

Abad Royo JM*

Abad Marco A**

Tisner Giraldo B***

Pérez Sánchez A*

Chamizo García J J*

*Médico del Servicio ORL
del Hospital Militar de Zaragoza
**Servicio Rehabilitación del
Hosp. Universitario Miguel Servet.
***Logopeda

Análisis digital de la voz. Conceptos básicos

Digital analysis of the voice. Basic concepts

RESUMEN

El presente artículo tiene como objetivo exponer los conceptos básicos más utilizados del análisis digital de la voz. Contiene ejemplos gráficos procedentes del programa SFS para Windows con los que se facilita la comprensión de estos conceptos.

PALABRAS CLAVE:

Trastornos de la voz. Análisis acústico.

SUMMARY

The present article has as objective to expose the used basic concepts in the digital analysis of the voice. It contains graphic examples coming from the program SFS for Windows with those that the understanding of these concepts is facilitated.

KEY WORDS:

Voice disorders. Acoustic analysis.

Introducción

El análisis digital de la voz es una pieza más en el diagnóstico de los trastornos de locución. El estudio de los mismos no es sencillo, porque implica aspectos de apreciación subjetiva tanto por parte del que habla, como por parte del que escucha. En definitiva podemos decir, siguiendo la definición del profesor Wendell Johnson (1967) en su libro *Speech Handicapped School Children*, (1) que "un sujeto tiene problemas de locución cuando el que le escucha desaprueba, se muestra ansioso, critica, en definitiva no le resulta agradable su voz, hasta el punto de que no presta debida atención al contenido por culpa de la forma del mensaje emitido". Los primeros años de la vida tienen una transcendencia fundamental en el desarrollo del lenguaje, siendo este uno de los aspectos más importantes en la vida del ser humano en sociedad.

En el tratamiento de estos problemas intervienen pediatras, psicólogos, neurólogos, logopedas, foniatras y, por supuesto, otorrinolaringólogos. La audición y todos los órganos que intervienen en la fonación, especialmente la laringe tienen una importancia vital en estos trastornos, por lo tanto es nuestra obligación como especialistas ORL conocer ampliamente estos temas.

Clásicamente los trastornos de la locución se clasifican en (1):

- 1) Errores de la articulación de palabras, bien sea por omisión, sustitución o distorsión.
- 2) Errores en el ritmo de conversación o lectura. Demasiado rápida o, por el contrario, demasiado lenta.

3) Falta de fluidez, por ejemplo el tartamudeo.

4) Trastornos de la voz: Alteraciones de la calidad de la voz por fallos en intensidad, tonalidad, timbre, flexibilidad, etc.

En este último apartado es donde tiene utilidad el análisis digital de la voz, al que hace referencia este artículo. El estudio ORL previo, tanto audiológico, como de vías respiratorias y fundamentalmente de laringe por observación directa y si es posible mediante técnicas de estroboscopia puede completarse con la ayuda de un ordenador y el programa correspondiente de análisis digital para mejorar el diagnóstico clínico y valorar posteriormente los resultados de los tratamientos aplicados ya sean médicos, quirúrgicos o foniátricos.

Material y métodos

Los registros sonoros empleados en este artículo fueron realizados a pacientes de la consulta cotidiana, con la única finalidad de completar su estudio laríngeo ORL, no cumplen, por tanto, las condiciones óptimas propuestas por "The Committee for Phonatory Function Test of the Japan Society of Logopedics and Phoniatrics" (2) en cuanto a las condiciones de recogida de los archivos sonoros: El paciente emite la voz, en este caso una letra "a" mantenida durante 2 segundos y en diferentes frecuencias, a 20 cm del micrófono, en una cámara sonoamortiguada, medida en dyn/cm². Sin embargo, en la práctica nos parecen suficientes.

Concretamente fueron recogidos por un ordenador PC Pentium II a 400 Mhz, con tarjeta Sound Blaster 16 bit, utilizando un micrófono convencional a una distancia de diez centímetros aproximadamente de la boca del paciente con la grabadora de sonidos de Windows'98. El local con aparente escaso ruido ambiente no está, sin embargo, especialmente insonorizado. Posteriormente se almacenaron como archivos "nombre.vaw".

A los pacientes se les hizo decir una frase determinada, las cinco vocales y una vocal mantenida. Para este artículo se han seleccionado diversos archivos "*nombre.vaw" procedentes de un paciente con voz normal, para compararlos con los de otro, que presenta una disfonía grave tras cirugía oncológica sobre glotis. Ambos constituyen ejemplos extremos que mejoran la comprensión de los conceptos expuestos.

El análisis de la voz se realizó con el programa SFS para Windows. Los gráficos adjuntos ayudan a la explicación de los conceptos básicos más utilizados.

