

MÉTODO PARA LA EVALUACIÓN DE LA VOZ PARA EL LENGUAJE HABLADO ESPAÑOL

DESCRIPCIÓN

5

OBJETO DE LA INVENCION (DESCRIPCION BREVE)

La presente invención hace referencia a un método para la evaluación de la voz del lenguaje hablado español, mediante el uso de un teléfono dispositivo móvil
10 inteligente, que incluye la implementación e interrelación de métodos: a) acústicos, b) aerodinámicos y c) perceptuales, para el diagnóstico funcional de la voz, integrando un sistema experto soportado en lógica fuzzy, el cual ha sido modelado a partir de una base normativa de voces evaluadas en español. El sistema fuzzy incorpora una base de datos normativa de voces del español, que
15 interrelaciona automáticamente mediciones multidimensionales para inferir en la evaluación integral del diagnóstico clínica de la voz, reduciendo el grado de subjetividad por parte personal de salud especializado del área de fonoaudiología (vocología) en la valoración la evaluación. El sistema implementado y posibilita la visualización en un dispositivo móvil de mediciones
20 (cuantitativas y gráficas) de parámetros acústicos, aerodinámicos y perceptuales de la voz, ejecutadas de forma automática desde un sistema cliente - servidor. La invención también presenta de forma automatizada un análisis de la valoración clínica funcional de la voz en el dispositivo móvil inteligente. un resultado de mediciones, gráficos y análisis de la voz desplegadas de forma
25 automática en el dispositivo móvil.

SECTOR TÉCNICO DE LA INVENCION

La presente invención está relacionada con el campo de la física a través de
30 instrumentos o de acústica que analizan la voz o el habla, utilizando técnicas o diferentes parámetros.

ESTADO DE LA TÉCNICA ANTERIOR O ANTECEDENTES DE LA INVENCION

Las enfermedades de la voz son cada vez más frecuentes, estas son diagnosticadas en muchos casos como enfermedades laborales que producen una limitación en la producción de la voz. Por esta razón, es importante realizarse un examen donde se evalúe el estado de todo el aparato que permite el habla y la incorrecta pronunciación de palabras.

Una evaluación foniátrica, o de la voz, consiste en una evaluación del lenguaje que requiere de un procedimiento complejo. Tiene el objetivo de evitar daños o complicaciones sobrevenientes relacionados con el habla, la voz y la audición; si se realiza con un correcto proceso permitirá la reducción y rehabilitación del lenguaje mismo. Por esta razón, es importante realizar periódicamente una evaluación de la voz en adultos.

Los trastornos de la voz, en los seres humanos, manifiestan la dificultad para producir sonidos al tratar de hablar o reflejan un cambio en el tono o calidad de la voz, que puede sonar débil, excesivamente velada, chillona o ronca.

En la actualidad, existen diferentes factores de riesgo físicos, ambientales, laborales y psicológicos, entre otros, que generan afectaciones directas en el aparato vocal de los seres humanos y ocasionan trastornos de la voz, que deben ser identificados y tratados con el mayor cuidado y profesionalismo en tanto que estas patologías ocasionan e incrementan, en los individuos que la padecen, sufrimiento o incapacidad de adaptarse al medio social o laboral.

Entre dichos factores desencadenantes se encuentra el agotamiento profesional, los trastornos músculo - esqueléticos, y las enfermedades infectocontagiosas, que afectan la calidad de vida, en particular para quienes la voz es su instrumento de trabajo (cantantes, locutores, consultores, docentes, solo por citar algunos ejemplos.).

Las principales causas del alto índice de alteraciones se relacionan con que los profesionales que emplean la voz como medio de trabajo, y las personas en general, no acostumbran a realizarse regularmente un examen foniátrico, el cual tiene principalmente componentes acústicos, perceptuales y aerodinámicos, su duración promedio es de 40 minutos y comúnmente los resultados son registrados manualmente en folios escritos o planillas; esto resulta dispendioso y el largo tiempo de duración es agotador para el paciente y el profesional evaluador.

- 10 El test perceptual es la primera herramienta utilizada en la práctica fonoaudiológica, dado que informa sobre la existencia de un trastorno de la voz; sin embargo, este componente parece ser muy limitado, toda vez que una misma característica perceptual puede ser causada por una variedad de procesos y factores totalmente distintos que, por este medio no pueden ser
- 15 diferenciados, adicionalmente presenta la limitación de la subjetividad con la que el oyente o evaluador juzga las condiciones de la voz, llevando a discrepancias intra e interjuicios, y dependencia de la experiencia y pericia del oyente.

- Estos inconvenientes ocasionan que no se detecten de forma precisa las alteraciones en la voz, y por lo tanto no se pueda determinar la efectividad de un tratamiento, además se pueden pasar por alto alteraciones que podrían desencadenar enfermedades mucho más graves como pólipos o cáncer.
- 20

- Teniendo en cuenta lo anterior, es de resaltar que para poder llevar a cabo un proceso de evaluación confiable y dar mayor precisión al diagnóstico, se requiere una combinación adecuada de los componentes fisiológico, acústico, perceptual y aerodinámico, realizando pruebas objetivas que midan tanto los parámetros perceptuales como efecto káiser y tiempo máximo de fonación, como los acústicos jitter, shimmer y la relación armónico-ruido. Sin embargo, hasta el
- 25
- 30 momento no existe ningún sistema que proporcione esta información de manera fácil, completa, segura y diagnostique posibles patologías vocales.

De otra parte para emitir un diagnóstico no se sabe exactamente cómo combinar o ponderar la información de anamnesis, personal, fisiológica, acústica, aerodinámica y perceptual del paciente.

- 5 Es muy difícil y requiere mucha práctica realizar una evaluación perceptual, y se reitera además que los resultados obtenidos son subjetivos, pues dependen de la pericia del profesional que realiza la evaluación.

De otra parte, para la evaluación del componente acústico del examen foniátrico se han propuesto a lo largo de la última década algunos sistemas computacionales para determinar con objetividad parámetros como frecuencia, amplitud, sonoridad y energía de la voz, que brindan información útil para el diagnóstico clínico.

- 15 Dentro de estos sistemas se encuentra un software de escritorio capaz de analizar entonación, intensidad, formantes, cocleograma, entre otros, además de mostrar gráficamente los resultados, no obstante, es complejo y de difícil manejo e interpretación, por lo que presenta desventajas significativas para el personal de salud a la hora de utilizarlo. Otro inconveniente, es que estos sistemas son
- 20 de alto costo y de difícil acceso de parte del sistema de salud para algunas regiones de países en vía de desarrollo, y aunque el uso de estos es indispensable por los profesionales de salud, representan una alta inversión económica que no en todos los casos se tiene.

- 25 Lo anterior, ha ocasionado que se reduzca su uso a sectores que tienen una alta capacidad adquisitiva y se suprima de las organizaciones de escasos recursos. Adicionalmente, la accesibilidad, en términos geográficos, en algunas zonas rurales es tan diversa, que se presentan dificultades a la hora de realizar la atención de pacientes en sectores alejados, ya que muchos de los dispositivos
- 30 para la realización de evaluaciones foniátricas son estáticos y no se pueden movilizar de acuerdo a las necesidades de las poblaciones.

Para solucionar la accesibilidad, han surgido aplicativos para dispositivos móviles, los cuales tienen grandes ventajas por su alta portabilidad. Estos son capaces de calcular algunos parámetros como la frecuencia, Jitter y Shimmer, además determinar el nivel normal o anormal de la voz de acuerdo a un umbral establecido, sin embargo, se pueden encontrar desventajas como: la edad de las personas, ya que solo admite el análisis de personas adultas, excluyendo a los niños; cuenta con más herramientas de análisis como el espectrograma y el máximo de tiempo en la pronunciación, pero se deben pagar para adquirirlos.

10 Observando la necesidad de realizar una evaluación más exhaustiva con instrumental y de modo objetivo de la calidad de la voz, se han creado software y hardware médico dirigido al diagnóstico de patologías vocales, los cuales han optado por diseñar nuevos métodos de análisis acústico como es el caso de los programas software PRAAT, VOXmetria y ANAGRAF, los cuales son sistemas de común uso en Latinoamérica, en ámbitos clínicos y de investigación; además, salió al mercado el dispositivo (hardware y software) VISIPITCH que cuenta con características especiales para la toma de las muestras, el procesamiento de señales y con una interface de usuario para computador de escritorio.

