en [cm H₂O], VFss(es la diferencia algebraica entre el voltaje de salida a la presión especificada y el voltaje de salida a la presión nominal mínima de 0 kPa) hace referencia a la presión máxima del sensor 3.92 kPa.

Vs(Vcc)	Voltaje Medido[v]	P[kPa]	VFss[V]=Vmax-Vmin	6.25%VFss[V]=(0.0625)(VFss)	VoutReal=Vs*[(0.2*P)+0.2] (P=kPa)
5	1.198	0.000	3.712	0.232	1.000
5	1.403	0.267	3.507	0.219	1.267
5	1.650	0.533	3.260	0.204	1.533
5	1.880	0.800	3.080	0.193	1.800
5	2.200	1.067	2.710	0.169	2.067
5	2.510	1.333	2.400	0.150	2.333
5	2.750	1.600	2.160	0.135	2.600
5	2.970	1.867	1.940	0.121	2.867
5	3.300	2.133	1.610	0.101	3.133
5	3.510	2.400	1.400	0.088	3.400
5	3.910	2.666	1.000	0.063	3.666
5	4.120	2.933	0.790	0.049	3.933
5	4.440	3.200	0.470	0.029	4.200
5	4.630	3.466	0.280	0.018	4.466
5	4.890	3.733	0.020	0.013	4.733
5	4.910	3.920	0.000	0.000	4.920

La relación Voltaje[v]/Presión [cm H₂O] es nuestra sensibilidad se salida ya que el voltaje real es la conversión a presión en cm H₂O que es la presión que se busca para el cálculo de la presión

subglótica. La figura 13 muestra la calibración total del sistema de presión. El intervalo de medición que ocupamos fue el valor mínimo 0 cm H₂O y máximo 40 cm H₂O, como se explicó anteriormente, la presión atmosférica se suma a la presión que medimos así que se utiliza una técnica de autozero o autoajuste relacionada con la ecuación 6. La función autozero es una técnica de compensación basada en el muestreo del sensor a la presión de referencia (presión atmosférica) es una referencia cero para una medición de medida que corrige la salida del sensor para la variación del offset durante la medición, realizar un autozero reduce los errores del sistema de presión.

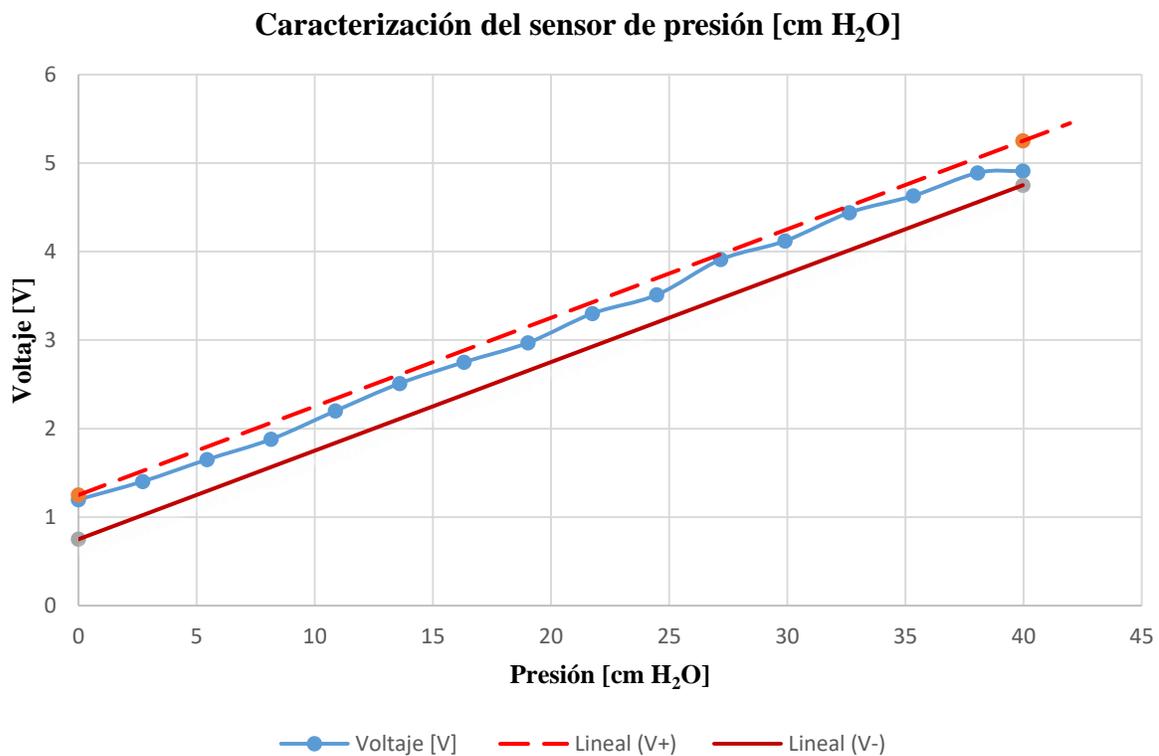


Fig 13. Caracterización del sensor de presión, la gráfica muestra la relación voltaje[v]/presión[cm H₂O], la línea punteada o segmentada (superior) es el voltaje máximo de salida del sensor en condiciones ideales en relación a la presión [cm H₂O] que el fabricante indica en la hoja de datos del sensor, la línea recta (inferior) es el voltaje mínimo de salida del sensor en condiciones ideales, la línea azul con puntos (Centro) es la más importante ya que esta indica los valores que se obtuvieron durante la calibración del sensor. Se observa que los resultados obtenidos son satisfactorios ya que estamos dentro del rango ideal considerando el sistema de medición comercial de presión que se utilizó, VF_{ss} es la diferencia entre el voltaje de salida del sensor con respecto al voltaje máximo menos el voltaje mínimo, esto lo utilizamos para medir el error que existe en comparación a la salida de voltaje

ideal que el fabricante considera (Línea central con puntos) es la salida del sensor caracterizado dentro de las condiciones ideales de error, se ajustó un error del sensor del 5%.

3.7 Interfaz gráfica del sistema del sistema de registro.

La interfaz gráfica de usuario (GUI) está desarrollada en Matlab® (2012a, *Mathworks*), en ella se realiza la recepción de los datos del sensor de presión en tiempo real, así como el registro de voz del micrófono, tenemos una comunicación vía serial con la tarjeta de adquisición Arduino, la figura 14 muestra la interfaz gráfica entre Arduino y Matlab para la comunicación.

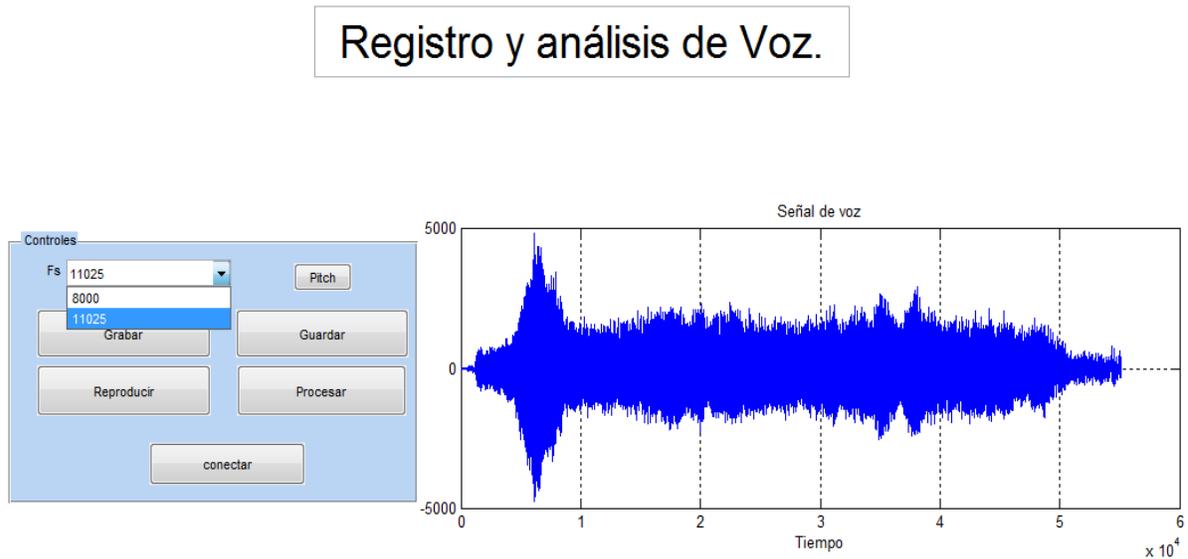


Fig 14. Interfaz gráfica para el registro de la voz. Se tiene un tablero de control para el tratamiento de la señal, se emplearon botones para reproducir y guarda el audio así como la comunicación serial.

En el manejo de muestro se propuso dos frecuencias de 8000 Hz y 11025 Hz, la señal de la presión subglótica se captura en tiempo real cuando se toma la muestra de voz del paciente. El botón “*pitch*” muestra la gráfica de la frecuencia fundamental del paciente por lo que se guarda automáticamente en una archivo.txt en nuestra base de datos. Las señal de voz es una señal variable en el tiempo, sin embargo, al ser analizada en periodos de tiempo cortos de 5- 500 ms,

se puede considerar que tiene características bastantes fijas en frecuencia en periodo y forma de la misma, es una señal invariante en tiempo y no en frecuencia se considera en periodos de tiempo cortos, existen diferentes técnicas en el dominio del tiempo y de frecuencia; algunas de las técnicas en el dominio del tiempo se encuentran la medición de la energía, la cual puede ser útil para distinguir la señal de una vocal (alta energía) con una constante (menor energía) además de conocer que partes de la señal hay un sonido presente (vocal) y en tramos en donde no existe sonido (consonante).

3.8 Sistema interactivo de voz.

Se desarrolló un sistema interactivo orientado a objetos en el IDE Microsoft Visual Studio en lenguaje C#. Dicho sistema consiste en un programa interactivo que registra la presión subglótica del usuario durante la fonación sostenida y en función de esta evade obstáculos que se van presentado de manera secuencial. La longitud de los obstáculos a esquivar depende del valor máximo de presión subglótica alcanzado por el usuario durante la calibración del sistema a partir de una fonación sostenida. La finalidad de este juego es rehabilitar al usuario que presente alguna patológica común foniátrica de manera intuitiva.

3.8.1 Programa de control del sistema interactivo.

En la figura 15 se muestra el diagrama de flujo general del sistema interactivo implementado en una computadora con procesador Intel Core Duo a una velocidad de 2.53GHz, con memoria RAM de 4.00GB y el sistema operativo Windows 7 de 64 bits.

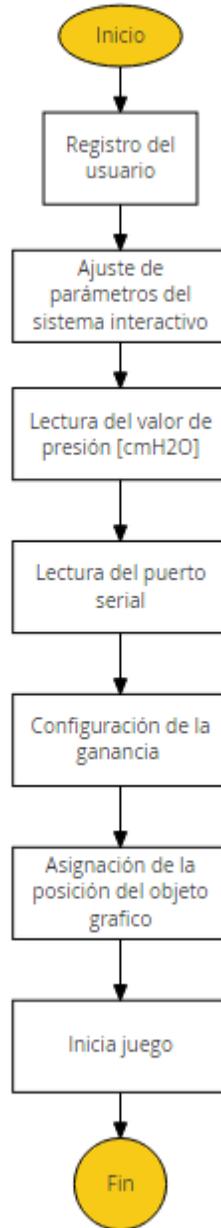


Fig 15. Diagrama de flujo general del sistema interactivo de voz

Mediante una conexión bluetooth el sistema interactivo puede recibir datos analógicos, esto por medio de un carácter establecido que ejecuta una instrucción para la comunicación entre módulos inalámbricos (figura 16).

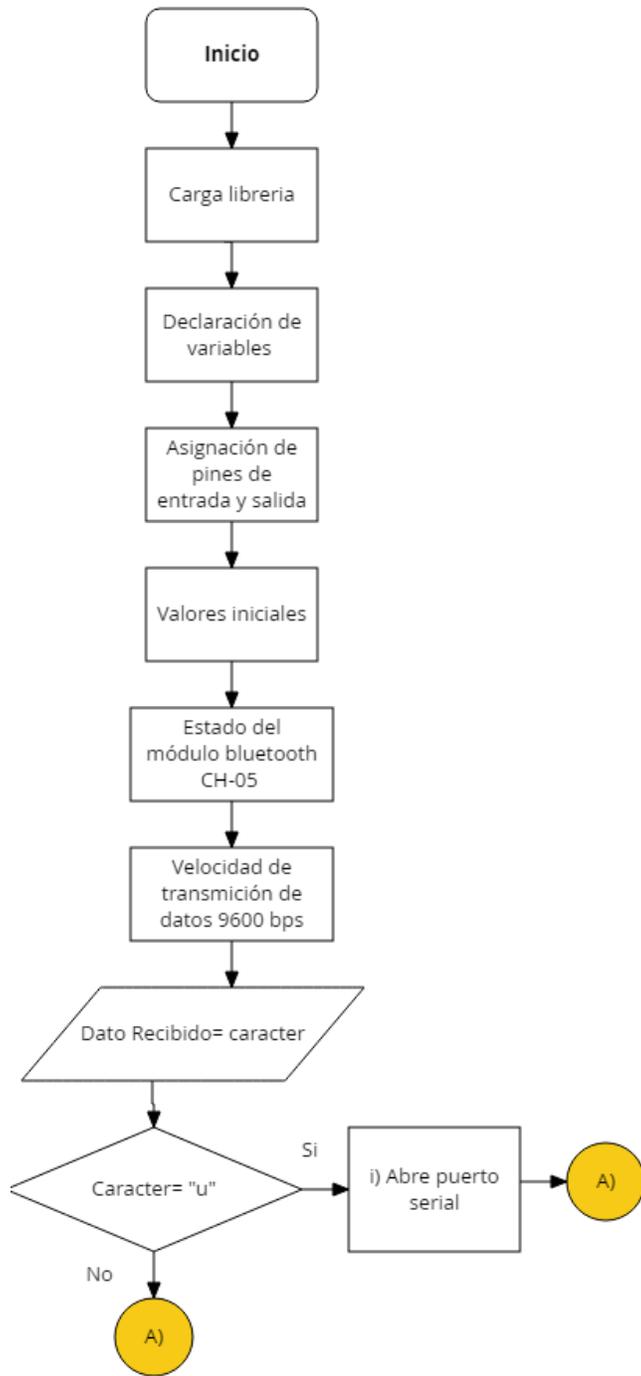


Fig 16. Inicialización del programa y lectura del puerto serial.

La figura 17 muestra un diagrama de flujo “general” del funcionamiento del juego desde que recibe una instrucción del bluetooth por medio de la fonación sostenida para el ajuste de control de mando (Bird. Top o Flappy Bird).

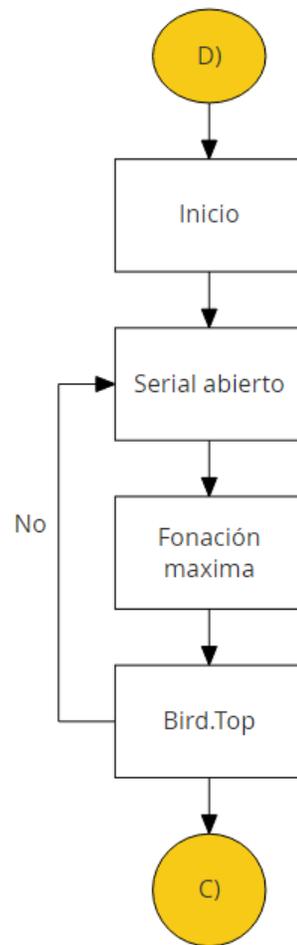


Fig 17. Diagrama de flujo inicial del juego interactivo.