Reconocimiento de la voz en un gráfico

El primer paso que vamos a dar para realizar nuestro estudio digital de la voz es reconocer la morfología que adopta ésta cuando empleamos un analizador de frecuencias. La voz es un fenómeno vibratorio que se representa sobre el papel o en el monitor como una mancha que contiene una frecuencia comprimida según los parámetros de intensidad y tiempo.

Las posibilidades de estudio dependen del programa de análisis de que se disponga:

- 1) En tiempo real. De gran interés para trabajar en presencia del paciente. Varios son los programas que nos ofrecen esta posibilidad: Dr. Speech (3) Exparam (4). Con ellos usamos las imágenes del monitor, que nos muestra las modificaciones de la señal acústica según es emitida por el paciente. Podemos estudiar frases completas, palabras o fonemas, que pueden tener utilidad en la identificación del mensaje, del sujeto que lo emite, del idioma. Desde el punto de vista clínico, la otra posibilidad de este método consiste en su utilización con fines de rehabilitación de la voz. También puede ser útil en la enseñanza del canto, haciendo que el paciente se ajuste a unos parámetros modelo (frecuencias, intensidad, formantes, armónicos).
- 2) Estudio de archivos de voz. La mayoría de trabajos publicados sobre este tema se basan en este sistema. Grabamos los sonidos que deseamos estudiar: vocales cortas o mantenidas algunos segundos, pequeñas frases, etc... Los datos, una vez elaborados, los guardaremos en un archivo informático o los reflejaremos a través de impresora. En principio, queremos valorar la estabilidad del sistema fonatorio y por tanto usaremos vocales mantenidas de manera constante durante algunos segundos. La vocal preferida depende del examinador y del idioma, siendo en nuestro idioma la "A" la más utilizada (5). Preciado López y col realizaron un estudio detallado de las características de las cinco vocales

españolas (6). En nuestros ejemplos se ha usado la "I" mantenida, vocal que empleamos para elevar y tensar más la laringe en la laringoscopia. La ventana de análisis será de algunos segundos, aunque el estudio puede limitarse a fracciones de éste.

En la Fig. 1 observamos diversos diagramas que corresponden a diferentes emisiones de voz. En este ejemplo el tiempo de estudio o ventana es de 6,4 segundos aproximadamente.

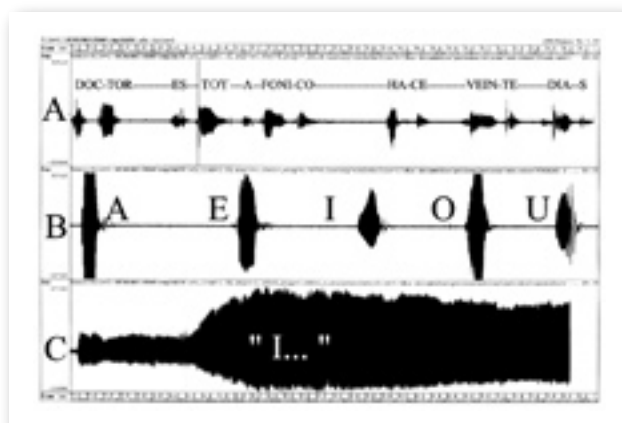


Fig. 1

En el diagrama A, el paciente ha emitido un mensaje: "Doctor, estoy afónico hace veinte días". Dicho gráfico representa doce fonemas distintos: "DOC - TOR - pausa - ES - TOY - A - FONI - CO - pausa - HA - CE - pausa - VEIN - TE - DIA-S ". La información contenida en esta frase se divide en golpes de voz o fonemas, representados como manchas de frecuencia comprimida dentro de unos parámetros de amplitud / tiempo que le dan esta morfología. El tiempo de emisión de esta frase es de 6,2 segundos. Dichos golpes de voz pueden ser distintos en la misma frase según el emisor y la entonación con la que los pronuncie.

El diagrama B corresponde a la emisión de las cinco vocales: "a - e - i - o - u" con los mismos parámetros amplitud / tiempo.

El diagrama C corresponde a la emisión de la vocal "i" mantenida durante unos segundos, en el que se aprecia un cambio de amplitud de la señal producido voluntariamente por el paciente.

LA VENTANA DE ANALISIS. El tiempo y en menor grado la intensidad, definen, por tanto, la ventana de estudio. De la misma forma que usamos un microscopio para fijarnos en un detalle anatómico, podemos limitar el tiempo a unidades menores a un segundo, por ejemplo: décimas, centésimas o milésimas de segundo, obteniendo así un gráfico de la misma voz aparentemente diferente pero esencialmente igual.

