



Fonoaudiología
Universidad de La Frontera

Manual Básico de Evaluación de la voz para estudiantes de Fonoaudiología

Autores

Gerson Jara Cabrera
Rachel Cabrera Fuentes



Manual Básico de Evaluación de la voz para estudiantes de Fonoaudiología

Este material fue desarrollado gracias al apoyo de la Dirección de Desarrollo Curricular y docente de la Universidad de La Frontera y sus proyectos de Desarrollo e Innovación Docente de Pregrado Código: DID21-0020

Autores:

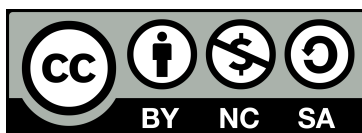
Gerson Jara Cabrera

gerson.jara@ufrontera.cl

Rachel Cabrera Fuentes

rachel.cabrera@ufrontera.cl

ISBN Versión digital: 978-956-414-157-2



Este material puede ser distribuido, copiado y exhibido por terceros si se muestra en los créditos y solo para fines educacionales y de divulgación científica. No se puede obtener ningún beneficio comercial y las obras derivadas tienen que estar bajo los mismos términos de licencia que el trabajo original.



Contenidos

<u>Presentación del manual</u>	1
<u>PARTE I: CONCEPTOS GENERALES DE LA EVALUACIÓN DE LA VOZ</u>	2
<u>Introducción de la sección</u>	3
<u>Evaluación Fonoaudiológica de la voz</u>	5
<u>Acústica de la voz</u>	10
<u>Análisis acústico de la voz</u>	17
<u>PARTE II: DATOS NORMATIVOS</u>	42
<u>Introducción de la sección</u>	43
1. <u>Notas Musicales y sus equivalentes frecuencias en Hertz</u>	44
2. <u>Clasificación de las voces según extensión fisiológica</u>	45
3. <u>Frecuencia Fundamental</u>	46
4. <u>Tono Medio Hablado</u>	48
5. <u>Valores promedio de F0 y formantes para las vocales del español Rioplatense</u>	49
6. <u>Rango fonacional máximo</u>	50
7. <u>Perturbación de la Frecuencia Fundamental</u>	51
8. <u>Intensidad vocal</u>	52
9. <u>Niveles de Intensidad máxima</u>	53
10. <u>Presión Sonora</u>	54
11. <u>Perturbación de la amplitud</u>	55
12. <u>Presión subglótica</u>	56
13. <u>Presión sonora a nivel intraoral</u>	56
14. <u>Tiempos máximos de fonación</u>	57
15. <u>Índice s/z</u>	58
16. <u>Valores índice s/a y su implicancia clínica</u>	59
17. <u>Flujo medio aéreo</u>	59
18. <u>Promedio de velocidad de flujo e implicancias clínicas</u>	60
19. <u>Presión subglótica e implicancia clínica</u>	61
20. <u>Umbral de presión de la fonación</u>	61
21. <u>Cociente de contacto</u>	62
22. <u>Valores normativos del PRAAT</u>	62
23. <u>Escala de pérdida de armónicos</u>	63
24. <u>Clasificación de asenia</u>	63

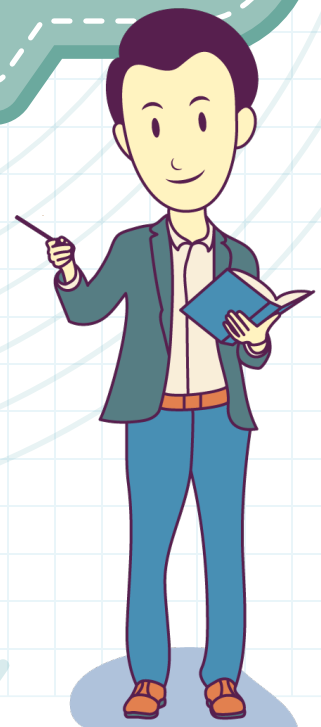
Contenidos

25.	Grados de severidad según valoración perceptual	64
MANUAL BÁSICO DE PRAAT PARA ANÁLISIS DE VOZ		65
Introducción de la sección		66
1.	Grabación de la señal	67
2.	Análisis de la señal	74
3.	Espectrograma	78
4.	Análisis de Formantes	80
5.	Análisis espectral: Espectro FFT (Power Spectrum)	83
6.	Análisis Espectral: Espectro LPC (Código de predicción lineal)	87
7.	Análisis Espectral: Cepstrum	92
8.	Análisis de la intensidad	97
9.	Análisis de la frecuencia	99
10.	Medidas de perturbación de la voz	101
11.	Espectro promedio a largo plazo: Long-Term-Average Spectrum (LTAS)	103
12.	Obtención de valores numéricos	109
REFERENCIAS		119
ANEXOS		122

Presentación del manual

El objetivo del presente trabajo, es servir de apoyo a los estudiantes y guiarlos en lo que respecta al proceso de evaluación de la voz, posterior análisis y realización de informes de evaluación fonoaudiológica.

Para tal propósito, el presente manual se dividirá en tres partes: La primera tiene por finalidad describir el proceso de evaluación vocal, cuales son las principales herramientas y procedimientos de evaluación y las medidas acústicas de la voz más utilizadas. En segunda instancia, se incluirá toda la data normativa de la voz existente en la literatura especializada. Por último, se describirá un manual de procedimientos de cómo realizar grabaciones acústicas y análisis de la voz utilizando el programa PRAAT.