3.8.2 Algoritmo de control de la aplicación.

En la implementación del algoritmo de control del sistema interactivo se considera el número de muestras que recibe de la comunicación así como una ganancia referente a la presión e intensidad de la vocalización que ajusta la escala del gráfico, la figura 18 muestra un diagrama de flujo general del ajuste de estos parámetros.

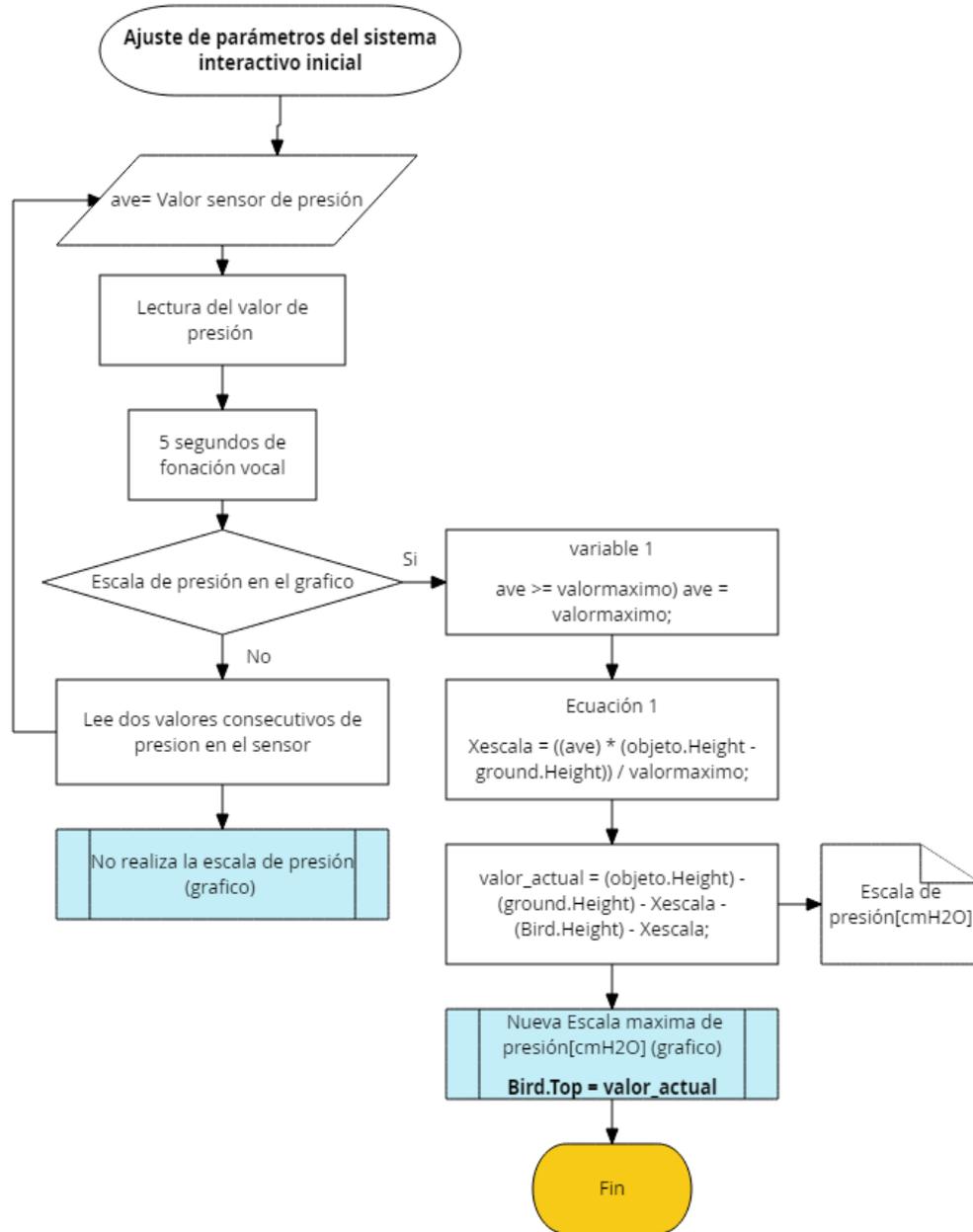


Fig 18. Algoritmo de ajuste de escala del grafico mediante la presión de fonación sostenida inicial.

La ganancia y ajuste de calibración inicial del juego se calculan a partir de la ecuación 7; se consideraron 100 muestras por segundo en cinco segundos iniciales (tiempo de muestreo) que corresponden a un tiempo de lectura de 50 ms. La fonación del usuario dependerá de la intensidad por lo que la ganancia se relacionó al ajuste escalar dentro del gráfico, entre menor sea el número de muestras obtenidas el control de mando (Flappy Bird) es más suave al momento de la fonación sostenida, esto lo hace intuitivo al usuario y le da una sensación de movimiento.

$$N^{\circ} \text{ de muestras} = \frac{\text{Tiempo de muestreo (ms)}}{50\text{ms}} \quad (7)$$

Para el cálculo de la ganancia se implementó una ecuación que determina la intensidad en la que se debe posicionar el control de mando y la relación con la presión de interés. El algoritmo para obtener la presión subglótica se considera la ecuación 7, la transformación de pixel le da una asignación de posición al objeto en el gráfico. La ecuación 8 representa el algoritmo de control del sistema para obtener la presión subglótica.

$$\text{Presión subglótica [cm H}_2\text{O]} = \frac{\text{Ganancia}_{\text{total}}(\text{ajuste de posición en el grafico})}{N^{\circ} \text{ de muestras}} \quad (8)$$

La presión subglótica está relacionada al rango de muestra cuando comienza la lectura de datos por lo que es posible calcular esta presión utilizando la posición del control de mando en el grafico (figura 19).

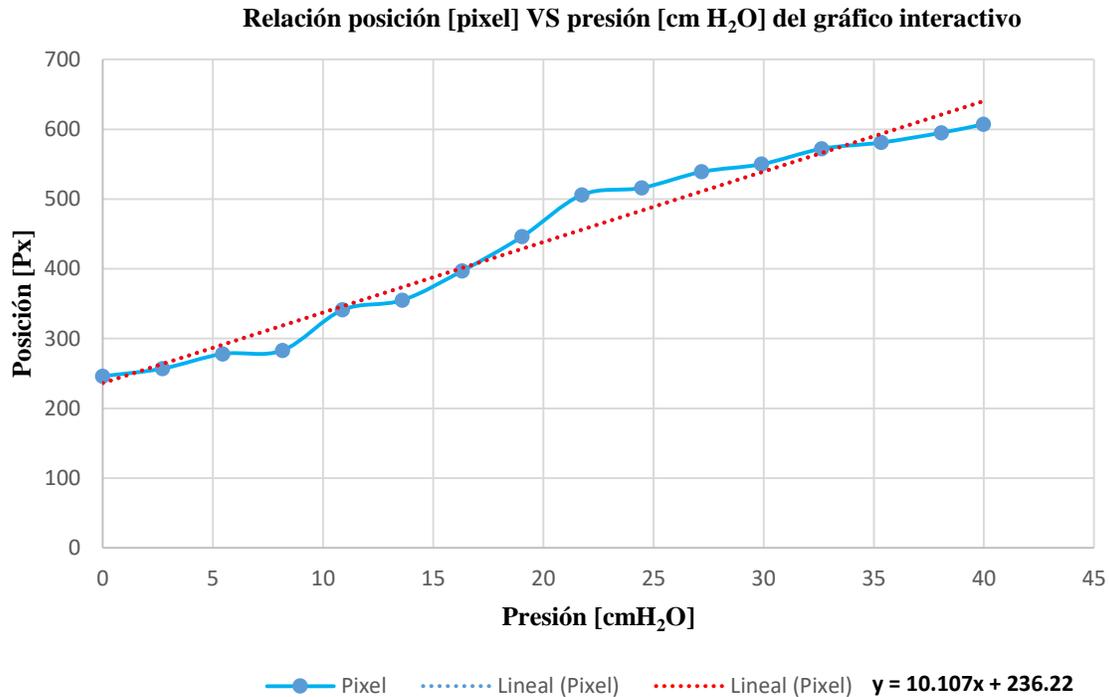


Fig 19. Relación de posición[Px] vs presión[cm H₂O] refiere al rango que ubica el control de mando (Flappy Bird) con referencia a la fonación sostenida durante 5s en el ajuste de posición máxima inicial. La línea roja (--) muestra los rangos ideales para obtener la presión subglótica. La línea de color azul (°--°) corresponde la salida de Presión subglótica con referencia al sistema medido anteriormente (figura 13), teniendo como valor máximo en posición 602Px del grafico total, la ecuación del control de mando es: Bird. Top= 602- valor máximo (fonación).

3.9 Desarrollo de la aplicación interactiva para rehabilitación.

Para el desarrollo de la aplicación del sistema interactivo de voz y registro de la señales de presión subglótica se utilizó la plataforma de Visual Studio Microsoft Visual C# Express 2010, utilizando un lenguaje de programación en C#.

El sistema tiene una interface que permite al usuario registrarse y crear su propio archivo txt. con sus datos generados durante la prueba de fonación. Cuando se captura el registro da inicio la fase de calibración o ajuste del control de mando (Flappy Bird) (figura 21) de presión

subglótica, mediante un botón de inicio del juego en donde el usuario tendrá que realizar ciertos tiempos de fonación que el especialista foniatra determinara dependiendo de la patología foniatrica, el tiempo debe ser en periodos cortos de 5s cuando se está rehabilitando el usuario. La figura 20, muestra la pantalla inicio de registro del sistema.

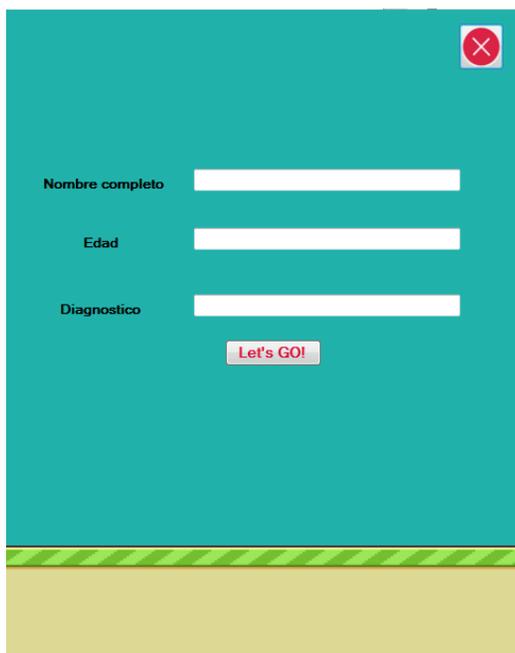


Fig 20. Pantalla de inicio del sistema interactivo de registro del paciente. Como se observa en la figura tenemos una serie de espacios en blanco que el paciente debe llenar con sus datos personales, la aplicación cuenta con un botón adicional que cierra el registro del paciente para dar inicio a la calibración de la presión subglótica e inicio del juego interactivo (botón “Let’s GO!”).

Con el objetivo que los usuarios de prueba puedan participar o hacer uso del sistema se realizó una compensación o ajuste del sensor de acuerdo a los parámetros de presión máxima de cada usuario como se muestra en la figura 21, esto da lugar a la posición de control de mando (Flappy Bird), los registros se guarda en un archivo .txt como se mencionó anteriormente dentro de una carpeta designada en un equipo asignado por el operador.

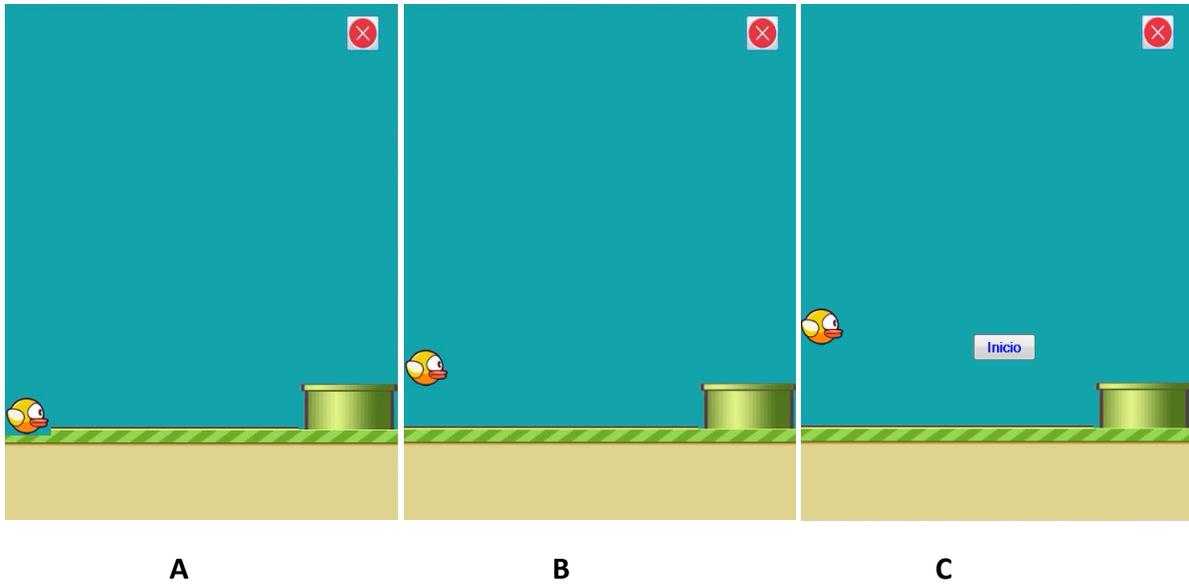


Fig 21. Ajuste de posición del control de mando (Flappy Bird). Figura A) Posición inicial del Flappy Bird de 0cmH₂O, la figura B) muestra una fonación media de 3cmH₂O, la figura C) muestra una fonación máxima 7cmH₂O.

Al termino del ajuste máximo de presión por el usuario como se mencionó anteriormente, se implementaron algunos obstáculos para que el usuario pueda realizar ciertos ejercicios de fonación máxima de forma intuitiva durante la rehabilitación de la calidad vocal y la presión subglótica, el cálculo de altura (ecuación 9) así como el ajuste de posición de los obstáculos (ecuación 10) tiene relación al porcentaje total del grafico para el ajuste del control de mando, la localización de altura (obstáculo) del grafico es 492Px estándar. Se consideraron los siguientes porcentajes de ajuste de escenas: 10%, 30%, 25%, 50%, 60%.

$$Altura_{Obstaculo} = \frac{(Y_{Localización (altura del obstaculo en el gráfico)})(Porcentaje de ajuste_{Obstaculo} (\%))}{100\%} \quad (9)$$

$$Posición_{Obstaculo} = Y_{Localización(altura del obstaculo en el gráfico)} - Altura_{Obstaculo} \quad (10)$$

En la figura 22 se muestran la secuencia de la aplicación durante el tiempo de ejercicios realizados por el usuario.