20 Los sistemas mencionados se enfocan en calcular información acústica de señales de voz por medio de parámetros como Jitter, Shimmer, Fo, HNR, formantes y espectrogramas. Todo ello se realiza con el fin de complementar los exámenes médicos y determinar normalidad o anormalidad en la producción de la voz, sin embargo, ninguno permite soportar la evaluación perceptual de la voz.

25 Por otra parte, dichos programas se han diseñado y calibrado para habla inglesa, brasilera y argentina; no para la fonética particular del español colombiano, lo cual influye en los rangos de normalidad. Adicionalmente, dichos programas son para computadores de escritorio, lo cual limita su portabilidad en zonas de difícil acceso.

30

Dentro de los aplicativos móviles dirigidos al estudio de la voz, se encuentran algunos que realizan análisis acústico tales como OperaVox, aplicación para sistema operativo iOS, con la cual se puede obtener mediciones de los

parámetros de Jitter, Shimmer y el tiempo máximo de fonación (TMF), pero con varias restricciones a la hora de usar sus herramientas, dado que requieren ser compradas. Por otra parte, está el aplicativo Voice Test para sistemas operativos Android, el cual permite tomar de una muestra, luego correlaciona los resultados obtenidos y los entrega en un símbolo grafico (semáforo), donde el usuario podrá observar su estado de salud vocal. Otra app relevante es Voice Clinical Systems, la cual permite realizar una evaluación acústica bastante completa y obtener un informe con todos los gráficos y parámetros respectivos, sin embargo, con esta aplicación se toma la muestra de voz y se envía, luego de aproximadamente un día se reciben los resultados, por lo tanto, es una aplicación demorada y costosa. Es de resaltar que estos aplicativos son diseñados para voces de habla inglesa, lo cual limita el uso al resto de público.

Han surgido otros programas como Evaper, el cual permite automatizar la escala GRBAS, calculando sus valores automáticamente, además permite realizar variaciones a cada parámetro y escuchar el resultado en una voz artificial. En este sentido, dicho programa se ha utilizado como simulador para que estudiantes logren percibir variaciones en cada parámetro como la aspereza o soplosidad. Otro simulador es el Auditory--Perceptual Analysis of Voice, que implementa una escala más objetiva llamada CAPE--V. Dicho simulador se implementó como una página web que presenta cinco pacientes, con sus respectivos cuadros clínicos y audios. Posteriormente al análisis del estudiante, se deben emitir los puntajes de la escala CAPE--V, y finalmente el simulador entrega los resultados pregrabados. Si bien el simulador presenta un modelo pedagógico interesante, se limita a los cinco pacientes sin posibilidad de capturar nuevas voces y resultados.

Otras tecnologías comerciales para la valoración de la voz como: el PRAAT, EVAPER, VISI-PITCH, OPERA VOX, no cuentan con la integración de todos los test aerodinámicos, perceptuales, acústicos y fonorespiratorios.

Adicionalmente hemos identificado dentro del estado de la técnica los siguientes documentos de patente, que si bien apuntan a resolver el problema técnico

abordado por nuestra invención, presentan una serie de diferencias, limitaciones y desventajas que son atendidas o superadas por el desarrollo reivindicado en la presente solicitud:

5 La patente KR20170045798 se refiere a un sistema y un método para el diagnóstico de un trastorno de la voz mediante el uso de un teléfono inteligente y, más específicamente, a un sistema y un método para un diagnóstico precoz de un trastorno de la voz mediante que proporciona diversos contenidos o servicios correspondientes a un resultado en pantalla de un estado de voz
10 mediante el uso de una base de datos y / o software almacenado en un servidor, basado en una voz introducida por un usuario. Según la invención, el sistema comprende: un terminal portátil que recibe la voz de un usuario; y un servidor de servicios que analiza la voz del usuario introducida a través del terminal portátil para generar datos de análisis.

15

Esta invención sin embargo no implementa la valoración de voces del lenguaje español. Carece de un módulo de registro y estado de los pacientes, no implementa e integra un sistema experto automático soportado en lógica fuzzy para inferir en la valoración del diagnóstico de la voz; adicionalmente no cuenta
20 con la implementación de un sistema para generar análisis perceptuales como fonetogramas (importante para obtener mediciones cuantitativas de rango vocal.) y el análisis efecto de un glissando importante en valoración de la voz cantada; No indica la interrelación de parámetros acústicos, perceptuales y aerodinámicos para generar análisis integral o automático de la valoración clínica
25 de la voz.

La patente US2008300867 revela Un método y sistema para la evaluación y diagnóstico de la voz en estados normales y enfermos que puede incluir la determinación de al menos una medida cuantitativa de la vibración de las
30 cuerdas vocales utilizando una grabación de imagen laríngea de las cuerdas vocales de un sujeto obtenida de un dispositivo endoscópico o una grabación auditiva de un sujeto durante una tarea fonatoria, y puede incluir el análisis posterior de una forma de onda seleccionada entre tipos de formas de onda que

comprenden a) una grabación acústica, y b) una forma de onda glótica que se extrae de la grabación de la imagen laríngea.

Esta invención carece de una interfaz gráfica simplificada y rápida con menús, y no implementa e integra un sistema experto automatizado soportado en lógica fuzzy para inferir en la valoración de diagnóstico de la voz. Adicionalmente no cuenta con la implementación de un sistema para generar análisis perceptuales como fonetogramas (importante para obtener mediciones cuantitativas de rango vocal.) y el análisis efecto de un glissando importante en valoración de la voz cantada; no indica la interrelación de parámetros acústicos, perceptuales y aerodinámicos para generar análisis integral o automático de la valoración clínica de la voz. Finalmente esta invención difiere de la reivindicada en esta solicitud porque no reivindica que exista la implementación de una base de datos normativa para la valoración de voces del español.

La patente JP2014130227 proporciona una técnica para evaluar objetivamente un método de vocalización en el que una voz se mezcla con una voz de fondo. (discriminar voces). Incluye un dispositivo con una unidad de evaluación que aprende del aprendizaje de datos en base a una onda fundamental y una pluralidad de ondas armónicas en un primer sonido (información sonora) que se obtiene de antemano y corresponde a una voz normal de al menos un sujeto y un sonido que produce persona distinta del sujeto; y un segundo sonido (información sonora) correspondiente a una voz de falsete; y una unidad de extracción para extraer la onda fundamental y la pluralidad de ondas armónicas en la información sonora introducida por un micrófono. El dispositivo introduce datos de evaluación basados en la onda fundamental y la pluralidad de ondas armónicas extraídas por la unidad de extracción, en una expresión de computadora, y calcula un valor de salida de la expresión de computadora como una primera tasa de fusión de la voz normal y la voz de falsete en la información de sonido, lo que permite a una persona comprender más fácilmente el porcentaje de mezcla de una voz normal y una voz de falsete en una voz pronunciada por la persona mediante procesamiento práctico.

Esta invención no integra diferentes pruebas en un único dispositivo, tampoco implementa la valoración de voces del lenguaje español; carece de opciones para ver los resultados de las pruebas y carece de módulos de registro y listados de pacientes; de igual forma no informa si habilita la conexión de datos en un servidor web con conectividad (3G, 4G y/o Wifi). Esta patente tampoco reporta la implementación e integración de un sistema experto automatizado soportado en lógica fuzzy para inferir en la valoración de diagnóstico de la voz, ni indica la interrelación integral de parámetros acústicos, perceptuales y aerodinámicos para generar análisis automático de valoración clínica de la voz. No relaciona la estimación de parámetros aerodinámicos, ni perceptuales para la valoración diagnóstica y no cuenta con la implementación de un sistema para generar fonetogramas, importante para obtener mediciones cuantitativas de rango vocal.

La patente CN104732977 describe un tipo de método y sistema de evaluación de la calidad de la pronunciación del lenguaje hablado en línea, el método descrito incluye: La voz probada recopilada por el cliente móvil es recibida por la red network La voz probada recibida se acopla para ser pre-procesada, La extracción del parámetro característico del habla se lleva a cabo para un habla probada pretratada, obtiene el parámetro característico de la voz probada; De acuerdo con el parámetro característico de la voz probada y el parámetro característico de la pronunciación recibida, la voz probada se evalúa, obtiene el resultado de la evaluación; El resultado de la evaluación se da a el cliente móvil mediante retroalimentación de red, y el resultado de la evaluación es mostrado por el cliente móvil. La presente invención puede realizar una evaluación de la calidad de la pronunciación del lenguaje hablado en línea, conveniente y precisa.