PARTE I: CONCEPTOS GENERALES DE LA EVALUACIÓN DE LA VOZ

Introducción de la sección

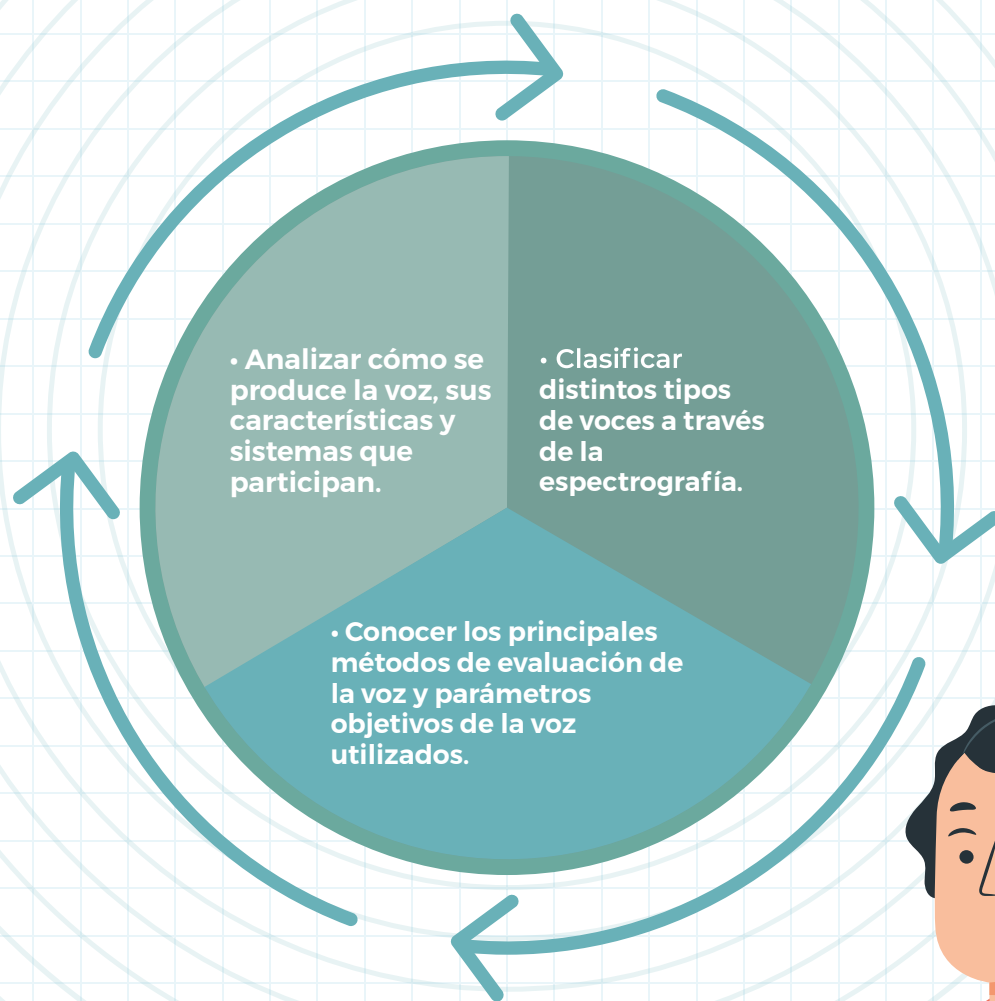
La evaluación fonoaudiológica de la voz cobra cada vez mayor relevancia clínica y requiere de profesionales bien entrenados con amplio conocimiento de diversas áreas como lo son anatomía, fisiología, física acústica, informática, procesamiento de imágenes, etc., ya que debemos ser capaces de relacionar aquello que escuchamos con su respectivo correlato fisiológico.

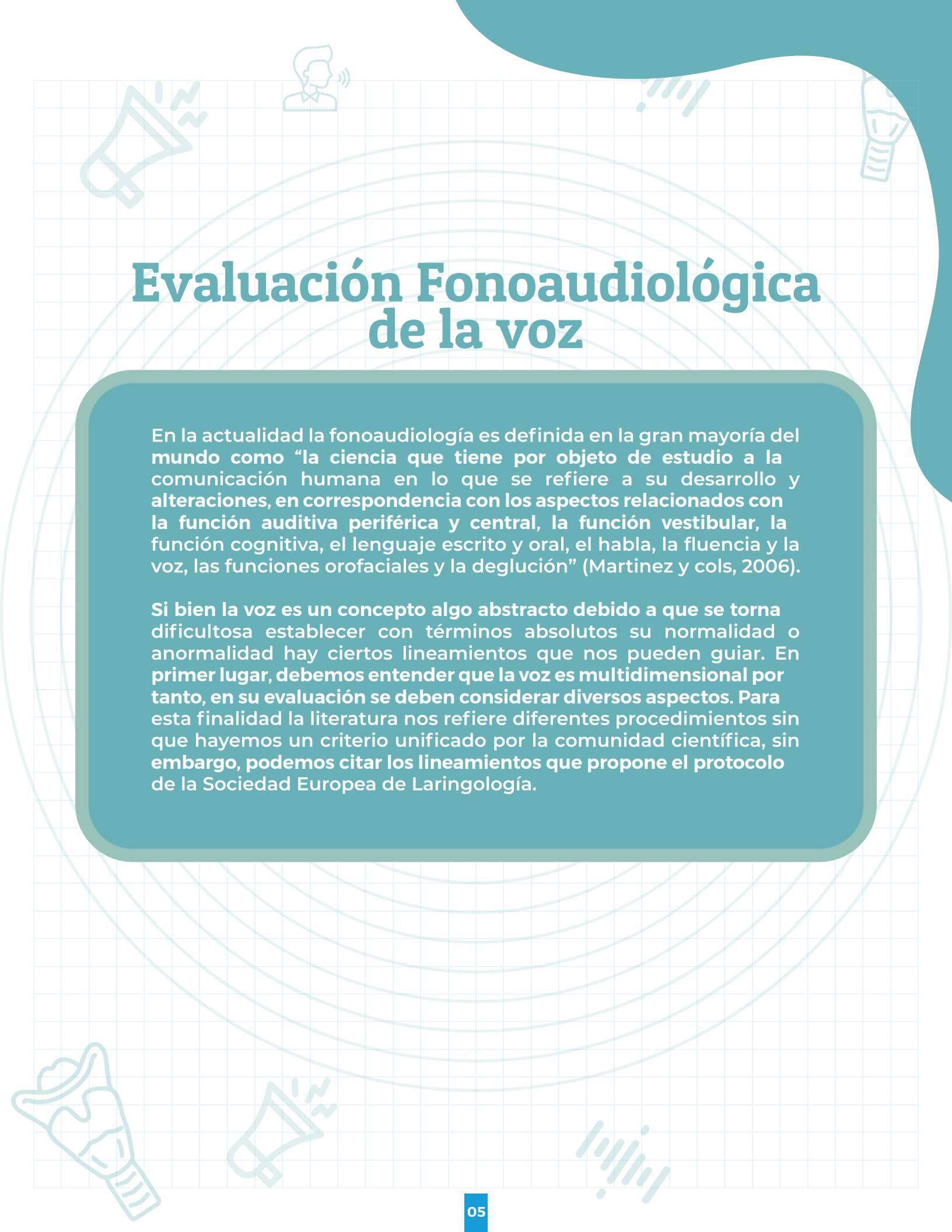
La evaluación de la función vocal o calidad vocal no es tarea sencilla, ya que como dice Aronson (1985) “el concepto de voz normal con un criterio objetivo absoluto no existe. Sin embargo, cada grupo étnico tiene parámetros psicoacústicos que lo acercan a la normalidad cuando un individuo tiene un timbre agradable, tono adecuado al sexo y a la edad, volumen apropiado a las necesidades vocales y flexibilidad que le permitan mostrar su propia imagen en un todo”.