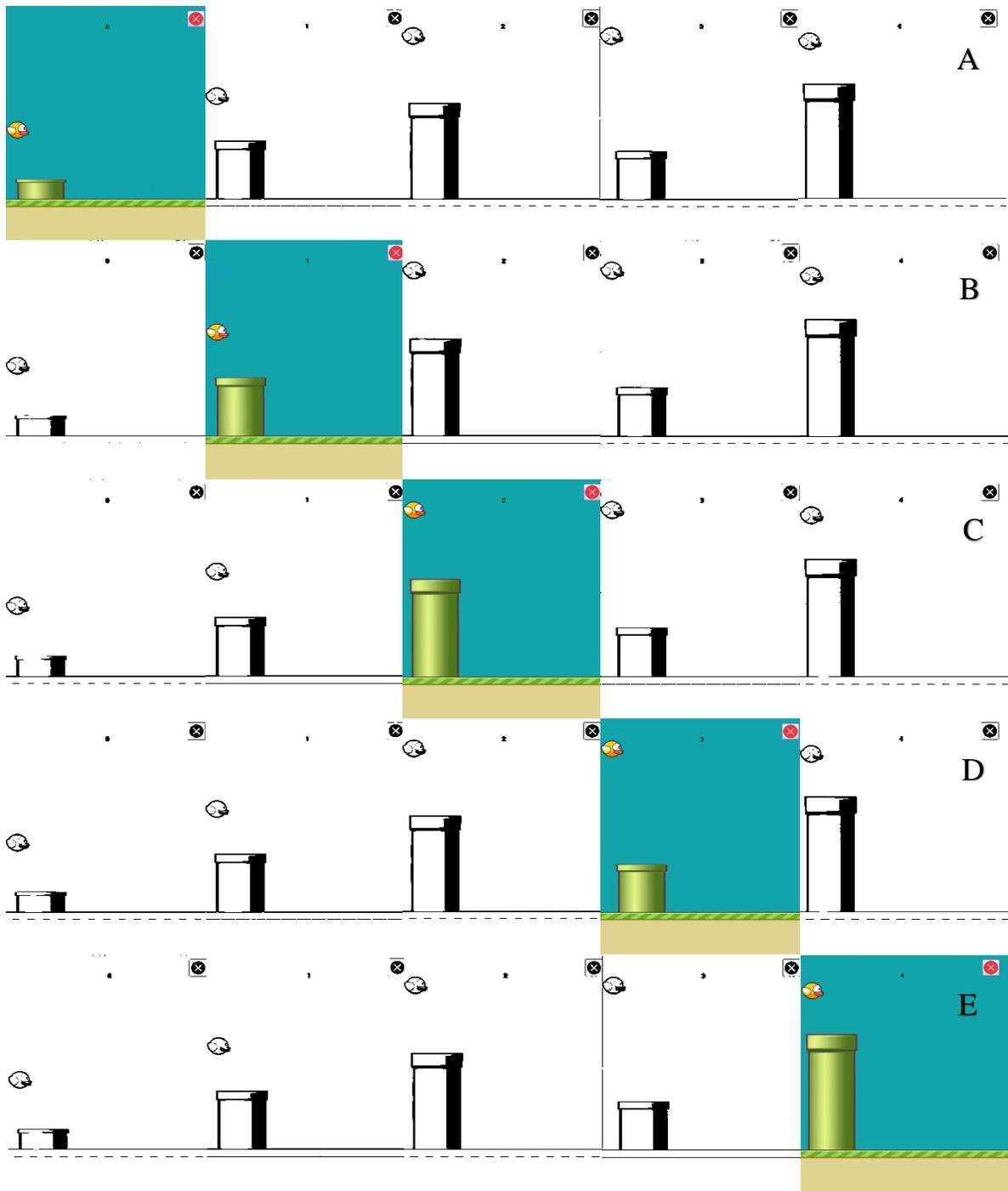


Fig 22. Secuencia de las escenas de la aplicación interactiva desarrollada a partir de cálculos dependientes de la posición del control de mando (Flappy Bird) así como la posición respecto al valor máximo de presión alcanzada por el usuario, da inicio el juego por lo que en las escenas se deben de alcanzar ciertos tiempos de fonación

necesario para poder alcanzar un objetivo, los tiempos e intervalo durante la secuencia se ajustan a la intensidad y esfuerzo de cada usuario dependiendo de la patología fonológica o no fonológica, las escenas representan varios obstáculos dependiendo de la posición máxima inicial de ajuste, la figura A) muestra el primer obstáculo con un porcentaje del 10 % la figura B) representa un obstáculo del 30% , figura C) representa un porcentaje de 50%, figura D) representa un porcentaje de 25% , la figura E) representa un porcentaje de 60%.

Capítulo 4. Pruebas

4.1 Protocolo de medición del grupo control.

En el Servicio de Foniatría del Instituto Nacional de Rehabilitación (INR) en la ciudad de México, se reclutaron voluntarios para el registro de la calidad vocal mediante la presión del aire durante la producción del habla (presión subglótica), por lo que se pretendió un grupo de 10 personas que presentaron alguna patología foniátrica y 10 personas que no presentaron ninguna patología vocal (figura 23) entre las edades de 27 a 50 años.

Se designaron los pacientes como vocalmente sanos a los que contaran con características vocales “normales”, las cuales están contempladas como tono, timbre e intensidad correspondiente a su género y edad, así como la falta de antecedentes vocales o patologías sistémicas que predispongan a las mismas.

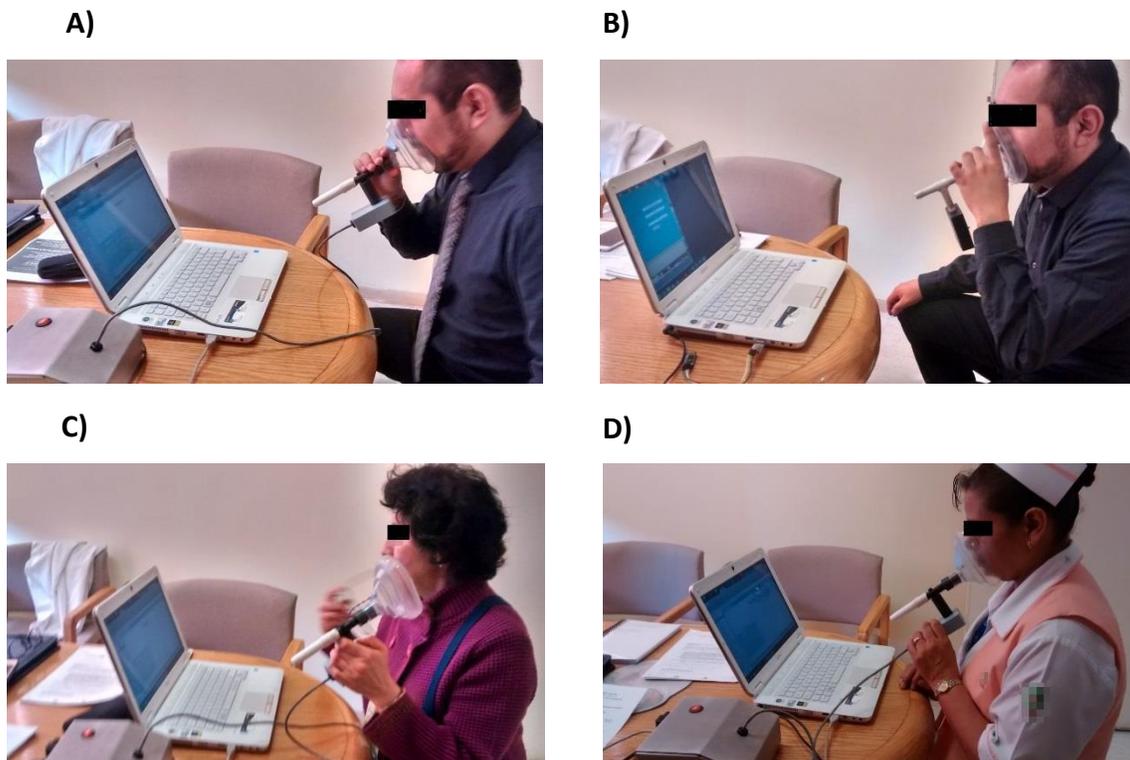


Fig 23. En la figura se muestra algunos de los usuarios que realizaron pruebas con el sistema correspondiente (registro e interactivo), en la figura A) se muestra un paciente sano (hombre) de 29 años de edad utilizando el sistema de captura y registro de la calidad vocal, la figura B) muestra el mismo usuario mencionado utilizando el sistema interactivo. Los pacientes con patología fonológica común se realizó en mujeres como se muestra a continuación. Figura C) Paciente mujer de 46 años de edad presenta una patología fonológica de Nódulos, como se visualiza se le realizó el registro de la voz, la figura D) muestra una paciente de 41 años de edad presenta una patología fonológica de disfonía de igual forma solo se realizó el registro de la calidad de voz.

4.1.1 Criterios de inclusión

En el grupo control de pacientes sanos se considera las siguientes:

- Clínicamente con voz normal.
- Edad entre los 27 años y 50 años.
- Ambos géneros.
- Sin antecedentes de patología fonológica.

En el grupo control de pacientes con patología fonológica se contempla lo siguiente:

- El paciente debe presentar solamente patología vocal.
- Edad entre los 27 años y 50 años.
- Ambos géneros.

4.1.2 Criterios de eliminación

- No haber firmado la carta de consentimiento informado. (Anexo 2)
- Que no deseen participar en el estudio.
- Padecer alguna patología de vías aéreas agudas.

En el procedimiento de captura y registro de las pruebas a los dos grupos consistió en lo siguiente:

1. Antes de comenzar la captura de las señales de voz se le explicó a cada uno en qué consistía la prueba a realizar y el tiempo de duración del registro, así como la vocal y tiempo de vocalización (5 s).
2. Se le pidió a cada usuario el nombre, edad y diagnóstico en cada registro previo a la prueba.
3. Se le indicó a cada paciente sentarse de forma cómoda, frente al dispositivo de registro.
4. Se colocó a cada paciente la mascarilla plástica que incluye (sensores de presión y un micrófono) a la altura de la boca.
5. Se le indicó al usuario vocalizar la vocal “a” durante un tiempo de 5 segundos. En la toma de datos solo se consideró la vocal /a/ por lo que la prueba en general demora 3 min en cada usuario.

Capítulo 5. Resultados

5.1 Resultados gráficos de la calidad de la voz.

Los resultados obtenidos del sistema de registro así como el sistema interactivo se realizaron de acuerdo a un protocolo respecto al grupo control en pacientes que presenta una patología fonológica como pacientes que no presentan patología fonológica.

Los resultados obtenidos del dispositivo de registro se muestran a continuación (pacientes sanos), la señal primordial es la señal de voz original (figura 24) ya que en base a esta podremos realizar un registro de la calidad vocal.

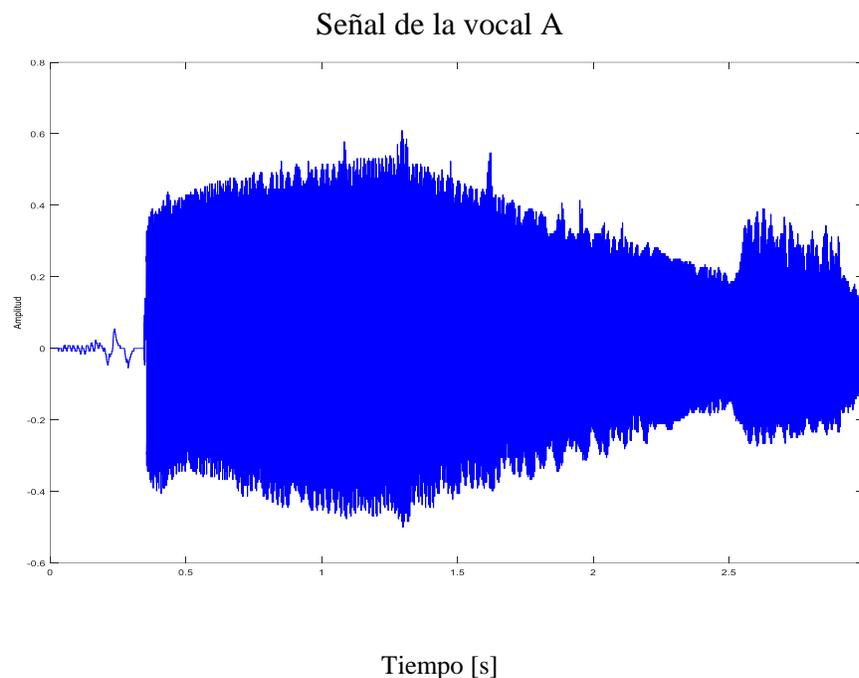


Fig 24. Señal de voz de la vocal /a/.

Ciclos de la señal de voz

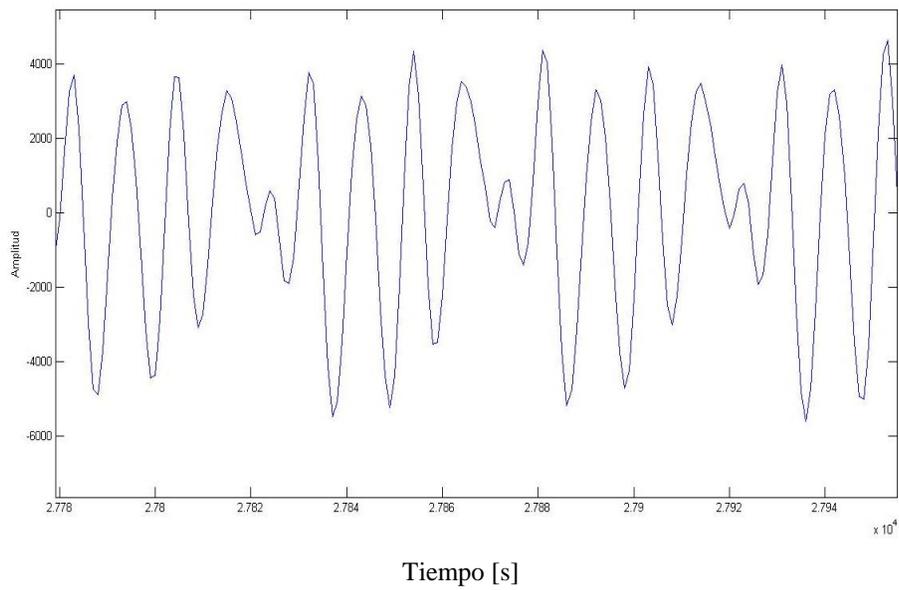


Fig 25. Ciclos de la señal de la vocal /a/.

Energía

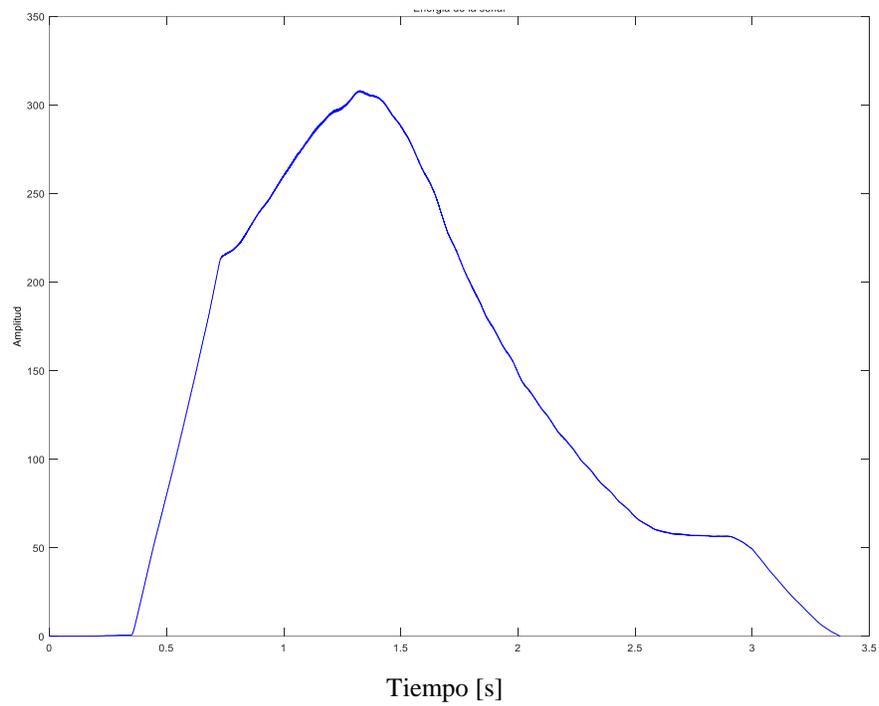


Fig 26. Energía de la vocal /a/. La energía de la señal distingue segmentos sordos y sonoros, debido a que los valores aumentan en partes sonoras.