Esta invención sin embargo no permite integrar diferentes test en un único dispositivo; no implementa la valoración de voces del lenguaje español. Carece de un módulo de registro y estado de los pacientes, no implementa e integra un sistema experto automático soportado en lógica fuzzy para inferir en la valoración del diagnóstico de la voz; adicionalmente no cuenta con la implementación de un sistema para generar análisis perceptuales como fonetogramas (importante para obtener mediciones cuantitativas de rango

vocal.) y el análisis efecto de un glissando importante en valoración de la voz cantada; No indica la interrelación de parámetros acústicos, perceptuales y aerodinámicos para generar análisis integral o automático de la valoración clínica de la voz.

5

La patente CN103093759 da a conocer un dispositivo de detección y evaluación de voz basado en un terminal móvil. Comprende un terminal de detección y evaluación de voz, en el que el terminal de detección y evaluación de voz integra una unidad de recolección de señales y almacenamiento de datos, una unidad de procesamiento y evaluación de señales de voz y una unidad de salida de datos. El terminal móvil es adoptado por el terminal de detección y evaluación de voz. La unidad de recogida de señales y almacenamiento de datos comprende un módulo de recogida de señales de voz, un módulo de almacenamiento de datos de señales de voz y un módulo de almacenamiento de resultados de detección y evaluación. La unidad de procesamiento y evaluación de señales de voz comprende un módulo de procesamiento de señales de voz, un módulo de detección de ciclo de señales de voz, un módulo de extracción de parámetros de características y un módulo de análisis y evaluación de características de señales de voz. La unidad de salida de datos comprende un módulo de salida de señal de voz, un módulo de dibujo de forma de onda de señal de voz y un módulo de visita del diario de detección. El dispositivo y el método de detección y evaluación de voz basados en el terminal móvil tienen las funciones de recopilación, proceso, análisis y evaluación de señales de voz.

25

Esta invención no permite la evaluación de voces en lenguaje español; no permite integrar diferentes test en un único dispositivo;. Carece de un módulo de registro y estado de los pacientes, así como de un módulo de visualización de resultados; no implementa e integra un sistema experto automático soportado en lógica fuzzy para inferir en la valoración del diagnóstico de la voz; adicionalmente no cuenta con la implementación de un sistema para generar análisis perceptuales como fonetogramas (importante para obtener mediciones cuantitativas de rango vocal.) y el análisis efecto de un glissando importante en

30

valoración de la voz cantada; No indica la interrelación de parámetros acústicos, perceptuales y aerodinámicos para generar análisis integral o automático de la valoración clínica de la voz. No informa si habilita la conexión de datos en un servidor web con conectividad (3G, 4G y/o Wifi); no tiene una interfaz simplificada y rápida con menús. Finalmente esta invención no informa sobre la implementación de sistemas para realizar valoraciones aerodinámicas de la voz soportadas en la medición del cociente fonatorio a/s y tiempos máximos de fonación que aportan al personal clínico en el cálculo de parámetros de eficiencia respiratoria y fonorespiratoria.

10

Finalmente encontramos la patente US2012116772, que proporciona un sistema y método para la terapia del habla que incluye un dispositivo móvil, un servidor y un cliente web. El dispositivo móvil captura y procesa señales de voz analizadas localmente y en el servidor y desde las cuales se coordina y administra una terapia del habla. El cliente web a través de la interacción con el dispositivo móvil y a través del servidor implementa una terapia del habla que puede ser monitoreada y administrada a través de una moderación clínica específica. El cliente web también proporciona un método alternativo para capturar y transmitir señales de voz al servidor para su análisis y desde el cual se coordina y administra una terapia del habla. La gestión de la terapia del habla puede implementar procedimientos de terapia, pautas y sesiones de comunicación uno a uno entre usuarios y proveedores en un entorno no clínico en tiempo real o en horarios programados. Se describen otras realizaciones.

25

Esta invención se centra en un método y herramienta con dispositivo móvil e interacción cliente web para la evaluación de terapia del habla midiendo parámetros tales como: la velocidad del habla y parámetros acústicos (espectrales, pitch, intensidad entre otros).; pero no permite la evaluación de voces en lenguaje español; no permite integrar diferentes test en un único dispositivo;. Carece de un módulo de registro y estado de los pacientes, así como de un módulo de visualización de resultados; no implementa e integra un sistema experto automático soportado en lógica fuzzy para inferir en la valoración del diagnóstico de la voz; adicionalmente no cuenta con la implementación de un

30

sistema para generar análisis perceptuales como fonetogramas (importante para obtener mediciones cuantitativas de rango vocal.) y el análisis efecto de un glissando importante en valoración de la voz cantada; No indica la interrelación de parámetros acústicos, perceptuales y aerodinámicos para generar análisis integral o automático de la valoración clínica de la voz. No informa si habilita la conexión de datos en un servidor web con conectividad (3G, 4G y/o Wifi); no tiene una interfaz simplificada y rápida con menús. Finalmente esta invención no informa sobre la implementación de sistemas para realizar valoraciones aerodinámicas de la voz soportadas en la medición del cociente fonatorio s/a y tiempos máximos de fonación que aportan al personal clínico en el cálculo de parámetros de eficiencia respiratoria y fonorespiratoria.

Teniendo en cuenta lo anterior, es de resaltar que para poder llevar a cabo un proceso de evaluación confiable y dar mayor precisión al diagnóstico, se requiere una combinación adecuada de los componentes fisiológico, acústico, perceptual y aerodinámico, realizando pruebas objetivas que midan tanto los parámetros perceptuales como efecto káiser y tiempo máximo de fonación, como los acústicos jitter, shimmer y la relación armónico-ruido. Sin embargo, hasta el momento no existe ningún sistema que proporcione esta información de manera fácil, completa, segura y diagnostique posibles patologías vocales.

LISTADO DE FIGURAS

Figura 1. Módulo de registro de usuario.

Figura 2. Menú principal desplegado. (Sistema Front End de la invención)

5 Figura 3. Interfaz de usuario para realizar la grabación de la señal de voz.

Figura 4. Interfaz para graficar señal de la voz grabada.

Figura 5. Interfaz resultado del Test de Glissando

Figura 6. Interfaz resultado del Test Efecto Kaiser

Figura 7. Interfaz resultado del Test Fonetograma

10 Figura 8. Interfaz del módulo de análisis acústico. Test de estimación de los
parámetros: Frecuencia Fundamental, Jitter, shimmer

Figura 9. Interfaz de usuario de evaluación de tiempo máximo fonatorio

Figura 10. Interfaz de usuario del Test que calcula el Índice Fonatorio E/S

Figura 11. Interfaz listado de Usuarios

15 Figura 12. Diagrama de paquetes de la tecnología VocalTest

20

25

30

DESCRIPCIÓN DE LAS FIGURAS

Figura 1. Módulo de registro de usuario

- 5 Interfaz para registro de datos del usuario (Nombre y Apellidos, Número de identificación, Edad, Género. Permite almacenar información digital en una base de datos local.

Figura 2. Menú principal desplegado (Sistema Front End de la invención)

- 10 Pantalla principal del sistema donde se visualizan las opciones de análisis de la voz que puede elegir el usuario. Se pueden seleccionar los diferentes test de análisis:

- Test # 1: reúne un grupo de exámenes que lleva a cabo la evaluación perceptual de la voz mediante el análisis de:

15 ○ Efecto Kaiser

 ○ Glissando

 ○ Fonetograma
- Test # 2: reúne un grupo de exámenes que lleva a cabo la evaluación acústica de la voz mediante el análisis de:

20 ○ Frecuencia Fundamental

 ○ Jitter

 ○ Shimmer
- Test # 3: reúne un grupo de exámenes que lleva a cabo la evaluación aerodinámica de la voz mediante la estimación del:

25 ○ Tiempo Máximo de Fonación

 ○ Índice fonatorio E/S

- En la figura 2 también se observa un botón para acceder a los usuarios registrados con sus respectivos datos y parámetros obtenidos y calculados desde los diferentes test aplicados al análisis de la voz. También se cuenta con
- 30

botón de envío de datos al servidor para posterior análisis en el sistema back end (sistema de lógica fuzzy).

5 En la parte superior se encuentran los datos del usuario, luego se despliega un menú en el que se muestran los diferentes Test que el usuario puede realizar. También se muestra el listado de pacientes. Tiene la posibilidad de configurar diferentes datos en su configuración y el comando para enviar datos a un servidor web para analizar normalidad.

10 Figura 3. Interfaz de usuario para realizar la grabación de la señal de voz.

Este módulo del sistema hace parte del Front end de la invención ya que permite la adquisición de datos (digitalización de la señal de voz) desde la activación del micrófono del Smartphone.

15

Figura 4. Interfaz para graficar señal de la voz grabada.