La evaluación de la voz se vuelve un tanto confusa al analizar la calidad vocal, sobre todo cuando el mismo concepto de calidad vocal tampoco es del todo claro en términos perceptivos y acústicos. Esto se ve incrementado además por la nomenclatura utilizada por los clínicos de voz y que tiene un enfoque más bien perceptivo y deja mucho a la subjetividad: áspera, soplada, tensa, estridente, chirreante, clara, oscura, raspada, etc (Farías, 2016).

Sin embargo, este proceso de evaluación vocal, ha ido evolucionando y cambiando a lo largo de los años, de la mano con el desarrollo de nuevas tecnologías tanto en la exploración visual como con el desarrollo de software especializados para dicho análisis.

Esta primera parte del manual está orientada a que el estudiante desarrolle los siguientes resultados de aprendizaje:



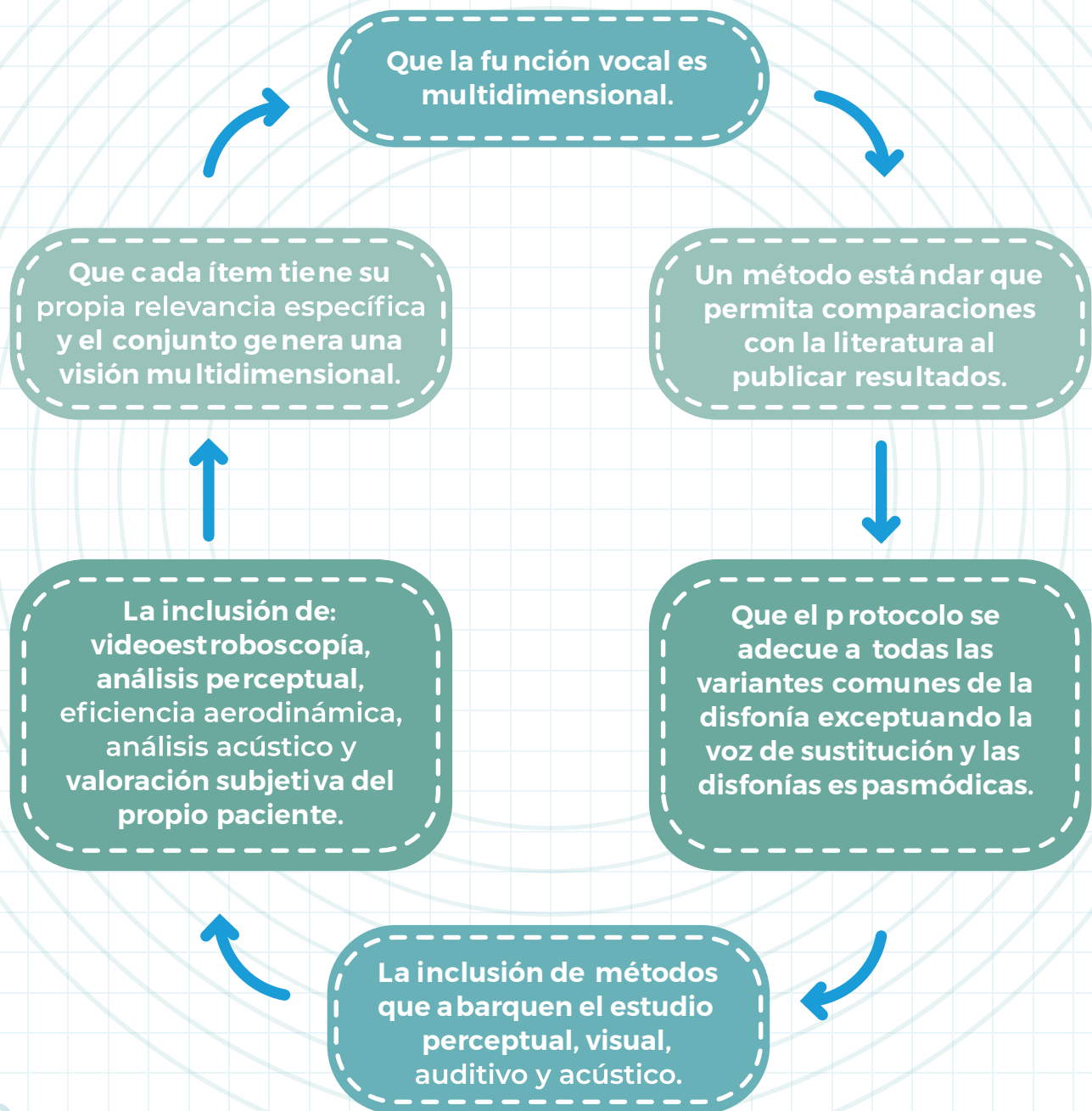


Evaluación Fonoaudiológica de la voz

En la actualidad la fonoaudiología es definida en la gran mayoría del mundo como “la ciencia que tiene por objeto de estudio a la comunicación humana en lo que se refiere a su desarrollo y alteraciones, en correspondencia con los aspectos relacionados con la función auditiva periférica y central, la función vestibular, la función cognitiva, el lenguaje escrito y oral, el habla, la fluencia y la voz, las funciones orofaciales y la deglución” (Martínez y cols, 2006).