Cuando cerramos la ventana usando unidades de tiempo menor al segundo, evaluamos las características de la voz en un segmento dado. En la Fig. 2 usamos como ejemplo, un pequeño fragmento de 0,1 segundo del diagrama

C de la figura anterior. Como decíamos arriba, el paciente emitió una "i" mantenida con un aumento de la intensidad de manera voluntaria, la parte seleccionada corresponde al fragmento de intensidad alta. Se trata de un paciente con una voz clara y armónica. Con esto pretendemos realizar fundamentalmente un estudio clínico elemental de la voz en cuanto a las características de frecuencia, intensidad, timbre o calidad. Siendo éste tipo de estudio el más utilizado por los especialistas que están interesados en este tema dando pie a trabajos, protocolos o publicaciones en las que se revisan los resultados de cirugía (7), de rehabilitación foniátrica en niños (8), de grupos de pacientes que tienen diversas alteraciones en los órganos implicados en la fonación, desde vías respiratorias bajas, laringe y órganos resonadores. (9)

Lo que sigue es un apunte sencillo de los términos más empleados con este fin.

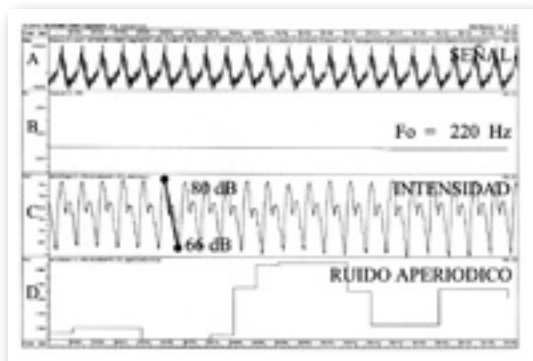


Fig. 2

Evaluación de la extensión de la voz

El estudio de la extensión de la voz ha sido más utilizado para el aprendizaje del canto, que para la evaluación clínica de la voz conversacional. Esta cualidad nos permite medir la capacidad individual de variación de frecuencia e intensidad para un sonido. Clásicamente se realiza eligiendo una nota musical con una frecuencia cómoda aumentándola hasta los valores más agudos que el examinado alcance y posteriormente disminuyéndola hasta los más graves. Tanto en hombres como en mujeres esta extensión suele estar entre 2 y 3 octavas (2). Fig. 3.

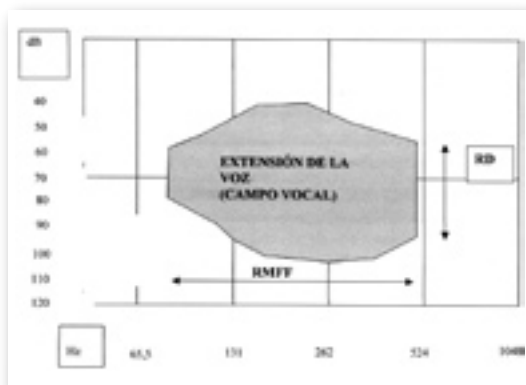


Fig. 3: Extensión de la voz.

La extensión cambia en los niños cuando entran en la adolescencia, se mantiene estable en la edad adulta y entra en declive cuando se sobrepasan los 60 años. (10)

El análisis clínico de la voz normal o patológica con un procedimiento similar constituye el CAMPO VOCAL del paciente, cuyos límites vienen definidos por el (RMFF) Rango Máximo de Frecuencia Fonatoria, que es el que va desde la frecuencia más grave, que es capaz de emitir el paciente, hasta la más aguda para una vocal dada, por ejemplo la "A" y el (RD) Rango Dinámico, que lo constituyen los valores más distantes de intensidad que puede alcanzar el sujeto explorado con un mismo tono. (11)

Evaluación de la frecuencia

Es el parámetro más estudiado, puesto que en esencia la voz y por extensión el sonido es un fenómeno vibratorio periódico.

Fluctuaciones no patológicas de la frecuencia fundamental de la voz. En la conversación humana, la lectura o el canto se producen cambios voluntarios en la entonación para mejorar su aspecto comunicativo y obtener un resultado más agradable para el oyente, por tanto, la frecuencia fundamental de los sonidos contenidos cambia continuamente. También lo hace dependiendo del fonema emitido.

Sin embargo, cada persona en razón de sus características anatómicas, edad y sexo habla con una voz en la que dominan unas frecuencias fundamentales determinadas.

Dicha frecuencia fundamental (fo) se obtiene emitiendo una voz, por ejemplo una vocal mantenida, de una manera confortable, sin la menor tensión laríngea y sin esfuerzo (2). Desde el punto de vista del análisis digital, es la frecuencia más baja con la que se repiten los mismos fenómenos ciclos de un sonido (12). Los analizadores nos la pueden dar como un número matemático, con su desviación estándar, tiempo de estudio, moda, promedio, máxima o mínima; o bien, como un gráfico lineal como el SFS – Windows, que aquí empleamos. Volviendo al registro del sonido de la Figura 2, en el diagrama superior (A) se aprecian los ciclos que constituyen la (fo) de ese fragmento de voz, aunque se adivina, por la falta de uniformidad del registro, que contiene otras frecuencias más altas incluidas en él mismo. El segundo diagrama (B) muestra una línea que corresponde a dicha (fo), que en este caso es 220 c / sg. Los siguientes diagramas (C) y (D) corresponden a fluctuaciones de la intensidad dependiendo de la fase del ciclo, en este caso entre 60 y 80 dB, y de análisis de ruido aperiódico respectivamente y serán comentados más adelante.