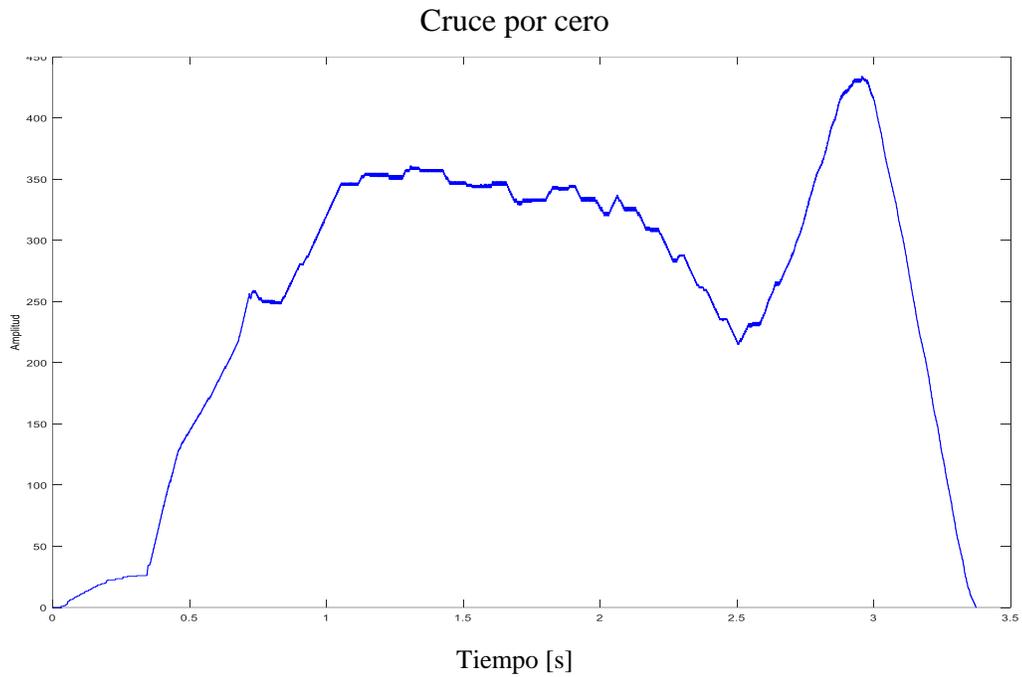


Fig 27. Cruce por cero de la vocal /a/. Estos indican el número de eventos que la señal cruza el eje de las x desde el inicio y el final en el punto 0, la amplitud mayor representa una vocal y una amplitud menor una consonante, en nuestro caso no hay ninguna consonante por lo que solo se está emitiendo la vocal /a/.

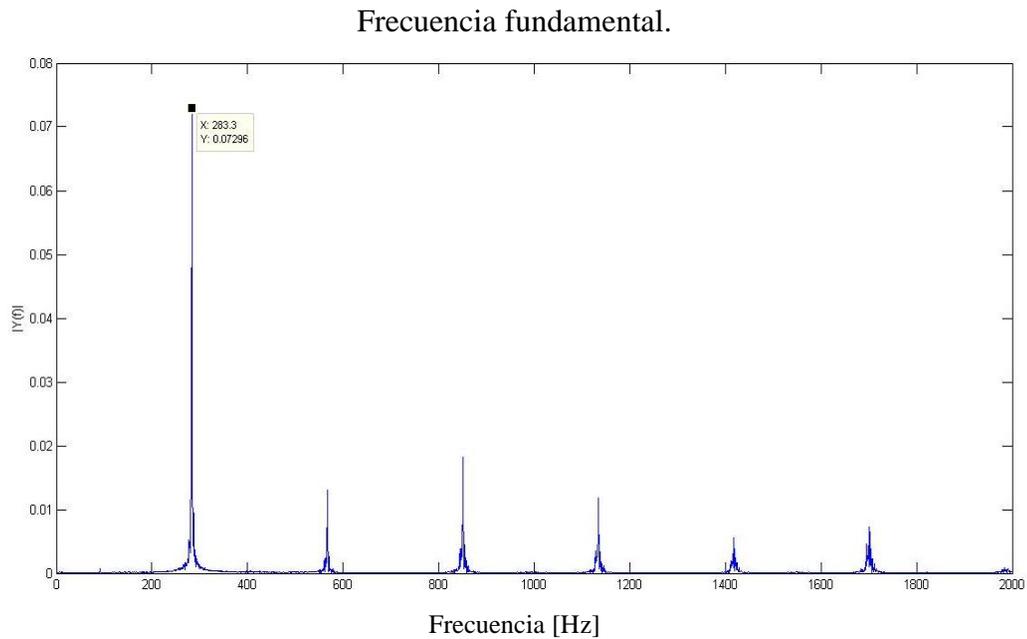


Fig 28. Frecuencia fundamental de la vocal /a/ "pitch". Se muestra una variación dependiente de la posición de los sonidos respecto al acento y entonación. El efecto de resonancia produce un énfasis en determinadas

frecuencias del espectro que se observa, denominadas “formantes”, solo consideramos la primera señal de frecuencia que está asociada a la cavidad oral y es la que nos brinda información de los armónicos.

5.1.1 Resultados gráficos No patológicos foniátricos

Las siguientes figuras muestran los resultados obtenidos en pacientes que no presentaron una patología foniátrica (sanos) así como no presentaron ejercicios de fonación previos.

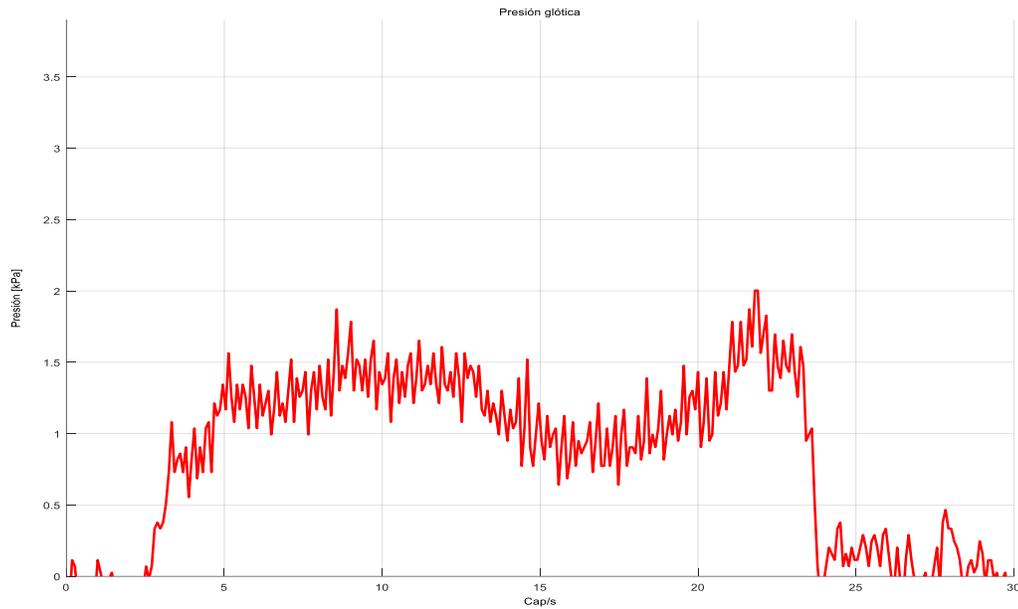


Fig 29. Presión subglótica (20.39 cm H₂O), paciente hombre de 46 años de edad.

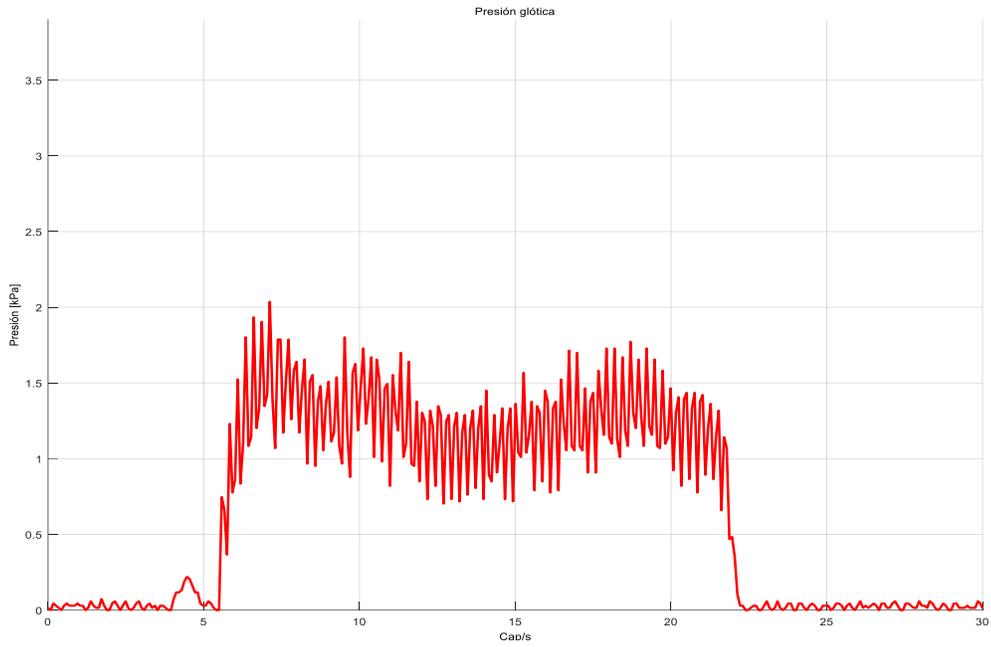


Fig 30. Presión subglótica (21.75 cm H₂O), paciente hombre de 42 años de edad.

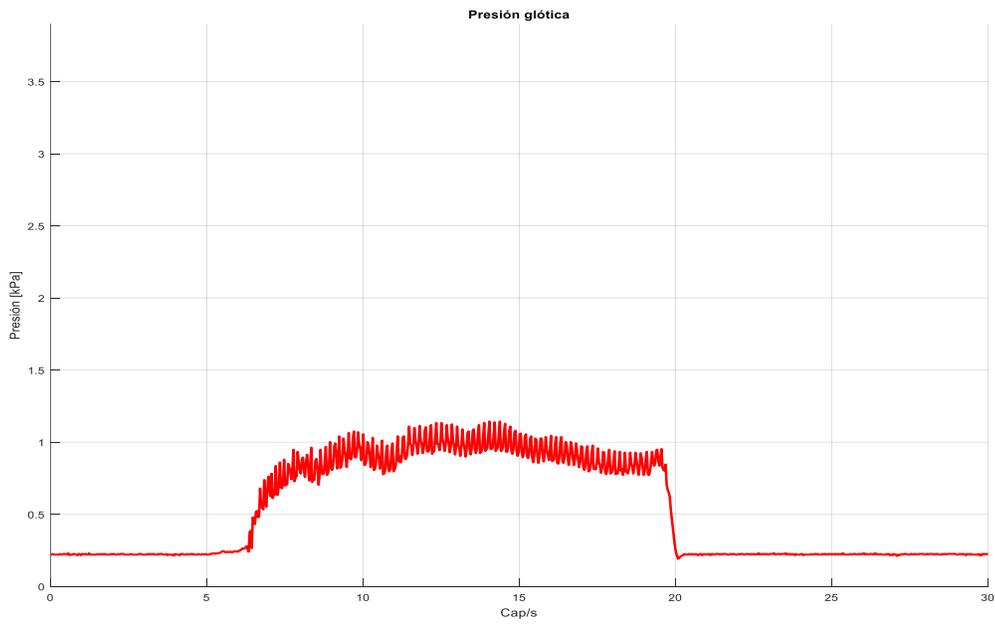


Fig 31. Presión subglótica (13.58 cm H₂O), paciente hombre de 52 años de edad.

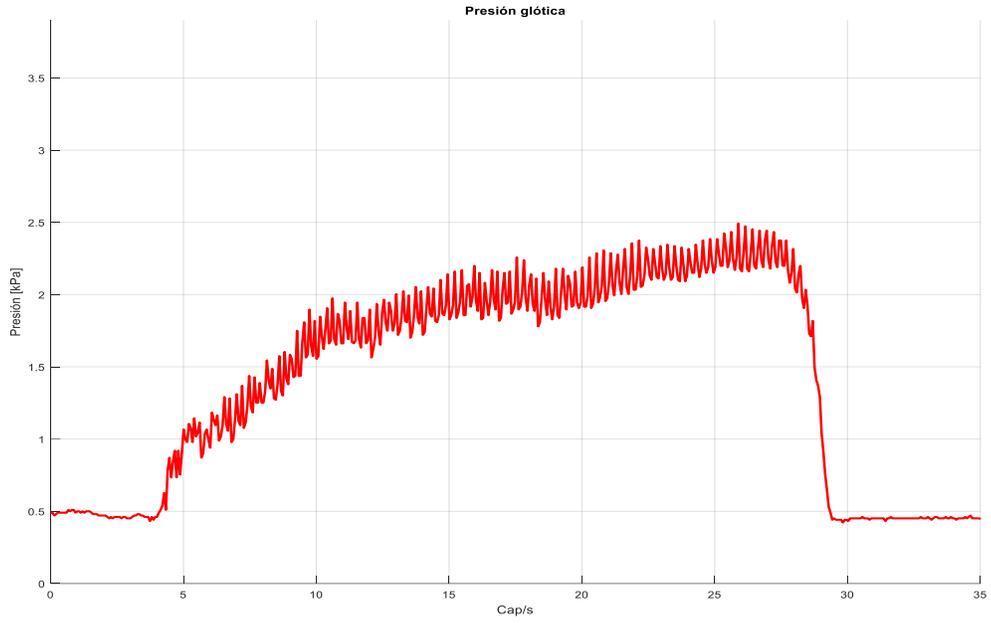


Fig 32. Presión subglótica (26.87 cm H₂O), paciente hombre de 29 años de edad.

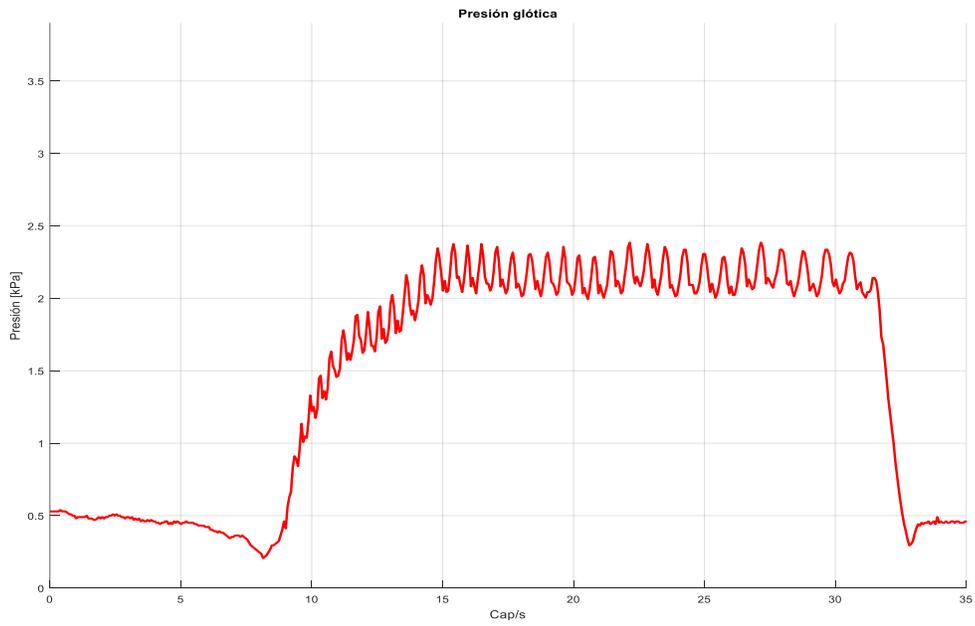


Fig 33. Presión subglótica (24.47 cm H₂O), paciente mujer de 37 años de edad.