En este módulo se cuenta con botón que permite graficar la señal de voz. Representa visualmente la morfología de la señal de voz adquirida por el
20 Smartphone.

Figura 5: Interfaz resultado del test de glissando

En la figura 5 se observa en la parte superior la gráfica de frecuencia contra
25 tiempo, para la prueba de glissando se espera que la frecuencia vaya incrementando de manera ascendente y descendente. En la parte inferior se muestra la gráfica de intensidad contra tiempo, para esta prueba se espera que la intensidad se mantenga aproximadamente en línea recta.

30 Figura 6: Interfaz resultado del test efecto Kaiser

En la figura 6 se observa en la parte superior la gráfica de frecuencia contra tiempo, para la prueba de efecto kaiser se espera que la frecuencia se mantenga

constante en línea aproximadamente recta. En la parte inferior se muestra la gráfica de intensidad contra tiempo, para esta prueba se espera que la intensidad se ascendente y luego descendente.

5 Figura 7: Interfaz resultado del test fonetograma

En la parte central de la figura se muestra una escala con las notas musicales de un piano y todas sus octavas, cada tecla reproduce el sonido respectivo. Una vez escuchado el sonido, en la parte superior se presiona el botón grabación para que el paciente imite el sonido iniciando con una baja intensidad y terminando con la máxima intensidad, en seguida se presiona el botón procesar, el sistema calcula y muestra los valores de frecuencia fundamental F0, y los decibeles (intensidad en dB) máximo y mínimo. Si los valores de F0 coinciden con la nota presionada, debajo del teclado se muestra el letrero: nota alcanzada, y se guardan los valores, de lo contrario muestra: Nota no alcanzada.

Figura 8. Interfaz del módulo de análisis acústico. Test de estimación de los parámetros: Frecuencia Fundamental, Jitter, shimmer

En este módulo el sistema permite grabar la señal de voz (sonido vocálico sostenido) grabada para el análisis. Posteriormente grafica la señal y permite el cálculo automático de los parámetros de perturbación Jitter (Jitta, Jitt, Rap y Shimmer (Shim, ShdB, Apq3) representando sus valores mediante diagramas de barra.

25 Figura 9. Interfaz de usuario de evaluación de tiempo máximo fonatorio

Interfaz de usuario que representa el modulo para aplicar el Test aerodinámico relacionado con el cálculo de los Tiempos Máximos de Fonación para sonidos vocálicos sostenidos y sonidos sordos como la transcripción fonética de la “s”. La interfaz permite grabar sonidos vocálicos sostenidos y sonidos sordos, también posibilita la reproducción acústica de los audios y por último permite estimar y visualizar el resultado obtenido de los tiempos máximos de fonación.

En la interfaz se indican los valores de referencia para géneros masculino y femenino.

Figura 10. Interfaz de usuario del Test que calcula el Índice Fonatorio E/S

5

Este módulo hace parte del test de evaluación aerodinámico. En este módulo se cuenta con un botón que graba un sonido vocálico sostenido y luego permite grabar un sonido sordo sostenido. Este módulo permite reproducir los archivos de audio de los sonidos grabados y posibilita la visualización graficar que representa la morfología de las señales de voz emitidas por el usuario. Este módulo permite estimar los tiempos máximos de fonación para cada sonido grabado y se estima el índice fonatorio desde la división de los tiempos máximos de fonación entre el sonido vocálico y el sonido sordo.

15 Figura 11: Interfaz listado de usuarios

En la figura 11 se muestra una lista con los nombres y apellidos de todos los pacientes que se han registrado en la aplicación. Cuando se presiona cada uno de los ítems de la lista se muestran los datos completos del paciente como cedula y género.

20

Figura 12: Diagrama de paquetes de la tecnología VocalTest

La figura 12 muestra el diagrama de paquetes de la tecnología VocalTest, según el UML (Unified Modeling Language) en el cual se observan cuatro partes fundamentales, las vistas que representan cada una de las interfaces de usuario descritas, los recursos donde se encuentran las imágenes y los audios generados durante los test, la lógica que contiene los algoritmos utilizados para procesar los audios y calcular los diferentes parámetros de la voz, y finalmente la base de datos, en donde se almacenan la información de los usuarios con sus respectivos parámetros y resultados.

25

30

DESCRIPCIÓN DETALLADA DE LA INVENCION

5

La presente invención hace referencia a un método para la evaluación de la voz para lenguaje hablado español, mediante el uso de un teléfono inteligente, que incluye la implementación de métodos: a) acústicos, b) aerodinámicos y c) perceptuales, para la valoración de la voz, integrando un sistema experto soportado en lógica fuzzy, el cual incorpora una base de datos normativa de voces del español, que interrelaciona automáticamente mediciones multidimensionales para inferir en la evaluación integral del diagnóstico clínica de la voz reduciendo el grado de subjetividad en la valoración por parte personal del área de vocología (fonoaudiología) y posibilitando un resultado de mediciones, gráficos y análisis de la voz desplegadas de forma automática en el dispositivo móvil.

La presente invención hace referencia a un método para la evaluación de la voz del lenguaje hablado español, mediante el uso de un dispositivo móvil inteligente, que incluye la implementación e interrelación de métodos: a) acústicos, b) aerodinámicos y c) perceptuales, para el diagnóstico funcional de la voz, integrando un sistema experto soportado en lógica fuzzy, el cual ha sido modelado a partir de una base normativa de voces evaluadas en español. El sistema fuzzy interrelaciona automáticamente mediciones multidimensionales para inferir en la evaluación integral de la voz, reduciendo el grado de subjetividad por parte personal de salud especializado en el área de fonoaudiología (vocología) El sistema implementado posibilita la visualización en un dispositivo móvil de mediciones (cuantitativas y gráficas) de parámetros acústicos, aerodinámicos y perceptuales de la voz, ejecutadas de forma automática desde un sistema cliente - servidor. La invención también presenta de forma automatizada un análisis de la valoración clínica funcional de la voz en el dispositivo móvil inteligente.

El método ha sido creado para ser implementado por medio de un teléfono o dispositivo móvil inteligente con sistema operativo Android, que permite aplicar a diversos grupos de pacientes, al menos tres tipos de test de fonoaudiología tales como: test acústico, test perceptual y test aerodinámico, donde dicho sistema está diseñado para presentar los resultados y almacenarlos en una base de datos.

La aplicación es nativa, por lo que las conexiones son rápidas y la interfaz está simplificada para el uso en un dispositivo móvil inteligente el teléfono. en configuración cliente-servidor.

Una ventaja técnica que propone la invención es el integrar en un dispositivo móvil diversos exámenes foniátricos automatizados, incorporando adicionalmente un sistema de análisis inteligente soportado en lógica fuzzy, los cuales permiten sistemáticamente la evaluación clínica instrumental de la voz, incorporando una valoración acústica, una valoración perceptual y el cálculo de parámetros aerodinámicos y fono respiratorios, posibilitando adicionalmente un examen para evaluar el rango vocal en voz cantada.

La invención cuenta con diversos exámenes foniátricos automatizados, customizados para realizar una valoración clínica funcional de la voz de forma autónoma. Cuenta además con la integración de un sistema experto soportado en lógica fuzzy para la valoración instrumental de la voz. La mayoría de las herramientas actuales no han sido adaptadas para realizar exámenes automáticos multidimensionales automatizados aplicados en la valoración clínica de la voz.

De igual forma brinda un conjunto de instrumentos para la valoración funcional de la voz: a) test acústico, b) test perceptual y c) test aerodinámico. Los resultados de los instrumentos de evaluación se presentan cuantitativamente, gráficamente y se almacenan en una base de datos. En relación al test acústico, inicialmente, se graba la voz de una persona en evaluación bajo niveles adecuados de sonometría que el sistema calcula, luego se reproduce y se presenta una gráfica de la señal de voz. Posteriormente la señal de voz es

enviada vía IP a un servidor donde algoritmos procedimientos computacionales desarrollados estiman automáticamente los parámetros acústicos que son luego regresados al dispositivo móvil y presentados en gráficos, indicando resultados de la evaluación acústica.

5

Por otra parte, en relación al test aerodinámico se calculan automáticamente tiempos máximos de fonación, emitiendo a partir de la emisión de sonidos vocálicos (“a”) y sordos (“s”) sostenidos.

10 La grabación de señales se obtiene con un botón de inicio y fin, acompañado de un cronometro. El test aerodinámico estima parámetros automáticos como el cociente fonorespiratorio para evaluar la función laríngea y su eficiencia vocal a partir de la relación entre el cierre glótico y la fuerza de espiración pulmonar durante un soplo fonatorio.