Perturbaciones de la (fo). Fluctuaciones patológicas de la frecuencia.

El término más conocido y estudiado es el JITTER (temblor), que mide las variaciones de la frecuencia de manera no voluntaria, por lo tanto se debe medir con un sonido sostenido de manera constante, por ejemplo una vocal. Estamos midiendo la estabilidad del sistema fonatorio del paciente obviando los cambios voluntarios.

Para ello se comparan las diferencias de frecuencia existentes entre ciclos cercanos, que en un sistema óptimo deberían ser "0". Cuanto mayor sea el porcentaje Jitter, mayor perturbación de la voz y por lo tanto menor calidad de la misma. (2)

FIGURA 4: El programa SFS Windows nos lo muestra como gráfico, tal como aparece en los diagramas B y E, frente a otros analizadores como Dr.Speech (2, 6, 7, 12, 13...), Multi-speech 3700 (14), Multi-Dimensional Voice (16), Sound Scope 1.0 GW Instruments (5), que lo hacen como datos numéricos con diferentes valores matemáticos: media (0,4 y 0,2 % en las vocales de pacientes normales) (6), máximos, mínimos, desviación estándar y múltiples índices o parámetros derivados de él. La valoración del efecto visual de nuestro ejemplo gráfico ha de tener en cuenta los parámetros en los que se dibuja: el tiempo analizado y sobre todo el porcentaje de perturbación (escala del margen izquierdo) siendo así muy dinámico para comparar diferentes muestras problema en un solo golpe de vista.

En esta Fig. 4 se comparan dos muestras de voz diferentes: Los diagramas A (señal), B (Jitter) y C (Shimmer) corresponden a una vocal "i" de 1,5 segundos de duración de un paciente con disfonía grave. El nivel de perturbación de frecuencia Jitter, medido en la escala adjunta, alcanza puntualmente valores de más de 75%, siendo su media aproximadamente del 25%. Los diagramas D (señal), E (Jitter) y F (Shimmer) corresponden a un fragmento de "i" mantenida, que sobrepasa el tiempo de ventana, del paciente con voz armónica. Sus valores de Jitter apenas sobrepasan el 1%.

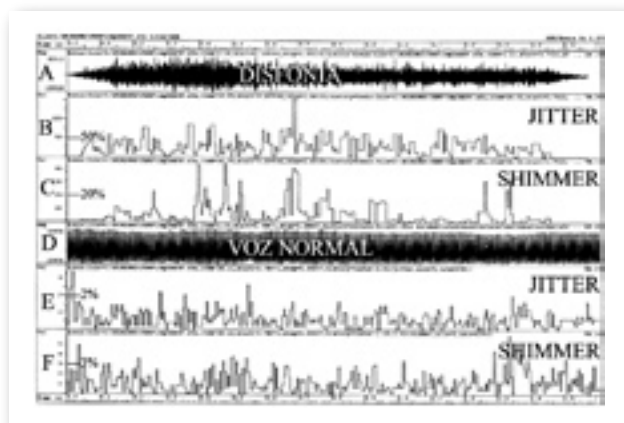


Fig. 4

Evaluación de la intensidad

La evaluación de la intensidad de la voz puede ser algo más compleja. En teoría debería seguirse un criterio uniforme como el comentado anteriormente (2), para la recogida de registros de voz. Sin embargo, en la práctica cada cual emplea métodos y medios ajustándose a lo que dispone.

Influyen en la intensidad de la voz:

- 1) La voluntad. El paciente puede modificar voluntariamente la intensidad de la emisión de voz. Figura 1, diagrama C.

- 2) La frecuencia de la voz emitida: en los tonos medios el paciente consigue mayor intensidad que en las frecuencias más bajas o graves y que en las más altas o agudas. FIGURA 3. Los propios ciclos del sonido producen continuas modificaciones de la intensidad, aunque se mantienen constantes en un sistema estable. FIGURA 2, diagrama C.
- 3) Retroalimentación auditiva (2): El paciente de manera inconsciente eleva o desciende la intensidad de la voz emitida según los estímulos auditivos ambientales. Es conocido el "fenómeno de Lombard", que consiste en el aumento de la intensidad de la voz emitida al estimular con ruido de enmascaramiento los oídos del paciente. Por el contrario, el paciente tiende a disminuir la intensidad de su emisión de voz al escuchar su propia voz por unos auriculares, constituyendo el "fenómeno de Fletcher". De la misma forma, todas las enfermedades que causan sordera pueden aumentar o disminuir esta intensidad.
- 4) Enfermedades de la personalidad y del sistema nervioso (depresión, neurosis, espasticidad) aumentan o disminuyen la intensidad.
- 5) Enfermedades de la glotis (parálisis, insuficiencia glótica, hipertensión) pueden también disminuir o aumentar la intensidad de la voz.