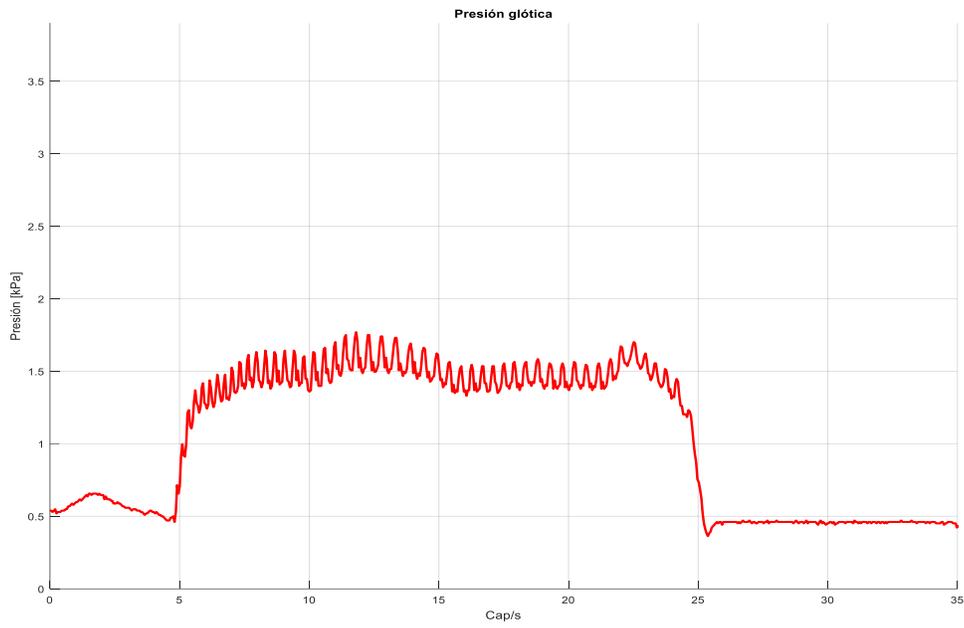


Fig 34. Presión subglótica (19.03 cm H₂O), paciente mujer de 46 años de edad.

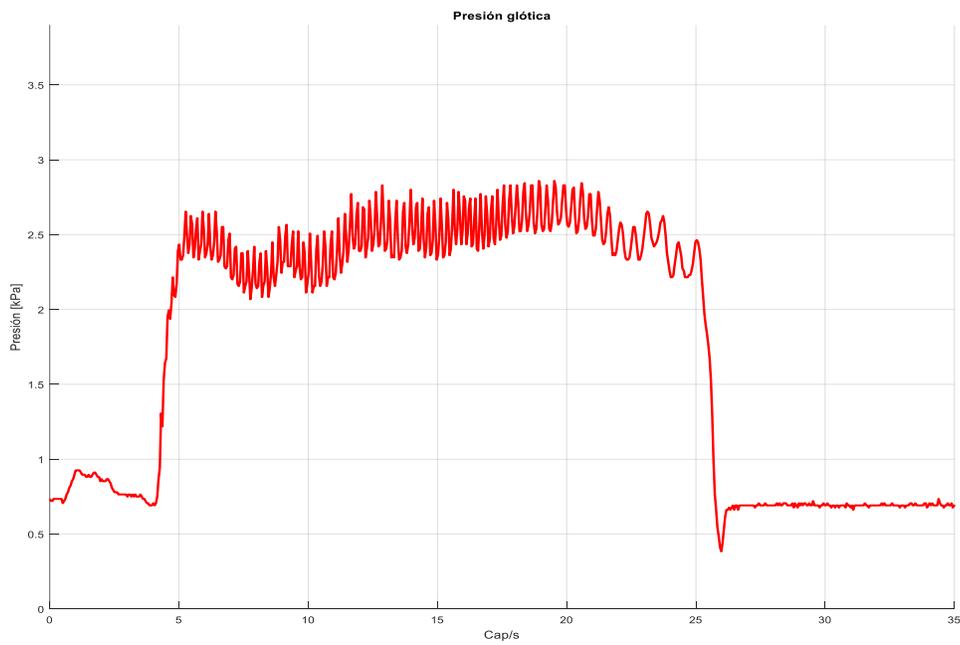


Fig 35. Presión subglótica (25.56 cm H₂O), paciente mujer de 29 años de edad.

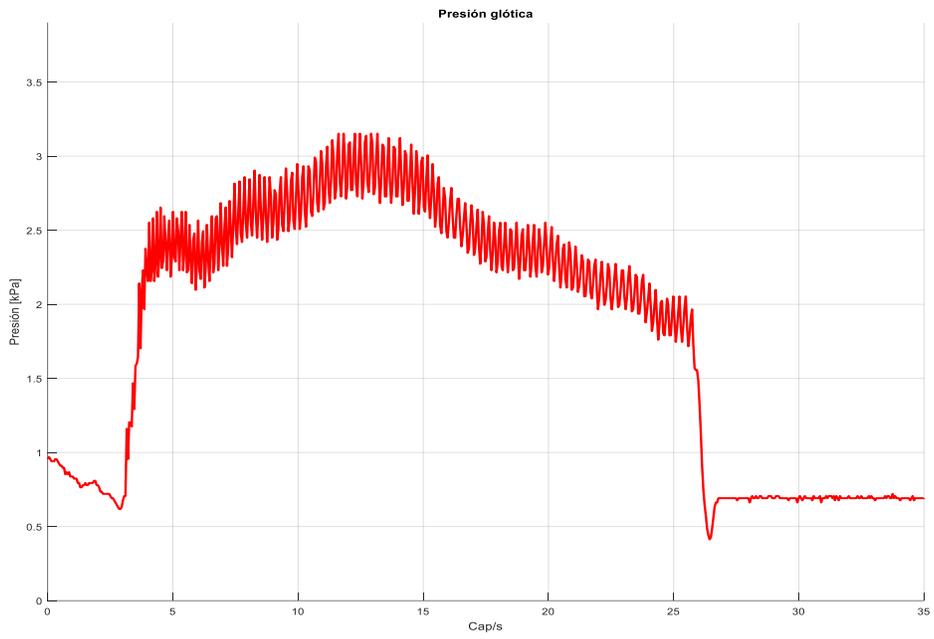


Fig 36. Presión subglótica (32.62 cm H₂O), paciente hombre de 32 años de edad.

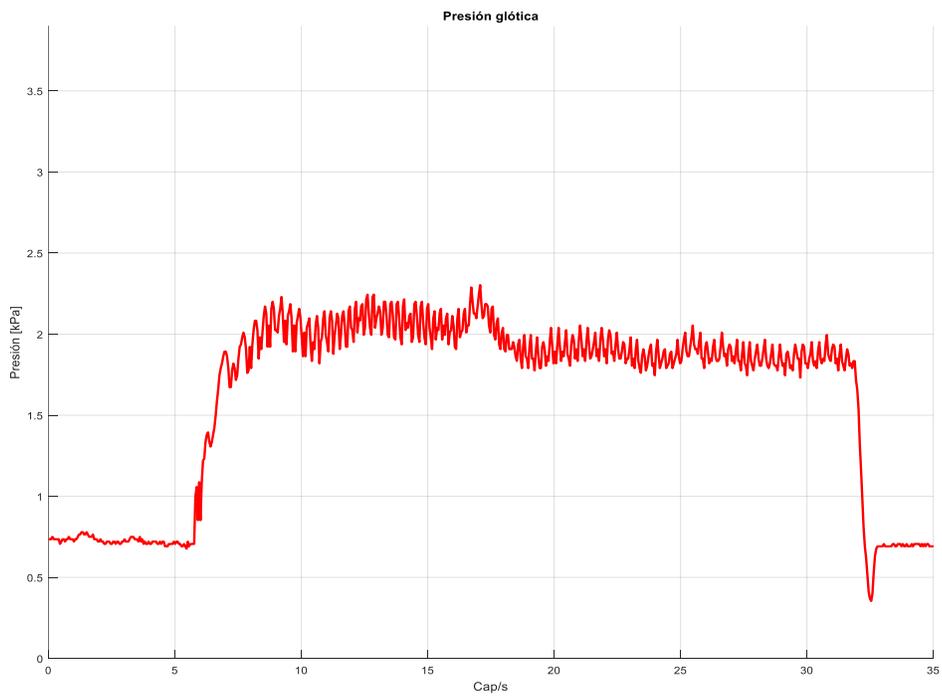


Fig 37. Presión subglótica (21.39 cm H₂O), paciente hombre de 37 años de edad.

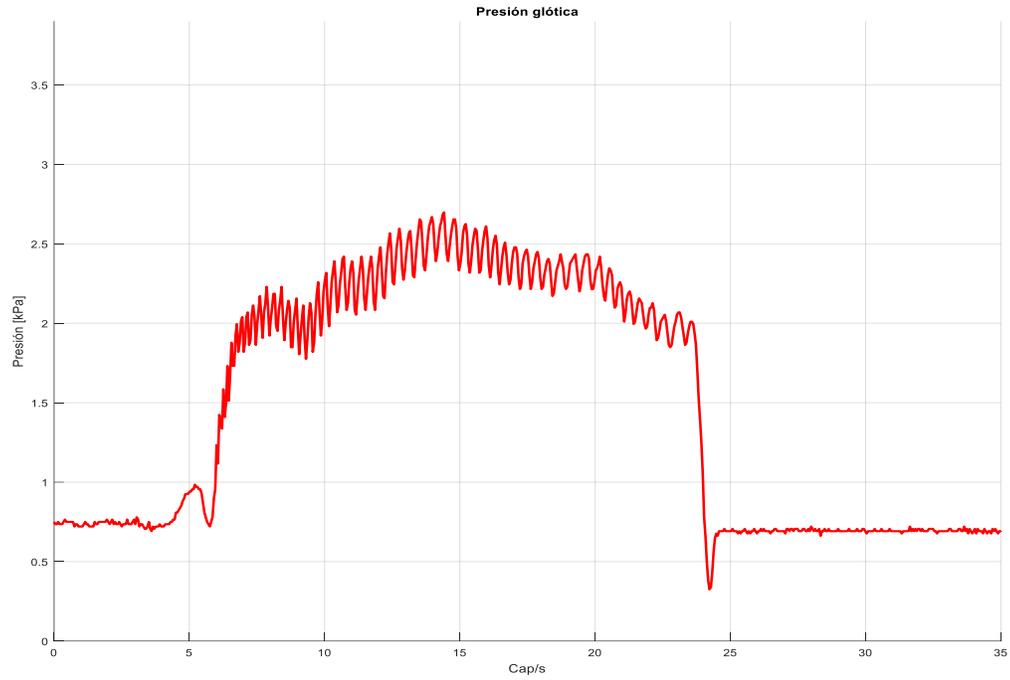


Fig 38. Presión subglótica (26.32 cm H₂O), paciente hombre de 34 años de edad.

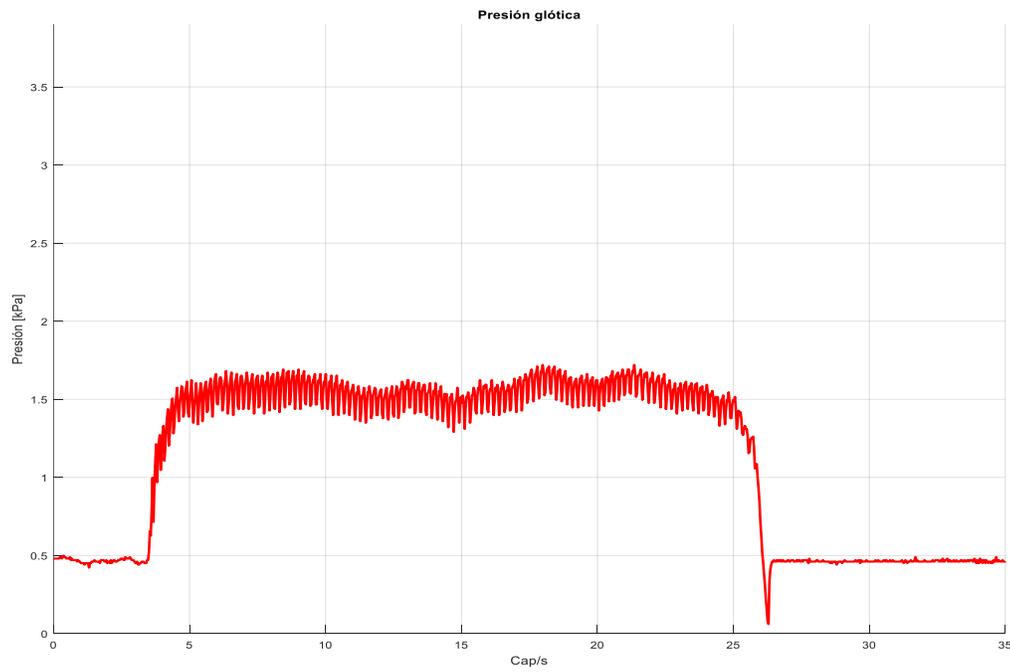


Fig 39. Presión subglótica (16.31 cm H₂O), paciente mujer de 29 años de edad.

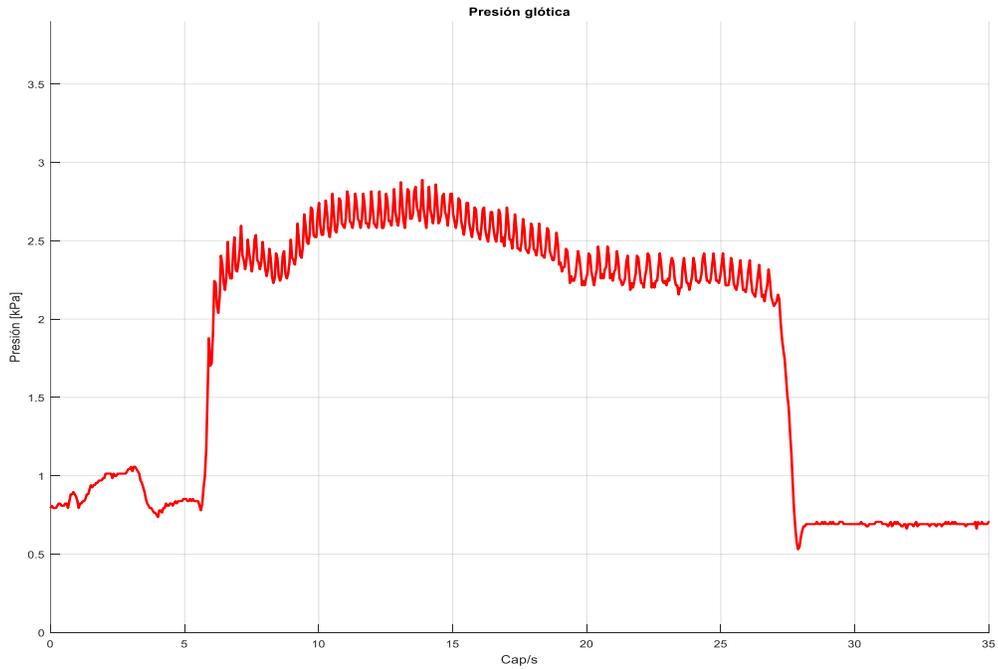


Fig 40. Presión subglótica (27.19 cm H₂O), paciente mujer de 33 años de edad.