Si bien la voz es un concepto algo abstracto debido a que se torna dificultosa establecer con términos absolutos su normalidad o anormalidad hay ciertos lineamientos que nos pueden guiar. En primer lugar, debemos entender que la voz es multidimensional por tanto, en su evaluación se deben considerar diversos aspectos. Para esta finalidad la literatura nos refiere diferentes procedimientos sin que hayamos un criterio unificado por la comunidad científica, sin embargo, podemos citar los lineamientos que propone el protocolo de la Sociedad Europea de Laringología.

Según el protocolo básico del Comité de Foniatría de la Sociedad Laringológica Europea, la valoración de la voz patológica (Dejonckere, 2000) deberá contemplar:



En la actualidad la fonoaudiología es definida en la gran mayoría del mundo como “la ciencia que tiene por objeto de estudio a la comunicación humana en lo que se refiere a su desarrollo y alteraciones, en correspondencia con los aspectos relacionados con la función auditiva periférica y central, la función vestibular, la función cognitiva, el lenguaje escrito y oral, el habla, la fluencia y la voz, las funciones orofaciales y la deglución” (Martinez y cols, 2006).

Si bien la voz es un concepto algo abstracto debido a que se torna dificultosa establecer con términos absolutos su normalidad o anormalidad hay ciertos lineamientos que nos pueden guiar. En primer lugar, debemos entender que la voz es multidimensional por tanto, en su evaluación se deben considerar diversos aspectos. Para esta finalidad la literatura nos refiere diferentes procedimientos sin que hayamos un criterio unificado por la comunidad científica, sin embargo, podemos citar los lineamientos que propone el protocolo de la Sociedad Europea de Laringología.

Escalas de valoración audioperceptual de la voz

Escala	Autores
GRBAS	Hirano, 1981.
RASAT	Pinho, 2002.
CAPE-V	Kempster et al. 2008
VPAS	Laver, Mackenzie, 2007.

Escalas de autoreporte vocal. (Farías 2016)

Escala	Autores
Voice Handicap index VHI-30	Jacobson et al. 1997
Voice Handicap iNDEX 10 VHI-10	Rosen et al. 2004
Pediatric Voice Handicap Index pVHI	Zur et al. 2007
Singing Voice Handicap Index SVHI	Cohen et al. 2007
Voice Activity and Participation Profile VAPP	Ma & Yiu, 2001
The Voice-related quality of life V-RQoL	Hogikyan & Sethuraman, 1999
Voice Outcome Survey VOS	Glickich et al. 1999
Pediatric Voice Outcomes Survey PVOS	Hartnick, 1999
VoSS	Deary et al. 2003

Farías, 2016

El inconveniente que pudiera presentar este tipo de evaluación es que depende mucho del oído experimentado del evaluador para emitir un juicio correcto. Es así como a lo largo de los años y con el desarrollo de nuevas tecnologías se han ido agregando nuevos procedimientos de evaluación vocal como son por ejemplo el análisis y exploración anatómico funcional de los pliegues vocales mediante Telelaringoscopia, videoestroboscopia, videoquimografía, etc. y el desarrollo de software y laboratorios de voz que permiten el análisis acústico de la voz, espectrografía, electroglotografía, análisis aerodinámicos que permiten al fonoaudiólogo complementar su evaluación dando una mayor objetivación a la alteración vocal de tal forma de emitir mejores diagnósticos y juicios clínicos.



No obstante, es importante recalcar que el oído bien entrenado del evaluador sigue siendo la principal herramienta de evaluación y todos estos otros procedimientos vienen a complementar la evaluación y aportar un grado de objetividad. Cualquier metodo utilizado de manera aislada tiene sus ventajas y desventajas por lo mismo es importante recalcar la utilización de ellos como un todo y que se vienen a complementar unos con otros.

De acuerdo a Jackson-Menaldi (2002), la evaluación persigue diferentes objetivos:

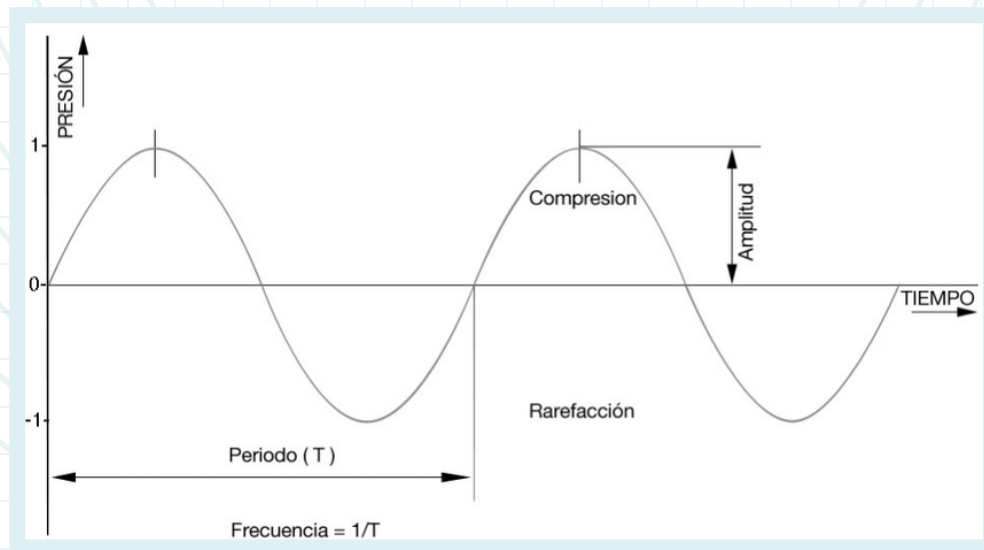
- **Completar y corroborar un diagnóstico médico.**
- **Ayuda a aclarar la causa del problema vocal.**
- **Describe la función de la voz en relación con la patología vocal y determina la severidad del problema y su pronóstico.**
- **Educa al paciente sobre el problema y acerca de su responsabilidad en el desarrollo de éste y en su resolución.**
- **Ayuda a apreciar la importancia del problema.**
- **Es el primer contacto con el paciente, que permite ponerlo al tanto de los componentes funcionales y orgánicos anormales.**
- **Permite evaluar el tratamiento.**