Medida de Fluctuaciones de la Amplitud o Intensidad: Se pueden medir los cambios en la amplitud de las diferentes voces a estudiar en el registro problema, independientemente de que tengan o no, significado patológico. Todas las voces tienen estas características, pero estos datos pueden poner en evidencia situaciones patológicas como alteraciones del ritmo, voces explosivas, etc.

- 1) Velocidad de Fluctuación de la amplitud: nº de fluctuaciones por segundo. FIGURA 2. Diagrama C
- 2) Tiempo de comienzo: Incluye desde cuando se alcanza el 10% de la intensidad de la palabra hasta el 90%. Figura 4. Diagrama A
- 3) Tiempo de intensidad estable. Mientras se mantiene el 90% de amplitud.
- 4) Tiempo de terminación. Desde el 90% hasta el 10% final.
- 5) Amplitud media de los fonemas, palabras, frases o incluso de la propia conversación.

Perturbaciones de la intensidad. Fluctuaciones patológicas de la intensidad:

Independiente de los cambios voluntarios o de entonación la estabilidad de la intensidad debería ser constante en una voz sin otros problemas. Los fallos de esta estabilidad se miden con el parámetro llamado "Shimmer" (luz trémula, brillo débil).

Por lo expuesto en los párrafos precedentes, se deduce que estas perturbaciones son más difíciles de valorar y estudiar que el Jitter frecuencial, aunque todos los sistemas existentes en el mercado dan estos valores dentro del análisis acústico de la voz.

El SHIMMER mide las pequeñas diferencias de intensidad existente entre las amplitudes de cada pico frecuencial. De la misma forma que el Jitter, cuanto menor sean estos valores mayor calidad de voz. Preciado (6) encuentra en su estudio sobre las vocales un porcentaje entre 2,08 y 1,9% en los pacientes normales. En la Figura 4, Los diagramas C y F corresponden al shimmer del ejemplo anterior. En la voz disfónica, diagrama C, el shimmer alcanza valores extremos cercanos al 20%, con una media cercana al 10 %. En la voz armónica de nuestro ejemplo, los máximos no llegan al 4%, siendo la media inferior a 1%. El efecto visual puede resultar engañoso, si no valoramos adecuadamente la escala porcentual situada en el margen izquierdo del gráfico.

Evaluación del timbre de la voz

El timbre de la voz está relacionado con la distribución acústica de la energía del espectro frecuencial. Y esta distribución es el resultado de la transformación del sonido laríngeo a su paso por las paredes y por las cavidades de resonancia. El resultado final es la estructura de la voz con diferentes FORMANTES. En definitiva estos formantes son refuerzos frecuenciales, que definen el resultado final de la voz y que hacen que el receptor la perciba como más o menos agradable.

Una voz rica, coloreada, brillante e incluso estridente en el peor de los casos, es así por tener abundantes formantes en frecuencias agudas. Las características físicas o constitucionales deciden dicho timbre, pero el entrenamiento puede desarrollar más determinados formantes. Así, algunos hablan de "formante de los cantantes" refiriéndose al que aparece reforzado en la frecuencia 2800 Hzs. (2, 15) Sin embargo, la voz tierna, sorda o velada es pobre en estos formantes agudos.

Por otro lado, la presencia de refuerzos frecuenciales periódicos constituyen los ARMONICOS, que están presentes en las voces normales, de acuerdo a las características del timbre de la voz del paciente o de posibles patologías que puedan alterar estos, hacer aparecer subarmónicos patológicos (16, 17) o, como ocurre frecuentemente, hacerlos desaparecer en determinadas frecuencias especialmente en agudos.

El programa de análisis de voz SFS Windows utilizado nos representa estos refuerzos frecuenciales o formantes de diversas maneras similares las presentaciones de otros programas:

En la Fig. 5, el diagrama A corresponde a la onda sonora del ejemplo de voz normal. Los FORMANTES son ofrecidos como datos numéricos (En nuestro ejemplo: F1 = 311 Hz, F2 = 2302 Hz, F3 = 2693 Hz, F4 = 3765 Hz), indicando la frecuencia concreta en la que se encuentra el pico más alto de cada uno de ellos. Sin embargo, este dato aislado tiene poco valor.