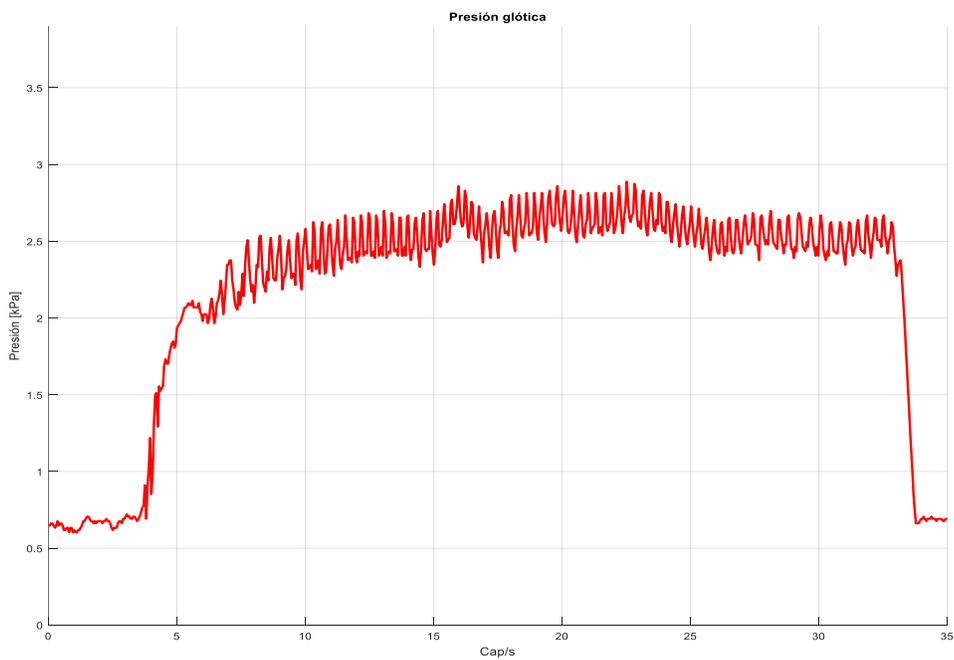


Fig 41. Presión subglótica (29.39 cm H₂O), paciente hombre de 28 años de edad.

La vocalización que realizaron los pacientes sanos resalta la intensidad y la fuerza de la fonación que atribuye al esfuerzo vocal durante las pruebas.

5.1.2 Resultados gráficos patológicos foniatricos.

Los resultados de las señales de presión subglótica [kPa] patológicos se muestran a continuación.

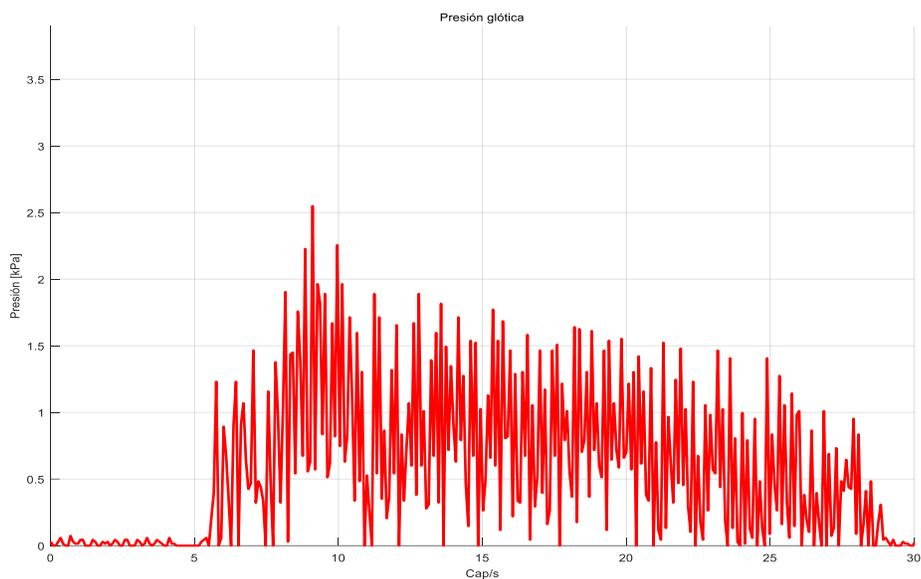


Fig 42. Presión subglótica (25.4 cm H₂O), paciente hombre de 51 años de edad que presenta una patología común (parálisis cordal del lado izquierdo cuerda vocal).

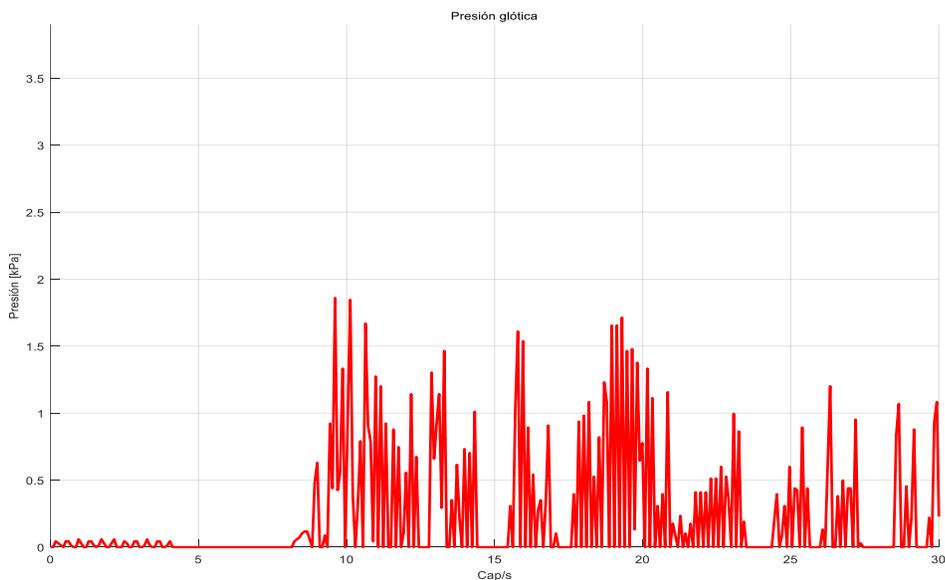


Fig 43. Presión subglótica (18 cm H₂O), paciente hombre de 47 años de edad que presenta una patología común (parálisis bilateral cordal).

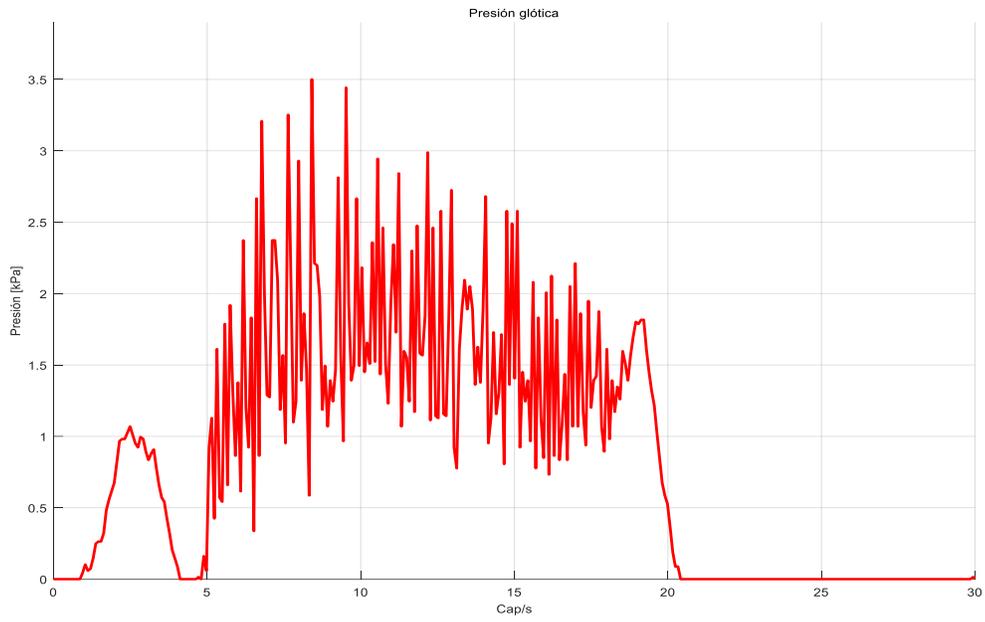


Fig 44. Presión subglótica (35.69 cm H₂O), paciente mujer de 33 años de edad diagnosticada una patología común (Disfonía espasmódica).

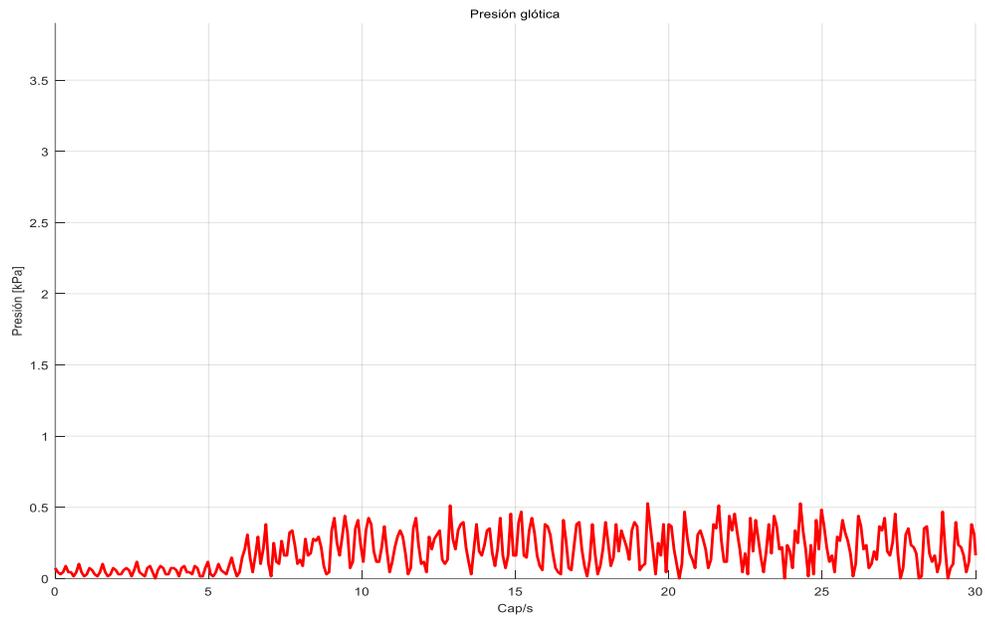


Fig 45. Presión subglótica (5.37 cm H₂O), paciente hombre de 43 años de edad diagnosticado una patología común (Disfonía Ventricular).

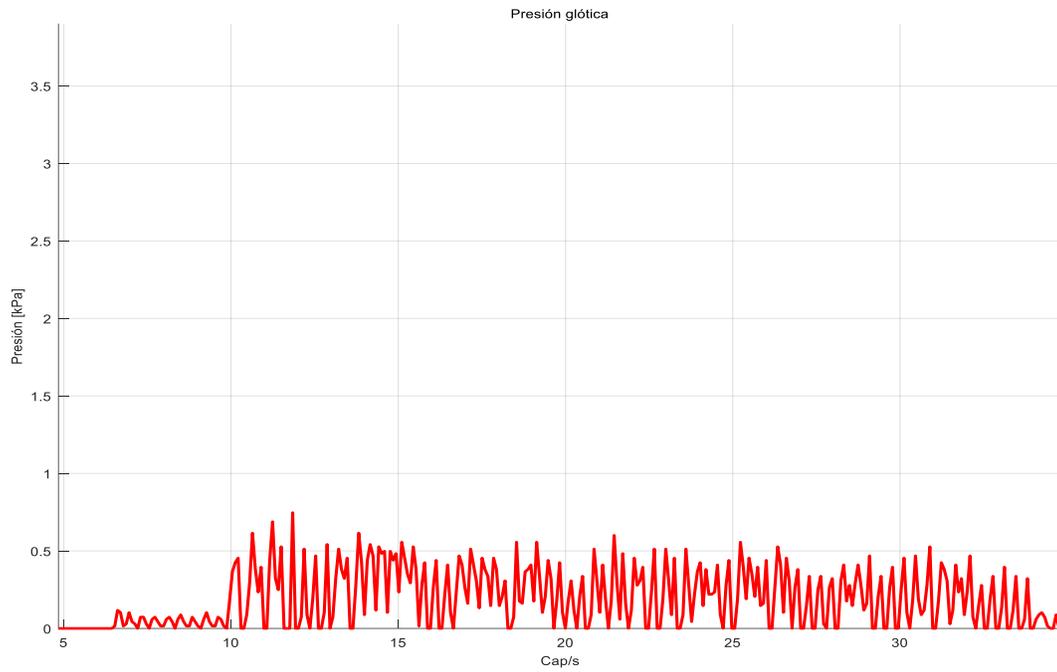


Fig 46. Presión subglótica (7.61 cm H₂O), paciente hombre de 30 años de edad diagnosticado con una patología común (Distrofia muscular cordal).

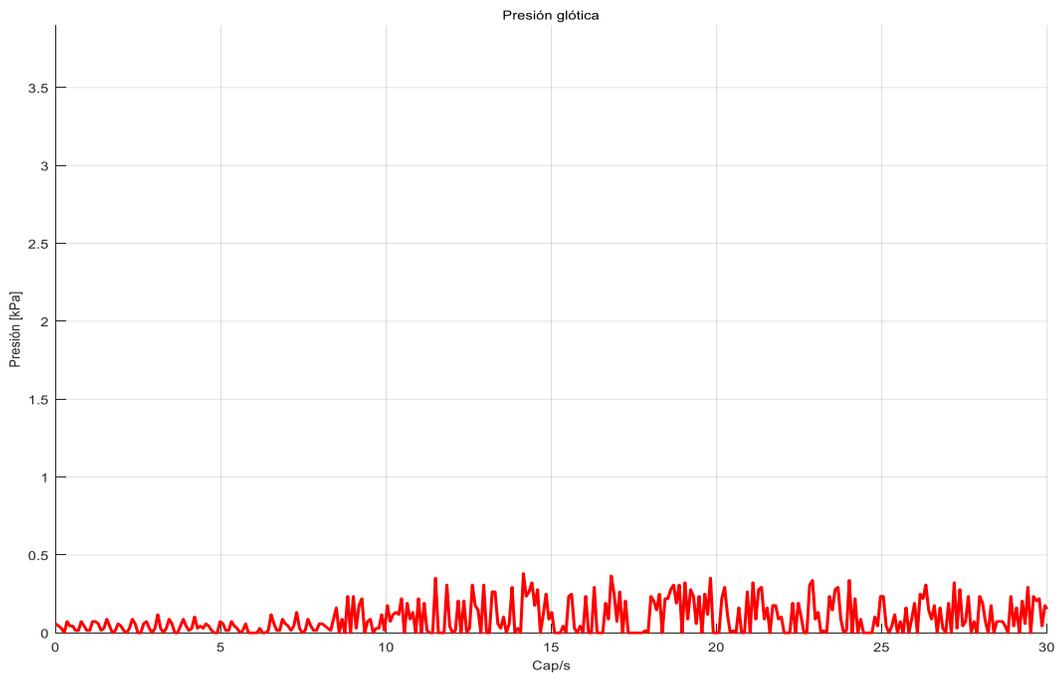


Fig 47. Presión subglótica (3.8 cm H₂O), paciente hombre de 36 años de edad presenta una patología común de (Distrofia muscular Miotónica).

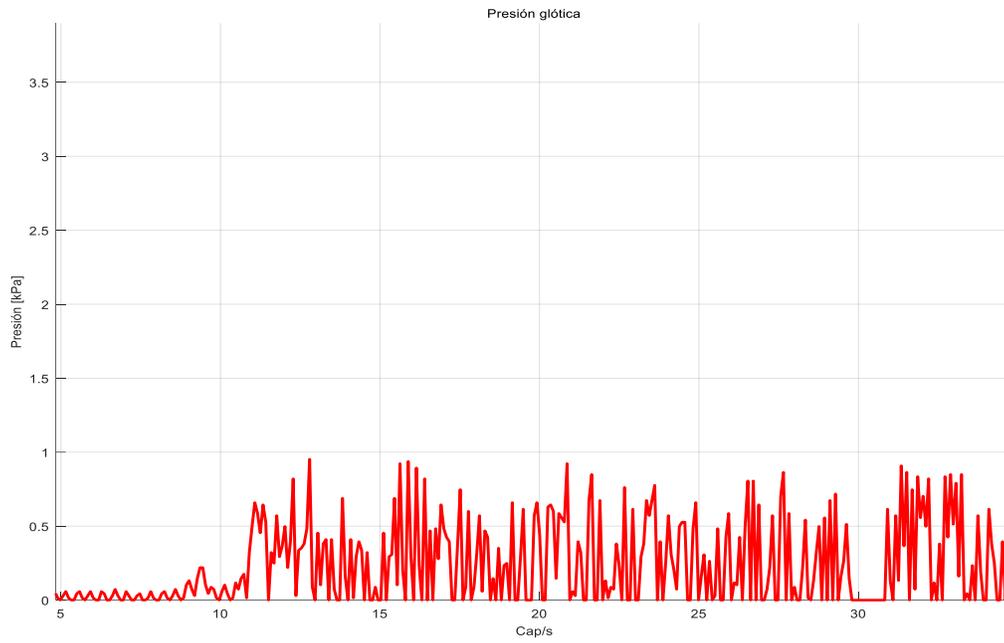


Fig 48. Presión subglótica (9.70 cm H₂O), paciente mujer de 31 años de edad que presenta una patología común (Nódulos).