15 Por último, el test perceptual, se realiza a partir de la generación de un direccional al paciente bajo un protocolo automatizado, instrucciones para la generación de un fonetograma para voz entrenada (voz cantada) y no entrenada. Adicionalmente posibilita la aplicación de pruebas de efecto glissando y efecto káiser para calcular automáticamente el rango de frecuencias y de intensidades
20 máximas y mínimas que puede alcanzar una persona que utiliza esta aplicación para la evaluación de su voz.

La presente invención es una tecnología portable y es al ser implementada en un dispositivo de tamaño reducido, permite valorar la voz sin necesidad de tener
25 contar con dispositivos grandes, robustos, dispendiosos y de escritorio, como algunos sistemas sofisticados que convencionalmente se encuentran en laboratorios de acústica.

Adicionalmente la invención ha sido implementada para valorar voces del
30 lenguaje hablado en español, incorporando el desarrollo de una propia normativa de evaluación de voces generada por fonoaudiólogos especialistas. Lo anterior es una ventaja, ya que la mayor parte de sistemas desarrollados provienen desde normativas de valoración de la voz para voces distintas al español.

La invención brinda una solución técnica ágil y eficiente, al permitir de forma automática la realización de diversos exámenes foniátricos para la valoración integral funcional de la voz y sistematizar los resultados con el propósito de sugerir una valoración de la voz soportada en el módulo de inteligencia computacional. Lo anterior posibilita que sea una herramienta portable al alcance de profesionales en la valoración de voz y también como una herramienta flexible en la valoración autónoma de personas que no cuentan con conocimientos foniátricos.

10

La tecnología desarrollada procesa las pruebas perceptivas perceptuales y aerodinámicas, y entrega los resultados de las mismas inmediatamente después de realizar las pruebas.

15 Permite calcular tres tipos de jitter y tres tipos de shimmer, entregando su valor y rango de normalidad; así como los parámetros perceptuales: fonetograma para cantantes con voz entrenada, y glissando y efecto kaisser para voces no entrenadas; y calcular Parámetros aerodinámicos: coeficiente E/Z cociente fonorespiratorio a/z y el tiempo máximo de fonación. Estos son calculados automáticamente mediante algoritmos a partir de las grabaciones de voz.

20

Utiliza un sistema experto automatizado soportado en lógica fuzzy alojado en un servidor, considerado como variables de entrada: parámetros acústicos, aerodinámicos, perceptuales para inferir en la valoración de un único diagnóstico clínico de la voz, teniendo en cuenta los resultados de los parámetros perceptivos, acústicos y aerodinámicos.

25

El test acústico se implementa grabando la voz del paciente, posteriormente se reproduce esta grabación y dicha señal de voz se representa en una gráfica; al finalizar, para cada parámetro se muestra una gráfica con su valor máximo y mínimo de normalidad.

30

En el test aerodinámico se calculan tiempos máximos en sostener una vocal 'A' y una 'S', se utiliza un grabador de sonidos, con un botón de inicio y fin, acompañado de un cronometro.

- 5 En el test perceptual, se realiza un fonetograma para voz entrenada y no entrenada, para calcular el rango de frecuencias y de intensidades máximas y mínimas que puede alcanzar un paciente.

10 Para efectos de la comprensión técnica del tipo de tests utilizados para la implementación de la tecnología, se entiende como Fonetograma una herramienta creada para el análisis y la rehabilitación de la voz a nivel clínico que posibilita conocer gráficamente y evaluar objetivamente el estado de los parámetros que conforman la gama vocal de una persona. Consiste en una prueba que valora el área vocal y mide los parámetros de intensidad y frecuencia
15 además de la interrelación entre éstos.

En 1952, Calvet definió el fonetograma como método para ilustrar rangos de tonos e intensidad de la voz. En 1981 la Unión Europea de Foniatra estandarizó el concepto y se definieron las bases para la generación de fonetogramas. El
20 fonetograma sirve para valorar voces normales y patológicas y para conocer las potencialidades de la extensión vocal en los cantantes. (Calvet, J. 1953. Etudes phonometriques. In La Voix, Cours international de phonologie et phoniatrie, pp. 67- 78. Paris: Libraire Maloine.)

- 25 De igual forma, los parámetros utilizados en el análisis acústico de la voz sirven para evaluar la voz subjetivamente (evaluar patologías laríngeas); acompañado de una base de datos normativa que posibiliten la caracterización de la calidad de la voz y utilizando sistemas inteligentes que combinan varios parámetros para distinguir entre normalidad y patología. Estas herramientas permiten disminuir el
30 grado de subjetividad en el análisis de la voz.

Las Mediciones Aerodinámicas por su parte relacionan la estimación de indicadores de apoyo a la valoración clínica de la voz. Contribuyen a la

evaluación de la aerodinámica de la función laríngea que se puede estudiar midiendo una serie de parámetros que se relacionan con la eficiencia del cierre glótico por una parte y con la fuerza de la eficiencia del cierre glótico por una parte, y con la fuerza de la espiración pulmonar que genera el soplo fonatorio.

5

El Método para la evaluación de la voz para lenguaje hablado español, mediante el uso de un teléfono inteligente con sistema operativo Android, que incluye la implementación de test automatizados de fonoaudiología: a) acústico, b) aerodinámico y perceptual, estima los siguientes parámetros:

10

- Parámetro Tiempo Máximo de Fonación (TMF): Relaciona el flujo aéreo espiratorio con las cuerdas vocales.

15

- Índice (s/z) – cociente fonorespiratorio: Es un indicador que resulta de dividir el tiempo máximo de producción de una consonante /s/ entre el tiempo máximo de fonación para la /a/. El numerador estará relacionado con el TMF del volumen pulmonar y el denominador con la fonación de cuerdas vocales. (Se obtiene dividiendo el tiempo máximo de fonación de la consonante /s/ entre el tiempo máximo de fonación de la vocal /a/ (tras inspiración profunda en un tono e intensidad cómodos). Su valor teórico normal es 1, si es mayor de 1.4 es patológico e indicativo de un cierre laríngeo defectuoso).

20

REQUERIMIENTOS MÍNIMOS PARA LA IMPLEMENTACIÓN:

25

Para que el método puedas ser implementado correctamente, es necesario disponer de un teléfono inteligente con sistema operativo Android (5.0 o superior), Procesador de 400 MHz, Memoria RAM de 512 Mb.50 Mb de disco duro.

30

Adicionalmente para enviar los datos al servidor web es necesario disponer de conectividad (3G, 4G y/o Wifi).

IMPLEMENTACIÓN DEL MÉTODO:

Las funciones de la aplicación son: registrar pacientes, ver los pacientes registrados, realizar diferentes test de fonoaudiología, ver los resultados de los test, almacenarlos en el Smartphone y enviar los datos a un servidor web.

La interfaz desarrollada permite que el usuario se registre y ejecute los diferentes tests a través de un conjunto de módulos diseñados para ello, una vez se accede a la interfaz de inicio del método se debe activar el registro del usuario en los módulos de inicio y de registro:

Módulo de Inicio: Para la realización de las evaluaciones el usuario debe ingresar su número de identificación y seleccionar el botón de continuar; si el usuario se encuentra registrado será dirigido a la pantalla principal. En caso de no estar registrado, será redirigido a la interfaz de registro.

Módulo de Registro (FIG. 1): En la interfaz de registro, el usuario ingresa los datos que solicitados, tales como: nombre, apellidos, identificación, edad y género. Para guardar la información debe seleccionar el botón de guardar. Si el proceso se realizó correctamente será dirigido al menú principal, y los datos son almacenados en la base de datos del dispositivo.

Menú principal (FIG 2.): Una vez el usuario se encuentre en la interfaz del menú principal, podrá acceder a las opciones seleccionando el botón de menú, comúnmente conocido como “botón de hamburguesa”. En el menú principal el usuario podrá visualizar la siguiente información:

- Datos del usuario
- Diferentes Test que el usuario puede realizar
- Listado de los pacientes
- Configuración de la aplicación
- Enviar los datos a un servidor web para analizar normalidad de la voz.

La solución también tiene un listado de pacientes: Para ver los pacientes que están registrados en la aplicación, el usuario seleccionara la opción de pacientes en el menú. Una vez seleccionada la opción, se mostrará un listado de todos los usuarios. Posteriormente, si se selecciona algún paciente del listado, se muestran los datos del usuario registrado y un listado de las evaluaciones que se le han realizado a dicho paciente con sus respectivos valores de las pruebas con su fecha de realización.