Acústica de la voz

Como se mencionaba anteriormente hoy en día el análisis acústico de la voz es una herramienta más que complementa la evaluación del fonoaudiólogo, dando cierto grado de objetividad que nos permitirá cuantificar la disfonía y además ser un medio para medir la efectividad de los tratamientos. Para esto es sumamente importante contar con el equipamiento y las condiciones necesarias para poder realizar grabaciones de alta calidad, lo cual veremos más adelante en el apartado de condiciones básicas de grabación.

El análisis acústico lo que hace es medir los parámetros físicos de la voz, los cuales tienen su correlato perceptual, de ahí la importancia como se mencionaba anteriormente del oído experimentado del evaluador. A continuación, se mostrará un cuadro de los principales parámetros acústicos y su correlato perceptual adaptado por Farías (2010).



Parámetros Acústicos	Parámetros físicos	Evaluación perceptual	Evaluación acústica
Intensidad	Amplitud	Intensidad habitual utilizada y variabilidad posible: voz suave, normal o fuerte.	Intensidad medida en dB (con decibelímetro o computadora). Energía. La fluctuación de amplitud se mide con el shimmer.
Altura o tono	Frecuencia	Tono de voz hablado determinado con teclado: normal, agravado o agudizado. Extensión tonal: posibilidad de emitir tonos desde graves hasta agudos (frito, modal y falsete).	F0 en hertz, rango de frecuencias medido en hertz. Su fluctuación se mide con el jitter. Fonetograma: gráfico de campo vocal que cruza todas las frecuencias con todas las intensidades posibles.
Duración	Tiempo	Tiempo máximo de fonación /a/ /i/ /u/ Índice s/z (s/a).	Tomado en mseg de la grabación digitalizada de voz.
Calidad	Timbre	Escala perceptual CRBASI: grado, aspereza, soplo, astenia, tensión e inestabilidad.	Espectrograma de banda ancha (formantes), de banda estrecha (armónicos). Relación armónicos y ruido.

Creo es importante que antes de adentrarnos en el análisis acústico y descripción de cada una de sus medidas, es necesario realizar una pequeña introducción respecto de la acústica de la voz.

La acústica es una rama de la física que se encarga de estudiar el sonido, las vibraciones y sus características particulares. El sonido se puede definir como una vibración mecánica que se propaga a través de un medio elástico y denso. Cuando esta vibración provoca una excitación en los mecanismos del oído se habla de percepción del sonido o sensación auditiva (Saldías, 2017).

La voz corresponde al elemento del habla que nos otorga la señal acústica vibratoria sobre la cual llevamos a cabo el mensaje hablado y que nos permite producir sonidos con intención (Stemple, Glaze y Klaben, 2010).

Para que se produzca esta vibración mecánica es necesario que exista una fuente de energía que dé inicio a la vibración lo que en el caso de la voz serían los pliegues vocales. La cantidad de ondas y la forma de estas, estarán determinadas por las formas en que este objeto se mueva, es decir la onda resultante de una fuente depende de los modos de vibración que ésta tenga (Saldías, 2017).

Los modos de vibración se definen como la diversidad de formas en que puede oscilar una onda estacionaria dentro de un sistema oscilante (Alton, 2006).

Una vez comenzada la vibración por la fuente, esta se transmite a las partículas de aire adyacentes, las que a su vez transmiten la vibración a nuevas partículas contiguas. Las partículas no se desplazan con la perturbación, sino que simplemente oscilan alrededor de su posición de equilibrio, de esta forma se crean zonas donde las partículas de aire se comprimen y la presión de aire es mayor a la atmosférica y otras zonas donde las partículas de aire se dispersan y la presión de aire es menor que la presión atmosférica. (Sataloff, 2005).