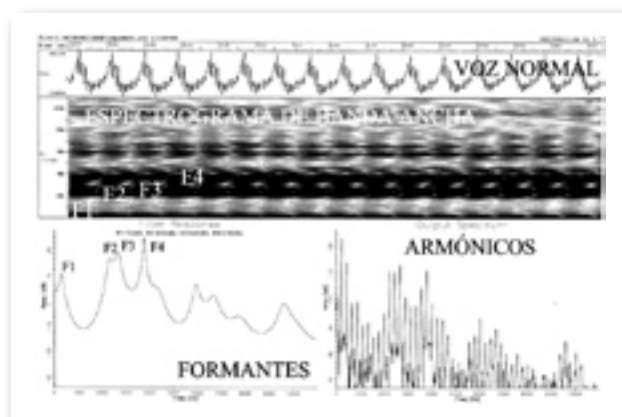


Fig. 5

Como diagrama lineal en el que podemos apreciar con un solo golpe de vista, la altura (dB) y la anchura (Hz) de la banda frecuencial, así como el pico más alto que define al formante, dando un efecto visual muy orientador sobre el tipo de voz que estamos analizando. Cuanto más alta y más estrecha sea la banda, y cuanto más agudo sea el pico, mayor claridad de la voz.

Como espectrograma de grises, en este caso de banda ancha de 300 Hz, en el que el realce gris define los distintos formantes al valorar con mayor precisión el tiempo que la frecuencia.

Los ARMÓNICOS se aprecian en esta misma figura en el diagrama de la derecha "Output Spectrum", donde aparecen como diagrama con picos en forma de agujas, que son muy evidentes y puntiagudos en un registro de voz normal.

Estos armónicos aparecen como resaltes grises, tal como se observa en la Fig. 6, en el espectrograma con filtro de 45 Hzs o de banda estrecha, puesto que valora con mayor precisión las frecuencias de la voz, que el tiempo en el que aparecen, a diferencia del de banda ancha. La riqueza en armónicos agudos y la mayor definición de los graves son cualidades positivas del timbre de la voz. Es conocida y empleada por algunos autores la clasificación de Yanagihara sobre estos espectrogramas según la definición de sus formantes y armónicos relacionándolos con diversos tipos de afonía. (18)

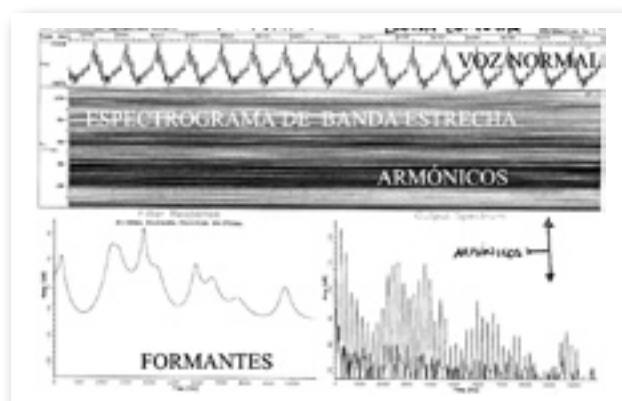


Fig. 6

La Fig. 7 representa el ejemplo de voz patológica. El diagrama A corresponde a la señal sonora, El B al espectrograma de banda estrecha, en el que no se aprecian prácticamente los realces grises de armónicos ni formantes. El diagrama lineal de la izquierda corresponde a unos formantes bajos, poco definidos y anchos, cuyo dato numérico es el siguiente: $F1 = 429 \text{ Hz}$, $F2 = 2098 \text{ Hz}$, $F3 = 3053 \text{ Hz}$, $F4 = 3457 \text{ Hz}$. El diagrama "Out Spectrum" de la derecha dibuja subarmónicos poco definidos y romos.

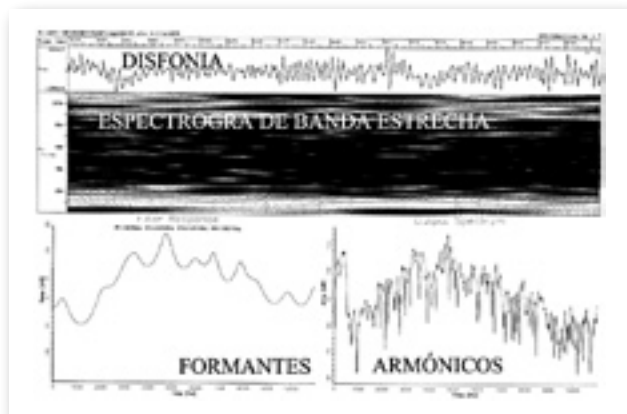


Fig. 7

El estudio de formantes nos proporciona amplias posibilidades de estudio en recién nacidos y niños en edad prelocutiva (19), en pacientes intervenidos en área ORL: Adenoides, amígdalas, uvulopalatoplastia, cáncer de laringe, etc. (14)

Gerard Guth (2) del Instituto de Fonética de Estrasburgo describió en su trabajo, como se puede comprobar que el timbre de todas las vocales es igual a nivel de las cuerdas vocales. Si se realiza esta misma medición a nivel del velo faríngeo, se observa como adquieren amplitud las frecuencias más agudas y disminuyen las graves. En la parte posterior de la boca, se realzan las frecuencias medias y en la parte anterior de la boca lo hacen nuevas frecuencias agudas y otros ruidos. Todas ellas han modificado el producto inicial procedente de la glotis hasta el resultado final por el que nosotros reconocemos las vocales y a su propio emisor.