APLICACIÓN DE LOS DIFERENTES TEST : (FIG. 2, 3, 4):

Desplegando el botón de menú, una vez el usuario se encuentre en la vista donde están todos los exámenes que ofrece la tecnología, es autónomo de escoger cualquiera de los test. En todos los casos se debe realizar un proceso de grabación de audio.

El dispositivo móvil debe contar con herramientas para la toma de la señal de voz, a través de una grabadora de audio, este punto es de importancia dado que, una buena toma de la muestra, significa la obtención de buenos resultados cuantitativos.

Para la toma de la muestra de voz, es importante que el usuario se ubique en un ambiente tranquilo y lo más silencioso posible, dado que la interferencia por ruido, puede causar resultados desagradables o falsos.

Para garantizar el nivel de ruido, la herramienta incorpora un sonómetro que calcula el nivel de ruido en decibeles.

Cuando el usuario inicia cualquier grabación, primero es redirigido al sonómetro, el cual pide silencio al usuario y graba automáticamente un segundo, si el nivel de ruido esta por debajo del umbral, le permite al usuario proseguir y grabar el audio para la prueba de voz

- Test de Glissando: Glissando hace referencia a un término que indica el deslizamiento continuo de una nota a otra. No es una palabra genuinamente italiana sino la italianización del verbo francés glisser, que significa “resbalar”, “deslizar”. Aplicada al piano, el efecto se obtiene deslizando rápidamente el pulgar o el costado del dedo índice a lo largo de una serie de teclas adyacentes, ascendentes o descendentes; hay también glissandos dobles, por lo general en octavas, y pueden darse glissandos triples también. Ravel y Debussy recurrieron al glissando en su música para piano. Esta técnica es muy común en la música para arpa, como en Introducción y allegro de Ravel. En los instrumentos de arco y en la voz, el glissando recorre un infinito número de microtonos. (Randel, 2003).

Para la realización de este Test, el usuario recibe las instrucciones para realizar el test, luego podrá seleccionar el botón de INICIAR GRABACION para empezar a grabar la nota de voz. Una vez finalizado el tiempo de grabación, el usuario podrá reproducir la grabación con el botón de REPRODUCIR GRABACION, ver un gráfico de la onda de voz con el botón de GRAFICAR y ver los resultados con el botón SIGUIENTE. (FIG 5).

Los resultados se presentan con 2 gráficos, el primero representa la variación de la frecuencia, a medida que transcurre el tiempo de la grabación. Con este grafico se calcula el rango de frecuencia que es capaz de realizar el paciente sin alterar su voz en intensidad.

Con el segundo grafico se calcula la variación de la intensidad, a medida que transcurre el tiempo de la grabación. Con este grafico se calcula que el paciente logre mantener la voz dentro de un rango de intensidad aceptable.

- Test Efecto Kaiser: El efecto káiser: es la ausencia de emisión acústica detectable hasta el momento en el que se excede el nivel de carga o intensidad máxima previamente aplicada. (Unidos, 1964) (Breckwoldt, 1957). Al igual que el Test anterior. El usuario podrá seleccionar el botón de INICIAR GRABACION para empezar a grabar la nota de voz. Una vez finalizado el tiempo de grabación, el usuario podrá reproducir la grabación con el botón de REPRODUCIR

GRABACION, ver un gráfico con el botón de GRAFICAR y ver los resultados con el botón SIGUIENTE. (FIG. 6)

5 Los resultados se presentan con 2 gráficos, el primero representa la variación de la frecuencia, a medida que transcurre el tiempo de la grabación. Con el segundo grafico se calcula la variación de la intensidad, a medida que transcurre el tiempo de la grabación. Con esta prueba se calcula el rango de intensidad que el paciente logra mantener sin alterar su voz, para una frecuencia conversacional. Mediante el primer grafico se verifica que la frecuencia se mantenga dentro de
10 un rango aceptable

Test de Fonetograma: Este tipo de test es ideal para pacientes con entrenamiento vocal como cantantes, de lo contrario pueden realizar el glissando y efecto káiser.

15

- Para iniciar el test, El usuario deberá seleccionar el botón de GRABACION, posteriormente el usuario presionará una tecla del piano y empezará a grabar la nota de voz por un determinado tiempo. Una vez finalizado el tiempo de grabación, el usuario podrá reproducir la grabación con el botón de
20 REPRODUCIR. Para ver los resultados, con el botón PROCESAR. (FIG. 7)

-

En detalle: El usuario deberá presionar, en primer lugar, una tecla de piano y reconocer el sonido, luego inicia la grabación y tratará de imitar el sonido escuchado en su frecuencia. Si la frecuencia de la nota de piano y la del paciente
25 son muy similares (se admite un rango de error) el sistema dice que alcanza la nota, de lo contrario, le dice que no alcanza y debe volver a intentar o terminar el examen. Adicionalmente el usuario cuando graba la nota debe iniciar en una intensidad baja e ir subiendo la intensidad, similar al efecto kaiser.

30 El proceso se repite para cada nota de piano, el sistemas va guardando los datos de frecuencias y máximos y mínimos de intensidad del paciente. Al final del examen, se le muestra un gráfico con el are vocal.

- Test de Frecuencia Fundamental, Jitter, shimmer: (FIG. 8). Para iniciar el test, El usuario deberá seleccionar el botón de EMPEZAR. Posteriormente el usuario empezara a grabar la nota de voz por un lapso de tiempo; El tiempo predefinido son 8 segundos, pero en los ajustes se puede configurar uno
5 diferente. Para ver una gráfica de la grabación, el usuario deberá seleccionar el botón de GRAFICAR. Para reproducir la nota de voz, el usuario deberá seleccionar el botón de REPRODUCIR.

Luego de reproducir, se da la opción de ver los resultados, el sistema muestra
10 una interfaz con los valores de frecuencia fundamental, jitter (absoluto y porcentual), shimmer (absoluto y porcentual), además para cada valor se presenta la escala de normalidad del parámetro, de modo que el usuario puede apreciar si esta dentro del rango normal

15 - Test de Tiempo Máximo de Fonación: (FIG. 9) Al igual que el Test anterior. Para iniciar el test, El usuario deberá seleccionar el botón de EMPEZAR. Posteriormente el usuario empezara a grabar la nota de voz por un lapso de tiempo. Aquí es de resaltar que el usuario debe emitir la vocal 'a' en un tono conversacional y frecuencia cómoda, y mantener lo máximo que pueda, cuando
20 se le acabe el aire y ya no pueda mas, debe para la grabación.

El sistema automáticamente elimina los silencios de inicio y de fin de la grabación, y calcula y muestra el tiempo que el paciente mantuvo su voz sin alteraciones grandes de intensidad y frecuencia.

25

Al final se muestra el valor calculado y el rango de normalidad

El usuario podrá pausar la grabación con el botón de PAUSA. Para reproducir la nota de voz, el usuario deberá seleccionar el botón de REPRODUCIR. Si el
30 usuario desea guardar el resultado, podrá hacerlo con el botón de GUARDAR RESULTADO. Si el usuario desea regresar al menú principal, deberá seleccionar el botón de VOLVER.

- Test de cociente fonorrespiratorio a/s Índice Fonatorio E/S: (FIG. 10).

Este test se divide en dos partes. La primera parte es pronunciando la letra 'E', y la segunda parte es pronunciando la letra S. El usuario deberá seleccionar el botón de EMPEZAR y posteriormente el usuario empieza a grabar la nota de voz por un lapso de tiempo. El usuario podrá pausar la grabación con el botón de PAUSA. Para reproducir la nota de voz, el usuario deberá seleccionar el botón de REPRODUCIR. Para ver los resultados del Test, el usuario deberá seleccionar el botón de CALCULAR.

- 10 Aquí es importante aclarar que el procedimiento es similar a la prueba de tiempo máximo de fonación, solo que aquí se repite para la 's'. El resultado es la división de los dos tiempos, el de la 'e' dividido el de la 's'.

Al final se muestra el valor calculado y el rango de normalidad

15

Es importante señalar que la información es almacenada directamente en el dispositivo, no en un servidor; sin embargo cuando se utiliza la opción del menú 'enviar datos', los parámetros calculados si son enviados a un servidor externo, el cual utiliza un clasificador de lógica fuzzy para determinar si el paciente tiene una alteración o es normal, este resultado es enviado desde el servidor hacia el dispositivo móvil y presentado en una interfaz.

20

Diagrama de paquetes modulo acústico y aerodinámico: (FIG. 12).