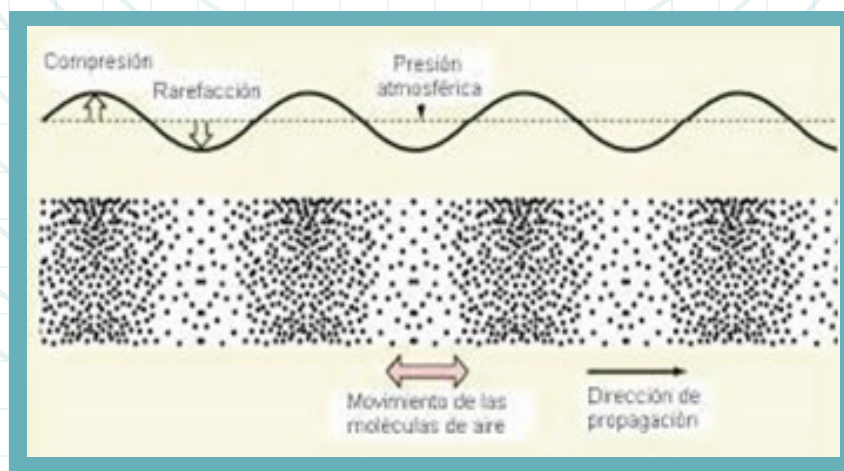


Imagen 1. Propagación de una onda

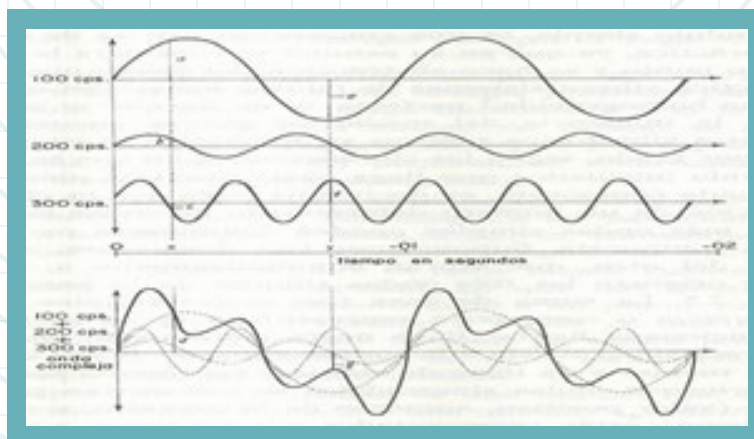


A esta vibración mecánica se le conoce como ondas y tiene ciertas características como son frecuencia, amplitud y longitud de onda. El tiempo que demora una partícula en realizar un movimiento completo de oscilación se le conoce como ciclo. La frecuencia, será el número de ciclos por segundo. La amplitud corresponde a la máxima distancia que alcanza la partícula en movimiento desde su punto de reposo y la longitud de onda corresponderá a la distancia en la que una onda viaja en el tiempo completando un ciclo (Alton, 2006).

Existen básicamente dos ondas, las simples y las complejas. Las simples corresponden a fuentes sonoras que tendrían solamente un modo de vibración y se representan por una onda sinusoidal que tiene solamente una frecuencia y amplitud determinada. También se les conoce como tonos puros y un ejemplo clásico sería el tono de un diapasón.

Las ondas complejas son aquellas que están compuestas por varias ondas simples. Una frecuencia fundamental o primer armónico y un conjunto infinito de frecuencias múltiplos de la frecuencia fundamental, denominados sobretonos. Generalmente la frecuencia fundamental es la que presenta mayor amplitud (Carrion, 1998).

Imagen 2. Imagen de una onda compleja a partir de la sumatoria de varias ondas simples.





A su vez las ondas complejas pueden dividirse en cuasi periódicas y periódicas. Las periódicas son aquellas que tienen periodos iguales y las cuasi periódicas poseen periodos aproximadamente iguales. Los sonidos vocálicos son cuasi periódicos.

Toda onda compleja periódica, puede representarse como la superposición de vibraciones periodicas simples, cuya frecuencia fundamental es la frecuencia de repetición de la onda. Esto es lo que se conoce como Teorema de Fourier. Mediante este análisis, por tanto, se pueden obtener los tonos puros que componen la onda, por el proceso inverso síntesis de Fourier, se reconstruye la onda compleja a partir de una serie de ondas simples (Farías, 2016).

Si llevamos todos estos conceptos a la producción de la voz, vemos que se requiere en un principio de la vibración de los pliegues vocales que serían la fuente fonatoria productora de sonido. Este sonido producido en la laringe es complejo y está compuesto de una frecuencia fundamental que sería la frecuencia más baja del espectro y un conjunto de frecuencias acompañantes múltiplos de esta frecuencia fundamental denominadas armónicos y corresponden a modos de vibración de los pliegues vocales.

Sin embargo, este sonido puede ser muy similar entre laringes de distintas personas, por lo que la individualidad de nuestra voz no está determinada necesariamente por las características de los pliegues vocales, sino que por las características resonanciales de cada individuo, las que están determinadas por el tracto vocal (McCoy, 2015).

El fenómeno de resonancia implica que un cuerpo que vibra hace vibrar a otro cuerpo próximo, como es el caso de la relación existente entre los pliegues vocales (fuente) y el aire contenido dentro del tracto vocal (cuerpo próximo) (Stemple y cols, 2010).



Estas vibraciones producidas, generan un fenómeno conocido como **sensación de vibración**, en el cual estas pueden ser percibidas en distintas partes del cuerpo. En la medida que este proceso de conversión de energía aerodinámica en acústica a nivel de la glotis sea más eficiente, es posible que estas vibraciones se distribuyan y se perciban en toda la cabeza, cuello y tórax. Al contrario, si la conversión de energía en la glotis es menos eficiente, entonces las vibraciones tenderán a permanecer más bien localizadas, como en la laringe por ejemplo (Titze y Verdolini, 2012). De esto nace el concepto de colocación de la voz, que en términos científicos es más una ilusión que un hecho físico real, ya que no es posible colocar el sonido en distintas partes, más bien lo que sucede es que podemos percibir el resultado de estas vibraciones en distintas partes del cuerpo.

Como se mencionaba anteriormente el sonido originado a nivel de la fuente (pliegues vocales) es débil, por tanto, la resonancia es la encargada de amplificarlo y dotarlo de características propias de la voz humana. En este proceso participan estructuras del tracto vocal como la laringofaringe, faringe, cavidad oral y cavidad nasal. Para que exista una resonancia adecuada debe existir un equilibrio en la resonancia de estas distintas estructuras.

Este sonido complejo compuesto de una frecuencia fundamental y un conjunto de armónicos al pasar por las distintas cavidades de resonancia se va modificando y es así como aparece otro concepto denominado **formantes**. Los armónicos producidos en la fuente van perdiendo energía acústica a medida que se van alejando de la F0 a razón de 12 Db por octava aproximadamente y se vuelven débiles. Sin embargo, al pasar por el tracto vocal hay zonas donde estos serán amplificados o atenuados dependiendo de la configuración de las cavidades de resonancia.

En ese sentido, los armónicos que coincidan con la resonancia de una cavidad determinada serán amplificados, mientras que los demás serán atenuados. De aquí nace el concepto de formantes que corresponden a resonancias del tracto vocal, es decir se asocian a la amplificación y mejora del sonido vocal (McCoy, 2015). Por tanto, los formantes corresponden a los modos de vibración de las partículas de aire contenidas en el tracto vocal (Saldías, 2017).