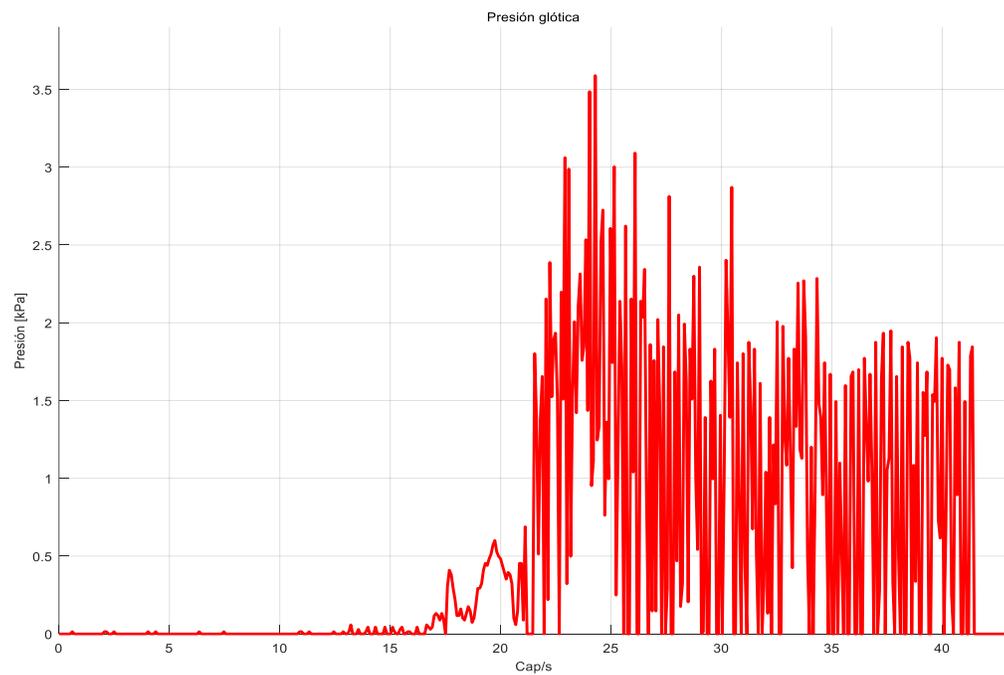


Fig 49. Presión subglótica (36.5 cm H₂O), paciente hombre de 42 años de edad diagnosticado con una patología común (Nódulos).

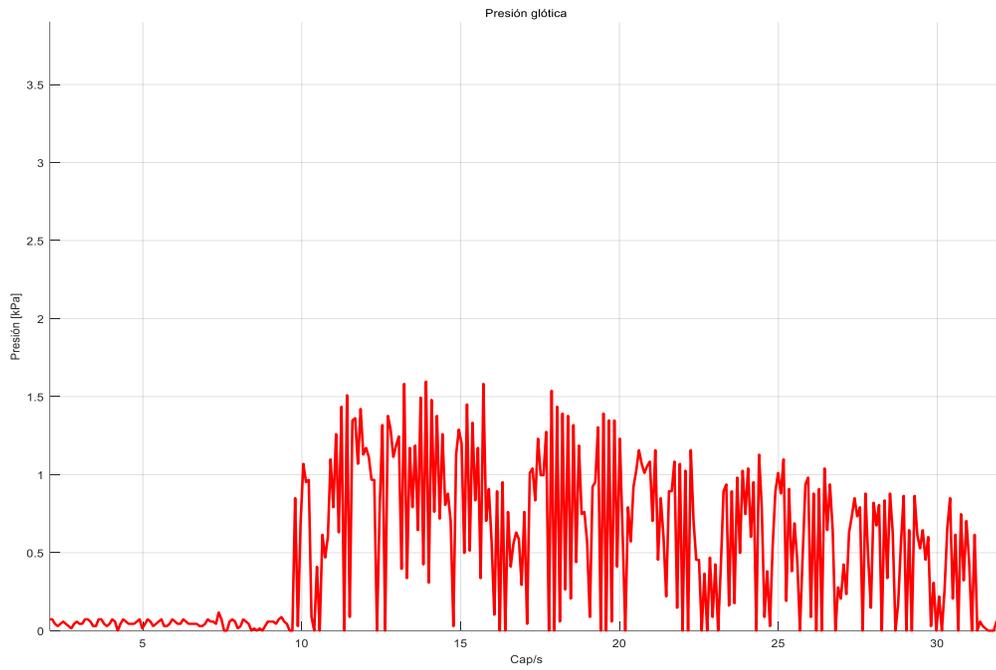


Fig 50. Presión subglótica (16.11 cm H₂O), paciente hombre de 32 años de edad que presenta una patología común (Hiperquinesia vocal).

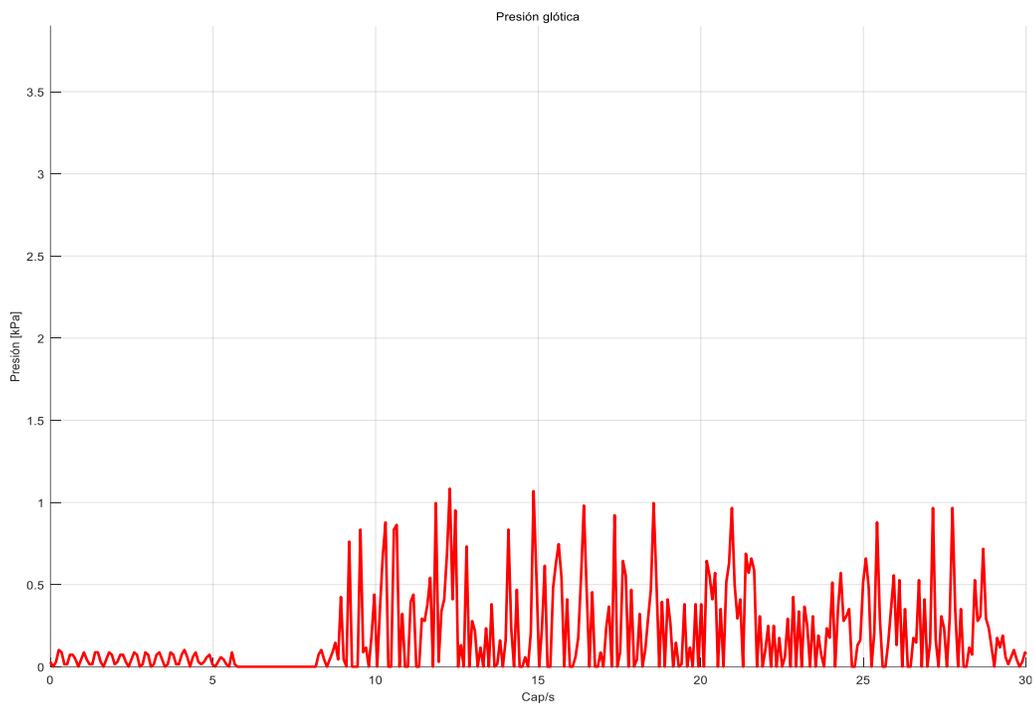


Fig 51. Presión subglótica (11.05 cm H₂O), paciente hombre de 43 años de edad que presenta una patología común (Disartria).

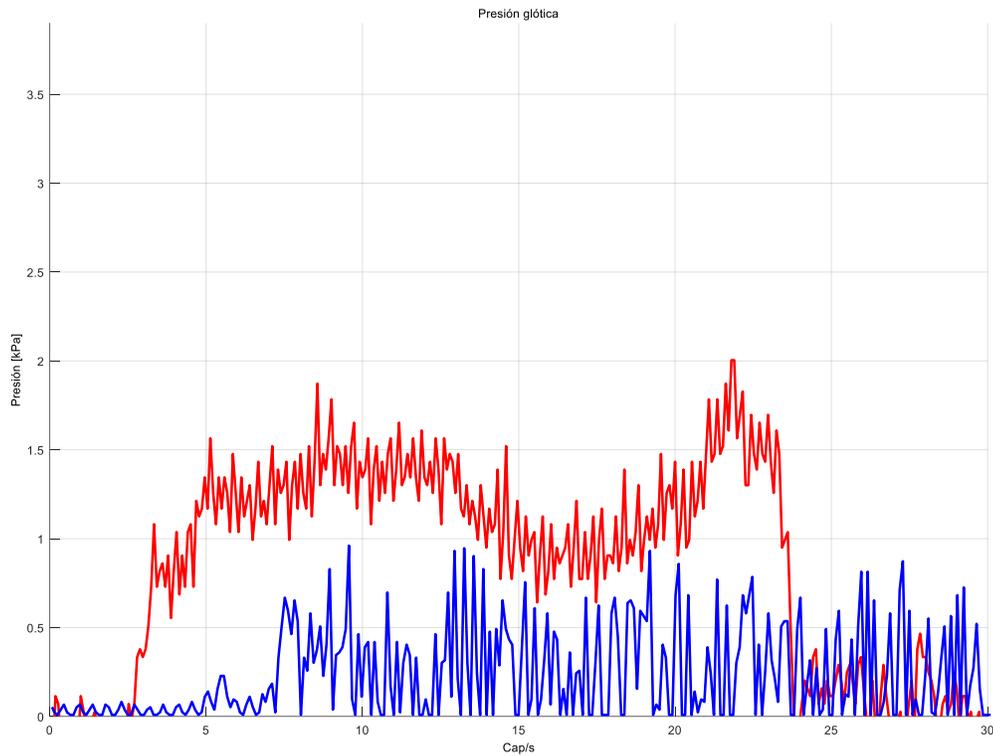


Fig 52. Se tomaron dos graficas de la presión subglótica estas representan la diferencia que existe entre un paciente patología foniatrico y uno paciente sano, se observa la diferencia de fonación sostenida así como las diferentes presiones obtenidas en cada uno. Presión subglótica (20.39 cm H₂O), paciente hombre de 46 años de edad sano (Roja). Presión subglótica (9.70 cm H₂O), paciente mujer de 31 años de edad que presenta una patología común (Nódulos) grafica azul.

Los resultados obtenidos de los pacientes patológicos así como los no patológicos muestran una diferencia en la forma de la señal por lo que si se vocaliza en el mayor tiempo posible se nota el esfuerzo e intensidad de la fonación así como la presión máxima alcanzada.

5.1.3 Resultados gráficos patológicos foniatricos del sistema interactivo.

Como se muestra en la figura 52 se observa el ajuste de posición del control de mando (Flappy Bird), como se puede visualizar se tiene una estrecha relación con las gráficas del sistema de registro esto es posible para identificar la presión subglótica [cm H₂O].

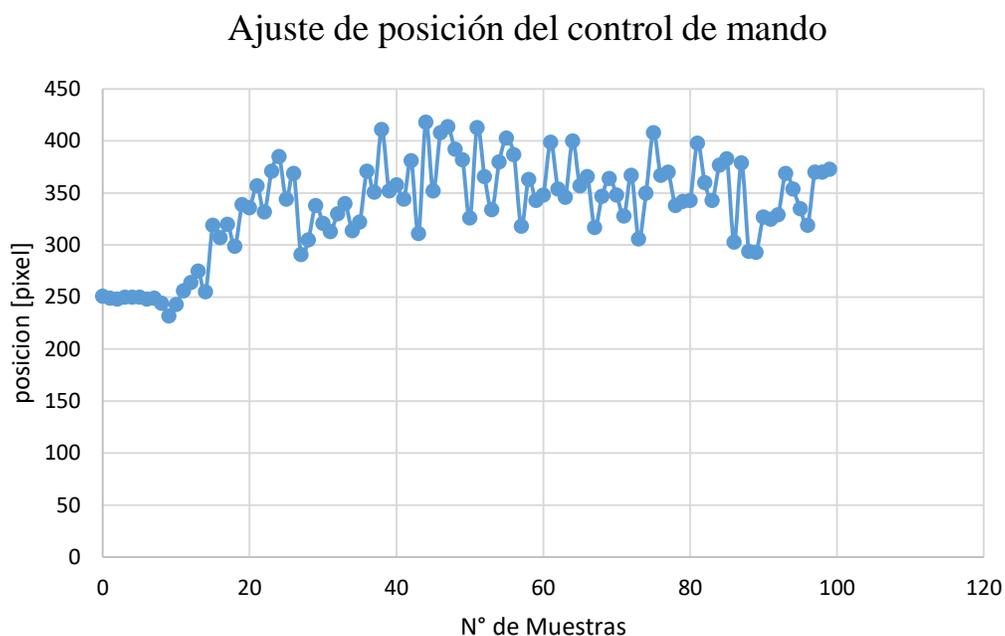


Fig 53. Ajuste de presión máxima por medio de posición [pixel], anteriormente se explicó cómo se obtiene la presión subglótica en base a la posición vs presión. Presión subglótica (17.12 cm H₂O), paciente mujer de 43 años de edad que presenta una patología común (Disfonía).

5.1.4 Resultados gráficos No patológicos foniatricos del sistema interactivo.

Como se muestra en la figura 53 al igual que las gráficas del sistema de registro existe una similitud de la posición, la señal es más suave por lo que el número de muestras que captura es menor por segundo, esto tiene que ver con la comunicación que se emplea.

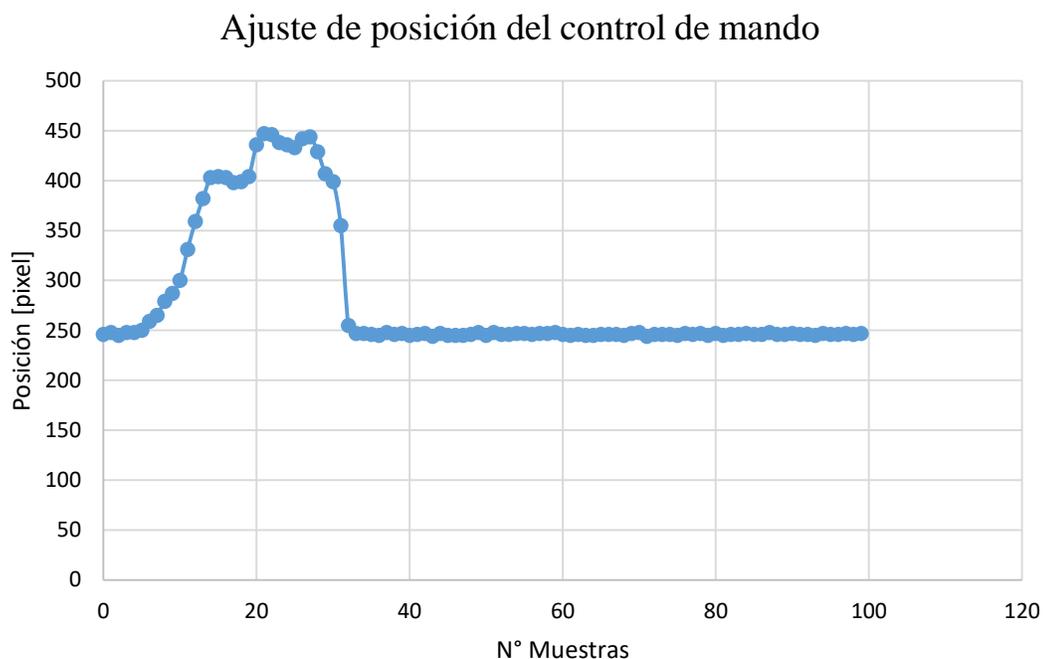


Fig 54. Ajuste de posición de control de mando de un paciente sano por lo que la posición que se ve es mayor que la que presenta un paciente patológico foniatrico, la presión máxima (P_{sub}) es mayor. Presión subglótica (20.31 cmH₂O), paciente hombre de 35 años de edad.