- 25 Contiene diferentes tipos de paquetes, donde cada uno cumple una función específica y se describen a continuación:

a) Vistas

- 30 Este paquete abarca el diseño de la interfaz gráfica y secuencial del aplicativo móvil, comprende todos los elementos tipo Layout del programa Android Studio, y con los cuales el usuario puede interactuar por el sistema.

Se trató en lo posible de cumplir el segundo objetivo específico planteado antes de iniciar todo el trabajo de programación.

b) Lógica

5

Con la ayuda de códigos externos (Harlan, 2014) para la extracción de datos codificados en los audios se logra tomar muestras de diferentes tamaños, para hacer análisis de voz, con lo cual se logra visualizar e implementar la señal de voz.

10

Todo se realiza implementando el lenguaje Java, donde se construye cada uno de los algoritmos utilizados para hacer el cálculo de los parámetros acústicos como: Frecuencia fundamental, Jitter y Shimmer; además para la obtención de las mediciones aerodinámicas como: Tiempo máximo de fonación y Cociente

15

Fonorrespiratorio a/s e/s.

c) Recursos

En cuanto al paquete de recursos, está compuesto por elementos que son importantes para visualizar, funcionar y almacenar información como son:

- Imágenes: son archivos con extensión .png, las cuales son necesarias en la interfaz de usuario y aportar mayor satisfacción visual al público.
- Descriptores: son archivos de uso programático, necesarios para que las partes del aplicativo funcionen con normalidad. Algunas funcionan como control central del sistema y en la mayoría de casos son con extensión .xml.
- Almacenamiento de audios: son archivos que se agregan a almacenamiento del dispositivo móvil después de haber sido grabados en la sesión del examen vocal.
- Otros recursos: elementos auxiliares necesarios para el funcionamiento del sistema, abarca: arreglos, colores, strings, entre otros.

d) Base de datos

Este paquete administra toda la información de los usuarios, en cuando a los
5 registros y resultados almacenados en cada sesión.

Está integrado por una clase donde se estructuran las tablas de relaciones entre
la información del usuario y los resultados, además de un controlador que se
encarga de mantener actualizado el sistema de bases de datos y de que no haya
10 perdida de la información.

Diagrama de paquetes modulo perceptual:

Se generaron diversos tipos de paquetes, cada uno cumple una función
15 específica descritas a continuación:

a.) Vistas

Dentro de este paquete se encuentran las interfaces gráficas y secuenciales de
la aplicación, comprende los componentes de tipo Layout del software Android
20 Studio, el usuario interactúa con estos a través del sistema.

b.) Lógica

Apoyado en códigos externos para la obtención de datos en las grabaciones de
audios se consigue tomar muestras de diversos tamaños, con lo que se logra
25 visualizar y analizar la señal de voz, a través de los algoritmos ya establecidos,
logrando realizar cálculos de parámetros acústicos como: intensidad(dB),
frecuencia (Hz) y frecuencia fundamental (Hz).

c.) Recursos

Este paquete se encuentra constituido por diversos elementos de gran
30 importancia para la aplicación ya que estos ayudan a un óptimo funcionamiento
como descriptores (archivos necesarios para que el aplicativo funcione
correctamente como es el control central del dispositivo son extensiones .Xml),

strings, arreglos, en visualización tenemos lo que son imágenes (necesarias para la interfaz son archivos con extensión .Pn) y colores, en almacenamiento de datos encontramos los audios ya que estos se almacenan luego de ser grabados y otros recursos estos son elemento auxiliares imprescindibles para el funcionamiento del sistema operativo.

Diagrama de clases Modulo Acústico y aerodinámico.

En este diagrama se visualizan las relaciones entre las clases involucradas. Cada clase encapsula toda la información de un objeto, y a través de ella se puede modelar el entorno de estudio como el caso .

La operación da inicio en la primera clase llamada <Bienvenida>, la cual lleva una asociación o secuencia hacia el segundo paso llamado <Opciones de Usuario>, en donde el usuario en libre es escoger una opción ya sea <Registrar Usuario>, <Repetir examen> o <Listar Registros>.

Si las condiciones para la operación se completan en este paso, se habilita la opción de <Seleccionar Tipo de examen> para que el usuario determine que examen desea realizarse y proceda a la toma de muestra de voz en la clase <Grabadora de Audio>, la cual hereda procesos de las clases <Algoritmos> (clase donde se procesa el audio para obtener los resultados paramétricos) y <Ruta Audio> (clase donde se realiza la conversión del archivo a formato .wav y sucesivamente se almacena en el dispositivo).

Si el proceso en las clases anteriormente mencionadas se completa, el usuario puede visualizar los resultados y finalmente quedan almacenados en la base de datos del sistema. Fin del proceso.

Diagrama de clases Modulo perceptual.

Este diagrama da a conocer visualmente como se relacionan las clases que comprende el sistema, encapsulando toda la información de un objeto en diversas clases, con lo que logra modelar un entorno.

- 5 Al ejecutar el método comienza con la primera clase <Presentación>, este nos traslada de forma secuencial al segundo paso al cual se le ha nombrado <Opción de Examen> donde el usuario establece o determina el examen a realizar seleccionando una opción entre <Glissando y E.kaiser> o <Fonetograma>.
- 10 Si selecciona la opción <Glissando y E.kaiser> para realizar el examen se procede a la toma de muestra de voz en la clase <Grabadora de Audio Glissando> esta hereda procesos de las clases <Algoritmos> (clase donde se procesa la muestra de audio para obtener los resultados paramétricos) y <Ruta Audio> (clase donde se realiza la conversión del archivo a formato .wav y su
- 15 almacenamiento en el dispositivo), si este proceso se culmina con éxito, el usuario puede visualizar los resultado a través de una gráfica.

- Si selecciona la opción <Fonetograma> para realizar el examen se procede a la toma de muestra de voz teniendo en cuenta que se debe seleccionar en el piano
- 20 presente la nota con la que se desea realizar la prueba en la clase <Grabadora de Audio Fonetograma> esta hereda procesos de las clases <Algoritmos> (clase donde se procesa la muestra de audio para obtener los resultados paramétricos) y <Ruta Audio> (clase donde se realiza la conversión del archivo a formato .wav y su almacenamiento en el dispositivo), si este proceso se culmina con éxito, el
- 25 usuario puede visualizar los resultado a través de una gráfica.

Alcance de la lógica fuzzy en la invención

- La lógica fuzzy en el método descrito para la evaluación automatizada de la voz desempeña un papel crítico en la reducción de la subjetividad y la complejidad
- 30 inherentes al diagnóstico foniátrico. El sistema experto soportado en lógica fuzzy permite interrelacionar automáticamente parámetros acústicos, aerodinámicos y

perceptuales para proporcionar un diagnóstico integral y objetivo. Este enfoque técnico ofrece una solución eficiente y reproducible a las limitaciones de los métodos tradicionales.

Cómo opera concretamente la lógica fuzzy en la invención

5 a. Ingreso de datos (señal de voz):

- La voz grabada por el usuario se procesa en el dispositivo móvil. Los algoritmos convierten las señales acústicas (ondas analógicas) en datos digitales, los cuales se parametrizan en diferentes dimensiones:

- 10 ▪ **Acústica:** Parámetros como *Jitter*, *Shimmer* y frecuencia fundamental.
- **Aerodinámica:** Tiempo máximo de fonación (TMF) y cociente fonorespiratorio (s/z).
- 15 ▪ **Perceptual:** Evaluación del rango vocal mediante fonetogramas.

b. Transformación y análisis inicial de datos:

- Estos parámetros se calculan a partir de algoritmos que interpretan las características acústicas y físicas de la voz del usuario. Los datos se visualizan en gráficos y tablas en el dispositivo móvil, mostrando rangos de normalidad o alteración.

20

b. Integración en el sistema fuzzy:

- Los parámetros calculados sirven como **entradas para el sistema fuzzy** alojado en un servidor web. Este clasificador fuzzy:
 - Evalúa la relación entre los parámetros (por ejemplo, correlaciona *Jitter* y *Shimmer* para detectar perturbaciones en las cuerdas vocales).
 - Asigna grados de pertenencia a conjuntos difusos (por ejemplo, "normal", "riesgo" o "alterado") en función de reglas

25

predefinidas basadas en normativas clínicas específicas para el español.

d. Evaluación y clasificación:

- 5 ○ El sistema fuzzy clasifica automáticamente la voz del usuario en una de las categorías posibles:
 - **Normal:** Todos los parámetros se encuentran dentro de los rangos esperados.
 - **En riesgo:** Algunos parámetros indican una ligera desviación, lo que puede requerir monitoreo adicional.
 - 10 ▪ **Alterada:** Los parámetros muestran desviaciones significativas, sugiriendo posibles patologías.

e. Resultados finales:

- 15 ○ La clasificación fuzzy se envía de vuelta al dispositivo móvil, donde se presenta en una interfaz gráfica fácil de interpretar. Los usuarios pueden acceder a gráficos, tablas y recomendaciones basadas en los resultados.