Evaluación de la calidad de la voz

La calidad de la voz es una apreciación que implica aspectos subjetivos. En realidad la valoración que se hace para seleccionar a profesionales como locutores de programas radiofónicos, de televisión, dobladores de cine, oradores o políticos depende de un aspecto global que puede resumirse en que dicha voz sea agradable e incluso atractiva para que el mensaje o contenido de la conversación llegue al que escucha. No obstante desde el punto de vista objetivo, una voz de calidad tiene que tener obviamente una buena relación Armónicos / Ruido.

En condiciones normales, cualquier señal acústica de voz tiene un cierto componente de RUIDO por limpia que sea. De todos los parámetros relacionados con ruido

los más utilizados son los siguientes: En el programa Dr. Speech aparecen entre otros, HNR -Harmonic to Noise Ratio- (Valores de normalidad por encima de 22 – 29 dB) (6), NNE -Energic Noise Normalized- (Valores de normalidad más negativos que -10 dB) (13). En el Multi- Diensional Voice aparece, entre otros el opuesto a HNR, el NHR - Noise to Harmonic Ratio- , cuyos valores de normalidad en este caso deben ser inferiores a 13 % (14). EL programa SFS Windows nos aparece información sobre ruido en forma de gráficos lineales como aperiodicity con valores inferiores a 0,1 en caso de normalidad, bien de normenergy entre 0,4 y 1, bien de logenergy, entre 130 y 160. En ellos cuanto más alto es el valor, mayor ruido y, por tanto, peor calidad. En la FIGURA 2, diagrama D, en este caso el ruido contenido en esta señal de voz es bajo debido a su alta calidad. En los espectrogramas de grises se interpreta ruido cuando faltan los realces de armónicos o de los formantes y puede ser útil la clasificación de Yanagihara, en cuyos diferentes tipos (I,II,III y IV) el ruido sustituye progresivamente a armónicos y formantes según la disfonía del paciente explorado. (18, 20).

Desde un punto de vista clínico, cuando hablamos de calidad de la voz pretendemos excluir ASPECTOS NEGATIVOS que, en definitiva, tienden a hacerla desagradable aunque sea compresible. En este sentido los más empleados son los siguientes:

- 1) ESCAPE DE AIRE. (BREATHY). Se puede producir por diversas alteraciones de las cuerdas vocales como vibración asincrónica de las mismas, parálisis uni o bilaterales en abducción y sin compensar, hipotonía, defectos de sustancia, etc. Este escape de aire parece ser el responsable mayor del componente ruido de los archivos de voz. Pero dicho ruido puede corresponder, además de a una patología fonatoria, a un fallo en el sistema de recogida de archivo sonoro o/y también por contaminación acústica ambiental.
- 2) ASPEREZA O TENSIÓN LARINGEA. (HARSH) Puede deberse a hipertensión en los músculos del cuello. Ataque vocal brusco. Aperiodicidad en el espectro frecuencial. En definitiva, la voz áspera se interpreta como no agradable para el oído.
- 3) RONQUERA. (HOARSE). Es un sonido peculiar fácil de reconocer aunque difícil de definir, que incluye carraspera, pérdida de aire e hipertensión laríngea. En sus diversos grados puede ser reconocido por el paciente en los casos leves "esta no es mi voz" o ser muy evidente, laringitis crónicas o tumores laríngeos.
- 4) CHASQUIDO. (VOCAL FRY). Se trata de un crujido o chasquido que se aprecia cuando el paciente usa las extensiones más graves de su voz, por ejemplo al pronunciar una vocal, y que suele desaparecer cuando eleva la frecuencia de emisión.

El programa de análisis "Dr. Speech" nos presenta estas cualidades en un formato gráfico mediante un diagrama de barras muy sugestivo y fácil de interpretar, comparando los datos del paciente examinado con los parámetros de normalidad. (2) Fig. 8.

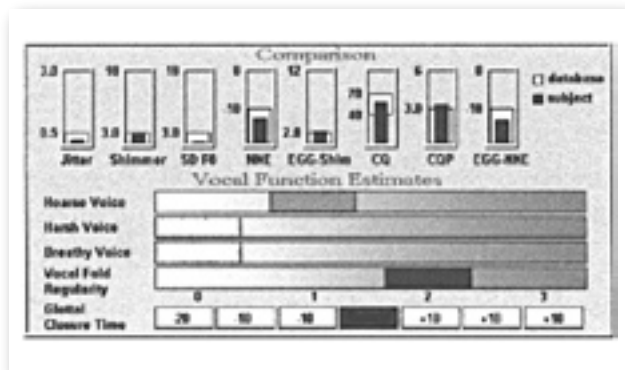


Fig. 8

Conclusiones

El análisis de la voz para el ORL tiene gran utilidad en la comprensión de lo que es la voz humana. Tiene gran interés para completar el diagnóstico en los pacientes con disfonía y sirve de gran ayuda en la rehabilitación de los mismos, así como para valorar objetivamente los resultados de los diferentes tratamientos propuestos.