El tracto vocal genera varios formantes desde las frecuencias más bajas a las más altas. A continuación, se describirá cada uno (Farías, 2016):



Los formantes F3, F4 y F5 se asocian a las características individuales de la voz de cada sujeto lo que se conoce como timbre extra vocálico y pueden aumentar el poder vocal o proyección del sonido (Sataloff, 2005).

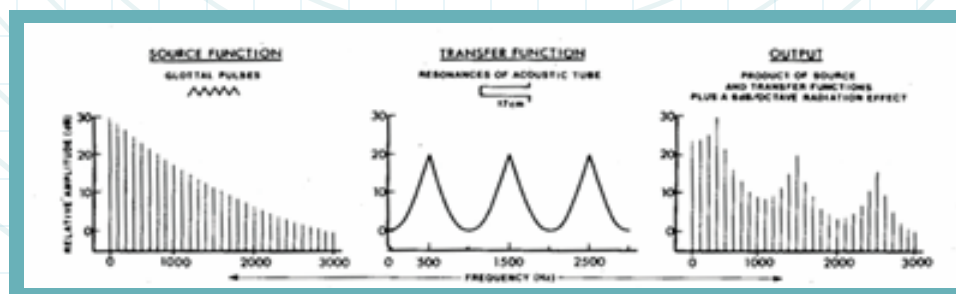


Imagen 3. Imagen del sonido al pasar por la fuente, tracto vocal y output final.



De lo anterior nace el concepto de color de la voz el cual es un aspecto perceptual y que se relaciona con la configuración del tracto vocal y cambios en la amplificación de los armónicos. Una voz clara es aquella que es producida por un tracto vocal más abierto y una laringe más elevada, mientras que una voz oscura es aquella producida por un tracto vocal más cerrado y una laringe más baja (Cobeta y cols, 2013).

Por último, cabe destacar que todo esto no sería posible si no hubiera armónicos que amplificar, por tanto, se necesita de una adecuada eficiencia glótica para convertir la energía aerodinámica en acústica. De aquí se desprende otro concepto que se denomina brillo, el depende del grado de aducción de los pliegues vocales.

Análisis acústico de la voz

Como expresamos anteriormente, el análisis acústico es una herramienta más, que viene a complementar nuestra evaluación. La persona que está aplicando el análisis acústico es importante que sepa lo que está haciendo y porque, así como también saber interpretar los resultados entregados por el software y correlacionarlos con lo perceptual.

El análisis acústico, por tanto, viene a dar cierto grado de objetividad a nuestra labor lo que trae ciertas ventajas, pero a su vez también tiene ciertos inconvenientes, los cuales se detallan a continuación (Casado, Adrián y cols, 2002):

Ventajas



- Identifica los componentes vocales responsables de la disfonía, aumentando la precisión del diagnóstico.
- Cuantifica los componentes vocales responsables de la disfonía, para así poder establecer una estadificación de la disfonía. En este sentido se le ha querido dar al análisis acústico una significación semejante a la que la audiometría tiene en la patología del oído.
- Valora la evolución de la disfonía tras los tratamientos aplicados (tanto a corto como a largo plazo). Esta valoración permitirá realizar cambios en el tratamiento.
- Es sencillo en la utilización de los equipos, poniendo al alcance de todos, los recursos necesarios para hacer análisis acústicos vocales. A esto hay que añadir la universalización del ordenador personal, que facilita enormemente el cálculo de los parámetros que hace años había que medirlos a mano.

Inconvenientes



- La emisión vocal por parte del paciente puede no tener siempre las mismas características, con lo que cabrían diferencias incluso entre dos exploraciones consecutivas. Esto crea una incertidumbre que resta credibilidad a la prueba.
- Inutilidad del análisis acústico vocal si no se hace en el contexto de un abordaje clínico global. La propia facilidad del análisis acústico puede llegar a crear confusión al no matizar e interpretar los hallazgos.
- Falta de unos estándares de normalidad para cada una de las pruebas, con lo que la normalidad se obtiene muchas veces para cada uno de los equipos.

Otro aspecto importante para el análisis acústico, es que si queremos que los resultados sean fiables se requieren de ciertas condiciones mínimas de grabación y equipamiento. En ese sentido la National Center for voice and Speech (NCVS) 2012, sugiere lo siguiente:

Micrófono



- Omnidireccional (ideal para ambientes silenciosos).
- Micrófono de condensador (omnidireccional o cardioide).
- Sensibilidad mínima de -60Db
- **Respuesta de frecuencia plana.**
- No usar filtro anti-pop.

Distancia de grabación



- Para la toma de muestras vocales sostenidas una distancia de máximo 10 cm entre la boca y el micrófono.
- Para habla conectada una distancia de 30 cm.
- Angulo de 45° o 90° desde el eje de la boca para así reducir ruido aerodinámico.
- Con sonómetro asegurarse que todos emitan a la misma intensidad.

Ambiente



- Para investigación debe ser en cámara silente. (construcción pesada, paredes dobles, piso flotante, iluminación y ventilación silenciosa y compartimientos para cables que permitan conexión entre equipos del exterior e interior.
- Que la cámara no este directamente unida a la estructura del edificio, sino más bien que flote.

Ruido



- El ruido ambiente debe ser menor a 50 decibeles.
- 40 decibeles es el ideal para cámaras cabinas silentes y menos de 30 decibeles para salas.
- Contar con una evaluación espectral del ruido presente en la cabina o habitación.
- Para determinar el nivel de ruido se sugiere el uso de un sonómetro de clase I siendo el C o el linear los más aptos para las grabaciones de voz.

Pre amplificador/interfaz



- Convertidor AD de 16 bit con amplificadores y filtros que posean una relación señal ruido en el rango de 85-95 decibeles.
- Frecuencia de muestreo puede estar en el rango entre 20 a 100 KHz.
- Se sugiere usar 44.100 Hz de frecuencia de muestreo.
- Que la interfaz entregue una alimentación Phantom 48v al micrófono.