Ventajas del sistema fuzzy en la invención

- **Reducción de la subjetividad:**
 - 20 ○ Al basarse en reglas claras y rangos predefinidos, la lógica fuzzy minimiza el impacto de la experiencia o juicio personal del profesional de salud.
- **Evaluación integral y multidimensional:**
 - 25 ○ Combina datos de distintas fuentes (acústicos, aerodinámicos, perceptuales) para realizar un análisis integral que no sería posible con métodos tradicionales.
- **Automatización y eficiencia:**
 - El sistema permite un diagnóstico rápido y automático, eliminando la necesidad de evaluaciones largas y complejas realizadas manualmente.

- **Adaptabilidad al contexto lingüístico:**

- Utiliza una base normativa específica para el español, asegurando que los diagnósticos sean precisos y culturalmente relevantes.

La lógica fuzzy en esta invención permite transformar un proceso altamente
5 subjetivo en un método automatizado, reproducible y adaptado a las
necesidades específicas. Su capacidad para manejar incertidumbre y
correlacionar parámetros multidimensionales asegura un diagnóstico preciso y
útil tanto para profesionales de la salud como para usuarios sin conocimientos
especializados. Este enfoque innovador optimiza el uso de recursos
10 tecnológicos, haciéndolo accesible en dispositivos móviles en cualquier contexto
geográfico o económico.

La patente describe un método para la evaluación de la voz en el lenguaje
hablado español mediante el uso de un sistema automatizado integrado en un
dispositivo móvil inteligente. Este método utiliza parámetros acústicos,
15 aerodinámicos y perceptuales para generar un diagnóstico funcional de la voz.
Un componente esencial de este sistema es la implementación de un sistema
experto soportado en lógica fuzzy, que desempeña un papel determinante para
alcanzar los objetivos de la invención por las siguientes razones:

- ❖ **Manejo de subjetividad en el diagnóstico:**

- El diagnóstico tradicional de la voz a menudo está influenciado por la
20 **subjetividad** de los evaluadores (dependiendo de la experiencia del
profesional y su juicio perceptual).
- La lógica fuzzy implementada permite modelar esta subjetividad mediante
la asignación de grados de pertenencia a diferentes conjuntos difusos,
25 como "buena calidad vocal", "calidad moderada", o "calidad deficiente".
Esto **reduce la dependencia de la experiencia humana** y permite
decisiones más uniformes y objetivas.

- ❖ **Interrelación multidimensional de parámetros:**

- Los parámetros acústicos (como Jitter, Shimmer), aerodinámicos (como
30 el cociente fonorespiratorio), y perceptuales (como los fonetogramas) no
tienen límites estrictos y pueden solaparse en ciertas categorías.

- La lógica fuzzy implementada facilita la **interrelación automática** de estos parámetros, permitiendo que el sistema realice inferencias basadas en combinaciones complejas de datos, sin necesidad de valores exactos.

❖ **Adaptación a incertidumbre y variabilidad:**

- Las mediciones de voz tienen **variabilidad inherente** debido a factores como edad, género, estado de salud, o contexto lingüístico.
- La lógica fuzzy puede manejar esta incertidumbre al trabajar con rangos y niveles de probabilidad, lo que asegura que el sistema pueda adaptarse a la diversidad de los usuarios.

10 ❖ **Optimización del diagnóstico clínico:**

- El sistema fuzzy combina los resultados de múltiples pruebas para generar un único diagnóstico clínico. Este enfoque sistemático y automatizado reduce la carga cognitiva para los profesionales de salud y minimiza el riesgo de diagnósticos inconsistentes.

15 ❖ **Automatización completa:**

- Integra pruebas acústicas, aerodinámicas y perceptuales en un dispositivo portátil, lo que permite un diagnóstico completo sin la necesidad de equipos de laboratorio voluminosos y costosos.
- El sistema fuzzy opera en un servidor remoto, permitiendo la **evaluación en tiempo real** desde cualquier lugar con conectividad.

20

❖ **Diseño específico para el español:**

- A diferencia de otros sistemas que se calibran para idiomas como el inglés o portugués, esta invención utiliza una **base normativa específica del español**, lo que garantiza una evaluación más precisa y contextualizada.

25 La lógica fuzzy es fundamental en esta patente porque **transforma un proceso subjetivo y multidimensional en uno sistemático y objetivo.**

30

REIVINDICACIONES

1. Método para la evaluación automatizada de la voz para lenguaje hablado español en un dispositivo móvil inteligente, caracterizado porque comprende los siguientes pasos:
- a.) Recibir la señal de voz;
 - b.) Grabar la voz, por medio de un dispositivo de grabación incorporado en el dispositivo móvil inteligente.
 - c.) Procesar automáticamente la información recibida mediante algoritmos que, a partir de las grabaciones de voz, realizan la representación gráfica y parametrización de la voz mediante tres tipos de test automatizados de evaluación vocal en fonoaudiología: a) test acústico, b) test perceptual y c) test aerodinámico. En donde se estiman al menos los siguientes parámetros:
 - Parámetros de perturbación Jitter y Shimmer (para evaluar las cuerdas vocales),
 - parámetro de Tiempo máximo de fonación (TMF) de sonidos vocálicos y consonánticos, el índice fonorespiratorio y el rango vocal de frecuencias que se realiza mediante el cálculo y visualización de un fonetograma para voz entrenada y no entrenada. Los parámetros estimados por los algoritmos mencionados, suministran los datos de entrada para el sistema fuzzy que permite la evaluación integral de la voz (ya que evalúa la voz a nivel acústico, perceptual y aerodinámico).
- Donde los test implementados son: Test de glissando, test efecto káiser, test de Fonetograma, test de frecuencia fundamental Jitter, shimmer; test de tiempo máximo de fonación y test de índice fonatorio;
- Donde la estimación del Parámetro Tiempo Máximo de Fonación (TMF) relaciona el flujo aéreo espiratorio con las cuerdas vocales.
 - Índice (s/z) – cociente fonorespiratorio: Que resulta de dividir el tiempo máximo de producción de una consonante /s/ entre el tiempo máximo de fonación para la /a/, donde el numerador estará relacionado con el TMF del volumen pulmonar y el denominador con la fonación de cuerdas vocales; que se obtiene dividiendo el tiempo máximo de fonación de la consonante /s/ entre el tiempo máximo de fonación de la vocal /a/ (tras inspiración profunda en un tono e intensidad

cómodos). Su valor teórico normal es 1, si es mayor de 1.4 es patológico e indicativo de un cierre laríngeo defectuoso).

- Donde el test acústico se implementa grabando la voz del paciente, posteriormente se reproduce esta grabación y dicha señal de voz se representa en una gráfica; al finalizar, para cada parámetro se muestra una gráfica con su valor máximo y mínimo indicando su normalidad o alteración.

- Donde el test aerodinámico se implementa calculando los tiempos máximos de fonación para una grabación de un sonido vocálico sostenido de la transcripción fonética de la vocal “a” en español y para una grabación durante el máximo tiempo de fonación de un sonido consonántico (sonido sordo) de la transcripción fonética de la “s” en español, para lo cual se utiliza un grabador de sonidos, con un botón de inicio y fin, mediante un cronómetro.

- Donde el test perceptual se implementa realizando un fonetograma para voz entrenada y no entrenada, para calcular el rango de frecuencias y de intensidades máximas y mínimas que puede alcanzar un paciente.

d.) Presentar de forma automática del resultado de las mediciones, gráficas y análisis de voz en una interfaz gráfica

e.) Almacenar los resultados en una base de datos

f.) Enviar los resultados de los parámetros estimados por los algoritmos de parametrización de la voz (algoritmos de análisis acústico, aerodinámico y perceptual), a un servidor web donde se aloja un sistema soportado en lógica fuzzy.

g.) Clasificar mediante el sistema de lógica fuzzy los parámetros calculados en las condiciones: normal o en riesgo/alterada, a partir de los parámetros de entrada.

2. Método para la evaluación automatizada de la voz para lenguaje hablado español en un dispositivo móvil inteligente caracterizado porque antes de grabar la voz se activa el registro del usuario y se selecciona el tipo de test a realizar.