Sin embargo, no hemos de perder de vista que el mejor analizador de voz que existe es el oído y el cerebro del que escucha y el que al final emitirá el juicio de valor definitivo: "me gusta o no me gusta".

Bibliografía

1. Melrose J. Trastornos de la comunicación verbal. En "Laringología II" de Paparella MM y Shumrick DA. Ed Panam. 2 Ed. Madrid. 1992: 7-18.
2. Jackson Menaldi C, Benvenuto M, Guevara W. Evaluación de la voz. En "La voz normal". Ed Panam. Buenos Aires. 1992: 145
3. <http://www.drspeech.com>
4. Suárez Guerra S. Procesamiento digital de voz, aplicaciones médicas para rehabilitación y entrenamiento. <http://www.galeon.com/exparam>.
5. Preciado López JA, Calzada Uriondo MG, García Cano FJ, Zabaleta López. Utilidad práctica del análisis físico de la vocal (A) en la valoración clínica del paciente disfónico. Libro de Resúmenes. Congreso extraordinario del 50º aniversario de la SEORL. Madrid 24-27 de octubre de 1999: 177

6. Preciado López JA, Calzada Uriondo MG, Zabaleta López M, García Cano FJ. Variabilidad en el análisis digital de la voz según la vocal analizada en pacientes normales y en pacientes disfónicos. Acta Otorrinolaring Esp 2000; 51(7) 618-828.
7. Núñez Batalla F, Baragaño Río L, Muñoz Pinto C, Maldonado Fernández M, Suárez Nieto. Aplicación del espectrograma de banda estrecha en fonocirugía. Libro de Resúmenes. Congreso extraordinario del 50º aniversario de la SEORL. Madrid 24-27 de octubre de 1999: 177
8. Boltezar IH, Burger ZR, Zargi M. Instability of voice in Adolescence: Pathologic Condition or normal development variation. J pediatr, 1997, 130: 185-190.
9. Huerta P, Vilaseca I, Fernández Planas A, Cardesin A, Bernal M. Calidad de voz en pacientes intervenidos de carcinoma glótico T1 por vía transoral con láser CO2. Comunicación. 53 Congreso Nacional SEORL. A Coruña. Junio 2002, 24-28.
10. Biondi S, Zappala M, Amato G, Consoli F. The senile voice. Acta Otorhinolaryngol Ital. 1992 Jan-Feb;12 (1): 69-79.
11. Nieto A, Del Palacio AJ, Lorenzo FJ, Vegas A, Cobeta I. Los ordenadores en el análisis de la voz: Aplicaciones clínicas. Acta Otorrinolaring Esp 1995; 46 (4): 241-245
12. López Gonzalo E, Hernández Gómez L. Tratamiento digital de la voz. www.gaps.ssr.upm.es/TDV/indexnfr.html
13. Fernández Liesa R, Damborenea Tajada J, Rueda Gormedino P, García García E, Leache Pueyo J, Campos del Alamo MA, Llorente Arenas E, Naya González MJ. Análisis acústico de la voz normal en adultos no fumadores. Acta Otorrinolaring Esp 1999; 50(2): 134-141.
14. Rosique M, Ramos JL, Campos M, Cervantes J. Análisis acústico de la voz en fistuloplastia fonatoria tras laringectomía total. Acta Otorrinolaring Esp 1999; 50 (2): 129-133.
15. Lundy DS, Roy S, Casiano RR, Xue Jw, Ewan J. Acoustic analysis of the singing and speaking voice in singing students. J Voice. 2000 Dec; 14 (4): 490-3.
16. González J, Cervera T, Miralles JC. Análisis acústico de la voz: Fiabilidad de un conjunto de parámetros multidimensionales. Acta Otorrinolaring Esp 2002; 53: 256-268.
17. Omori K, Kojima H, Kakani R, Slavik DH, Blaugrund SM. Acoustic Characteristic of rough voice: subharmonics. J Voice. 1997 Mar; 11 (1): 40-7.
18. Yanagihara N. Significance of harmonic changes and noise components in hoarseness. Journal of speech and Hearing Reseach. 1967; 10:531-41.
19. Wernike K, Monde W, Manfredi C, Bruscaglioni P. Developmental aspects of infant's cry melody and formants. Med Eng Phys. 2002. Sep-Oct; 24 (7-8) : 501-14.
20. Núñez Batalla F, Suárez Nieto C, Moro Melón M, Manrique Estrada C, Maseda Alvarez E, Carreño Villarreal M. Evaluación objetiva de la patología vocal en la infancia. Acta Otorrinolaring Esp 1999; 50(7): 525-529.