Capítulo 6. Conclusiones y perspectivas

Los resultados obtenidos se consideran satisfactorios ya que cumplen con los los objetivos específicos, cabe mencionar que el dispositivo que se desarrollo es económico y su uso es de fácil manejo. Se ha demostrado, que a partir de la señal eléctrica de la voz, es posible el registro de la presión subglótica durante la fonación sostenida en individuos con voces no patológicas. En pacientes con disfonía, se observó que el sistema de registro implementado es útil para el registro y análisis cuando se captura una máxima oclusión glótica durante la fonación sostenida. Sin embargo, embargo, no es posible asegurar que en todos los pacientes que presenten diferentes patológicas foniátricas sea el tiempo de terapia durante la fonación ya que depende del esfuerzo vocal de cada individuo cuando se le está rehabilitando en la presión máxima alcanzada.

Los resultados del análisis de la voz y del tiempo máximo de fonación así como los estudios de la valoración cuantitativa y cualitativa de la voz, sugieren un efecto positivo de la calidad de la voz, promoviendo un incremento del tono y volumen durante la fonación que deriva en una mejor oclusión glótica, mayor control de la fonación y mayor calidad vocal.

Requiere que la persona se le motive a pronunciar una vocal de manera sostenida durante el mayor tiempo posible, manteniendo constantes la tonalidad y el volumen en las grabaciones de voz durante 5 segundos. Las grabaciones que se realizaron durante las diferentes pruebas constituyen de manera indirecta una terapia de voz así que pueden considerarse por sí misma un tratamiento de rehabilitación, de modo tal, que no debe despreciarse su probable contribución a la tonificación muscular y al incremento de la capacidad pulmonar, misma que influye en el tiempo máximo de fonación.

Los tratamientos con terapias de voz suelen prolongarse por periodos de tiempo mayores, por tal motivo puede suponerse mínima la contribución de los ejercicios de fonación a los diferentes resultados. La basta y creciente diversidad de técnicas para la evaluación de la calidad vocal, así

como el gran número de variables implicadas en la metodología de cada una de ellas, es importante identificar los parámetros medibles (“*Pitch*”, energía, cruce por ceros y presión subglótica) que aportan mayor información de la voz.

En este trabajo se ha logrado analizar el comportamiento de la voz por medio de un sistema de análisis y registro, se ha demostrado que a partir de la señal eléctrica de la voz es posible estimar una valoración vocal por medio de diferentes procesos de análisis en personas de diferente edad y sexo, sin embargo, se requiere de un estudio más extenso para conocer las limitantes relacionadas con la presión subglótica, la señal se adquirió de manera indirecta en cada persona, se eligió aplicar el estudio de la presión glotal durante su máxima oclusión durante la fonación sostenida.

Se propone como trabajo futuro, evaluar la utilidad de registro por medio de un Electróglografo (EGG), podría proporcionar información más precisa sobre la presión subglótica y el número de eventos que suceden cuando se obtiene una energía total de la voz. Se requiere un estudio más extenso, en el que se pueda diferenciar el efecto fisiológico producido por efecto de ejercicios de fonación.

Referencias

- [1] M. Guzman, A. M. Laukkanen, P. Krupa and J. Hoácek. *Vocal tract and glottal function during and after vocal exercising with resonance tube and straw*. Journal of voice Vol.27 N° 4, 2013.
- [2] M. Weeks. *Digital signal Processing using MATLAB and Wavelets*. Electrical Engineering series, 2007.
- [3] Guyton, tratado de fisiología médica, 12 ediciones, 2011.
- [4] Tortora, Derrickson, Principios de anatomía y fisiología, 11 edición, 2010.
- [5] Border, G.J. y Harris, K.S, speech science, Physiology, acoustics and perception of speech, 1980.
- [6] J. J. Jiang, Y. Zhang, and C. McGilligan. *Chaos in Voice, From Modeling to Measurement*. Journal of Voice, Vol. 20. January 5, 2005.
- [7] G. Far. *The voice source in connected speech*. Department of speech and hearing, KTH speech communication 22. October 1997.
- [8] F. G. *Preliminaries to analysis of the human voice source*. STL-QPSR.1982; 23:1–27.
- [9] Sundberg J, Titze I, Scherer R. *Phonatory control in male singing: a study of the effects of subglottal pressure, fundamental frequency, and mode of phonation on the voice source*. Journal of Voice. 1993.
- [10] B. Gold and N. Morgan. *Speech and Audio Signal Processing*. John Wiley and Sons, 2000.

- [11] R. L. McAulay. *Speech Analysis/Synthesis Based on a Sinusoidal Representation*. IEEE Transaction on acoustics and signal processing, Vol Assp 34, N° August 1996.
- [12] S. Bjorklund and J. Sundberg. *Relationship Between Subglottal Pressure and Sound Pressure Level in Untrained Voices*. Article in press journal of voice, March 2015.
- [13] J. Sundberg, R.Scherer, M. Hess, F. Muller, and S. Granqvist. *Subglottal Pressure Oscillations Accompanying Phonation*. Article Journal of Voice, Vol 27, March 19, 2013.
- [14] T. Arai, Y. yoshida. *A Method of Speech Signal Analysis by Zero-Crossings*. Article Faculty of Science and technology, Sophia University, tokyo ,Japan. August, 1992.
- [15] Bachu R.G., Kopparthi S., Adapa B., Barkana B.D. *Voiced/Unvoiced Decision for Speech Signals Based on Zero-Crossing Rate and Energy*. Department of Electrical Engineering,University of Bridgeport. Springer, 2010.
- [16] P. S. Phillips, S. P. Hirani, and R. Epstein. *Peak flow and voice Pathology*. Article in press Journal of Voice, Vol 23, London, January 1, 2003.
- [17] M. R. Hoffman, C. D. Baggott, and J. Jiang. *Reliable time to estimate subglottal Pressure*. Article Journal of Voice, Vol.23, Madison, Wisconsin, september 13, 2007.
- [18] A. M. Laukkanen and J. Sundberg. *Peak to Peak Glottal Flow Amplitude as a Function of Fo*. Article Journal of Voice, Vol.22, january 2, 2007.
- [19] G. Muhammad. *Extended Average Magnitude Difference Function Based Pitch Detection*. The International Arab Journal of Information Technology, 2 April 2011.
- [20] M. Rothenberg (2014, August, 08) Glottal Enterprise Incorporated (1ts edition) 110(4). Disponible en: www.glottal.com.

[21] H. Group (2015, March, 12) The Phonatory Aerodynamic System (PAS) (1st edition) 243(34). Disponible: <http://www.pentaxmedical.com/pentax/en/99/1/Phonatory-aerodynamic-system-PAS-Model-6600>.

[22] J. Jiang, C. Leder and A. Bichler. *Estimating subglottal pressure using incomplete airflow interruption*. Laryngological. 2006, 116(1):89-92.

[23] G. Fat, A. Kruckenber, S. Hertegard and J. Liljencrants. Sub-and Supraglottal pressures speech. Department of Logopedics and Phoniatrics, unpublished.

[24] D. J. Beebe, "Signal conversion," en *Biomedical Digital Signal Processing*, W. J. Tompkins, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1993, ch. 3, pp. 61–74.

[25] A. L. Rosenthal, S. Y. Lowell, and R. H. Colton. *Aerodynamic and Acoustic Features of Vocal Effort*. Article Journal of voice, vol.28, September, 2013.

[26] B. H. Story. *An overview of the physiology, physics and modeling of the sound source for vowels*. Department of Speech and Hearing Sciences, University of Arizona. Acoust. Sci. & Tech. April 2002.

[27] J. A. Díaz, C. Sapienza, H. B. Rothman, Y. Natour. *Robust algorithm for the detection of the voice fundamental frequency based on the spectrogram*. Revista ingeniería UC. Vol.10, Diciembre 2003.

[28] S. Sheela. *Laryngeal aerodynamic analysis of vocal nodules*. Article Journal of Laryngology and Voice, Vol.3, June 2013.

[29] L. R. Rabiner and R. W. Schafer. *Digital Processing of Speech Signals*. Prentice-Hall Inc. 1978.

- [30] C. Ignacio, N. Faustino, F. Secundino. *Patología de la voz*. Libro de la sociedad de Otorrinolaringología y patología cérvico-facial 1° edición. Barcelona 2013.
- [31] P. John G., M. Dimitris G. *Digital Signal Processing*. 3° edición, Prentice-Hall international, INC.
- [32] B. jacob, S. M. Mohan and H. Yiteng. *Speech processing*. Springer Handbook 2008.
- [33] G. Fant. *Acoustic Theory of speech production*. Editorial The Hague, Netherlands, First printing 1960.
- [34] F. Alipour and R. C. Scherer. *Pressure-flow relationships during phonation as a function of adduction*. Article Journal of voice, Vol.11, june 1997.
- [35] T. Bäckström, P. Alku and E. Vilkmán. *Time–Domain Parameterization of the Closing Phase of Glottal Airflow Waveform From Voices Over a Large Intensity Range*. IEEE Transactions on Speech and Audio Processing, Vol.10, N° 3. March 2002.
- [36] J. Guónason, D. D. Mehta, T. F. Quatieri. *Close phase estimation for inverse filtering the oral airflow waveform*. IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP) 2014.
- [37] P. Yu, M. Ouaknine, J. Revis and A. Giovanni. *Objective Voice Analysis for Dysphonic Patients: A Multiparametric protocol including acoustic and aerodynamic measurements*. Journal of voice Vol.15, N° 4. June 1, 2001.
- [38] P. M. Balata, H. J. Silva, L. Araújo and G. Oliveira. *Electrical activity of extrinsic laryngeal muscles in subjects with and without dysphonia*. Journal of voice, March 2014.
- [39] J.S. Damico, N. Muller and M. J. Ball, “*The handbook of language and speech disorder*” N. Heinemann Medical Book. London, 1998.

- [40] R. Luchsinger, G.E. Arnold. *Voice Speech Language: Clinical Communicology- Its Physiology and Pathology*. Boston: Wadsworth 1 st U. S edition; 1965; p.118.
- [41]B. Fritzell, J. Gauffin, B. Hammarberg, I. Karlsson, J. Sundberg. *Measuring insufficient Vocal fold closure during Phonation*. En: Anders A, Si F, Erick J, Johan S, editors. SMAC 83 Proceeding of the Stockholm Music Acoustics Conference. Vol 1. N° 46:1 ed.Royal Swedich Academy of Music, Stockholm Sweden; 1985. P.175-186.
- [42] S. Hertegard, J. Gauffin, P. A. Lindestad. *A Comparison of Subglottal and Intraoral Pressure Measurements during Phonation*. *J. Voice*.1995; 9(2):149-155.
- [43] L. Boves and B. Cranen. *Spectral and temporal difference between glottal flow pulses as a function of speaker a supra-glottal load*. Institute of Phonetics, Nijmegen University (1933).Pag 87-102.
- [44] K. Kondo. *Subjective quality measurement of speech. Signal and communication technology*, Springer-Varlag Berlin Heidelberg, 2012, p.7-15.

Anexos. Especificación general de los dispositivos usados

Anexo 1 Especificaciones del sensor de presión Honeywell.

Anexo 2 Formato de consentimiento informado y aviso de privacidad.

Freescale Semiconductor
Data Sheet: Technical Data

Document Number: MPxx5004
Rev. 12.1, 05/2015



MPxx5004, 0 to 3.92 kPa, Differential and Gauge, Integrated Pressure Sensor

Freescale's MPxx5004 series piezoresistive transducer is a state-of-the-art monolithic silicon pressure sensor designed for a wide range of applications, but particularly those employing a microcontroller or microprocessor with A/D inputs. This sensor combines a highly sensitive implanted strain gauge with advanced micromachining techniques, thin-film metallization, and bipolar processing to provide an accurate, high level analog output signal that is proportional to the applied pressure.

Features

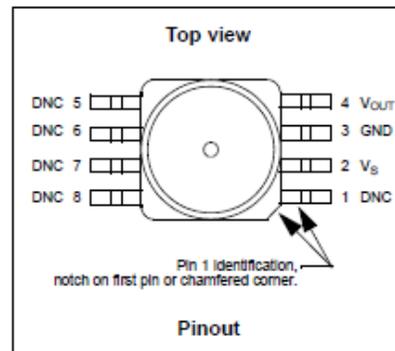
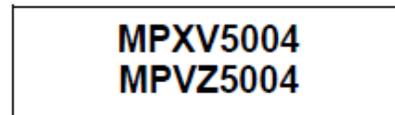
- 1.5% maximum error for 0 to 100 mm H₂O over +10 °C to +60 °C with autozero
- 2.5% maximum error for 100 to 400 mm H₂O over +10 °C to +60 °C with autozero
- 6.25% maximum error for 0 to 400 mm H₂O over 10 °C to +60 °C without autozero
- Temperature compensated over 10 °C to 60 °C
- Available in gauge surface mount (SMT) or through-hole (DIP) configuration
- Durable thermoplastic (PPS) package

Applications

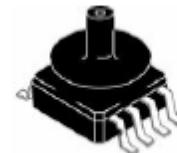
- Washing machine water level
- Ideally suited for microprocessor or microcontroller-based systems
- Appliance liquid level and pressure measurement
- Respiratory equipment

Table 3. Operating characteristics (V_S = 5.0 V_{DC}, T_A = 25 °C unless otherwise noted, P1 > P2)

Characteristic	Symbol	Min	Typ	Max	Units
Pressure range	P _{OP}	0	—	3.92 400	kPa mm H ₂ O
Supply voltage ⁽¹⁾	V _S	4.75	5.0	5.25	V _{DC}
Supply current	I _S	—	—	10	mAdc
Span @ 306 mm H ₂ O (3 kPa) ⁽²⁾	V _{FSS}	—	3.0	—	V
Full-scale span @ 400 mm H ₂ O (3.92 kPa) ⁽²⁾		—	3.92	—	
Offset ⁽³⁾	V _{OFF}	0.75	1.0	1.25	V
Sensitivity	V/P	—	1.0	—	V/kPa
Accuracy ^{(4) (5)}	0 to 100 mm H ₂ O (10 °C to 60 °C)	—	—	±1.5	%V _{FSS} With autozero
	100 to 400 mm H ₂ O (10 °C to 60 °C)	—	—	±2.5	%V _{FSS} With autozero
	0 to 400 mm H ₂ O (10 °C to 60 °C)	—	—	±6.25	%V _{FSS} without autozero



Small outline packages, surface mount



MPXV5004GC6T1/6U, MPVZ5004GC6U
Case 98ASB17757C

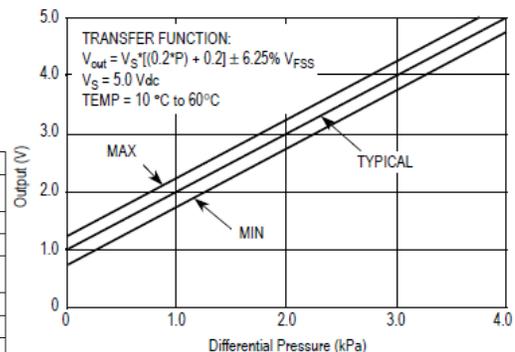


Figure 5. Output vs. pressure differential at ±6.25% V_{FSS} (without autozero, Table 3, note 5)