Computador



- Que cuente con diferentes opciones para almacenar información (CD, disco duro, USB, Pendrive, etc).
- Que cuente con puertos USB 2.0
- No se recomienda grabar audio directamente con la tarjeta de sonido del computador.

Almacenaje y formato de archivos



- Se recomienda utilizar formato wav sin compresión.

A continuación, describiremos las principales medidas utilizadas por los diferentes programas de análisis acústico existentes en el mercado y las más utilizadas para fines de investigación, evaluación y rehabilitación vocal.

Antes de describir cada una de ellas, es importante saber que no todas las voces son candidatas a ser analizadas mediante análisis acústico. En voces muy caóticas los valores cuantitativos pierden validez. En ese sentido la National Center for Voice and Speech (Titze, 1995) recomienda que la evaluación debe comenzar con un análisis espectrográfico para clasificarla con el fin de determinar los métodos más idóneos de evaluación y en base a eso podemos determinar qué tipo de medidas utilizar. En relación al espectrograma, la voz se puede clasificar en tres niveles:

VOZ TIPO 1: Normal y patológica con vibraciones casi periódicas (leve disfonía). Debería ser analizada por medio de los parámetros de perturbación a corto plazo (jitter, shimmer, HNR).

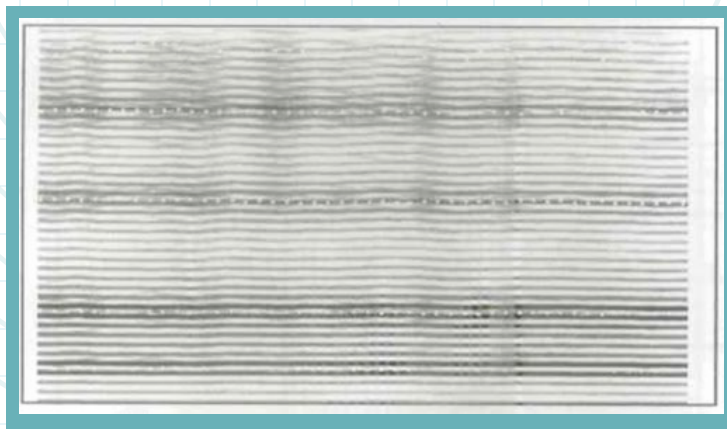


Imagen 4. Voz tipo 1

VOZ TIPO 2: Voces patológicas con subarmónicos y modulaciones. Perturbación mayor al 5%. No puede ser analizada de forma fiable con los parámetros de perturbación a corto plazo. Se estudiará mediante algún método perceptual como la GRBAS y por medio de un método visual como el espectrograma.

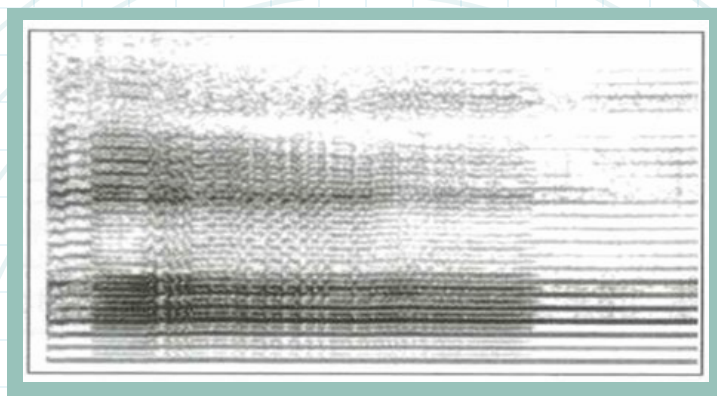


Imagen 5. Voz tipo 2

VOZ TIPO 3: Voces caóticas o random. Solo pueden ser estudiadas mediante métodos perceptuales GRBAS.

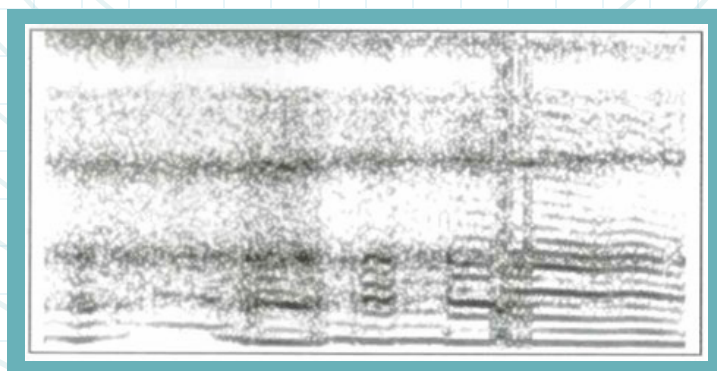


Imagen 6. Voz tipo 3

Principales medidas utilizadas

JITTER ABSOLUTO (JITTA):

Evalúa la variabilidad de la F0 ciclo a ciclo y es muy sensible a las variaciones de F0 que ocurren entre periodos consecutivos.

VFO(COEFFICIENT OF FUNDAMENTAL FREQUENCY VARIATION):

Corresponde a un tipo de medida de perturbación de la frecuencia que también es obtenido a través de un factor suavizante, promediando un número determinado de ciclos de acuerdo a las necesidades del evaluador.

JITTER ABSOLUTO (JITT):

Evalúa la variabilidad relativa de la F0 ciclo a ciclo y es medida en porcentaje.

1.-Análisis de la frecuencia fundamental y su perturbación.

SPPQ(SMOOTHED PITCH PERIOD PERTURBATION QUOTIENT):

Corresponde a un tipo de medida de perturbación de la frecuencia que también es obtenido a través de un factor suavizante, promediando un número determinado de ciclos de acuerdo a las necesidades del evaluador.

RAP (RELATIVE AVERAGE PERTURBATION):

Corresponde a un tipo de medida de perturbación de la frecuencia obtenido a través del promedio de 3 periodos consecutivos. Esto elimina la influencia de las variaciones a largo plazo sobre las de corto plazo, mitigando o suavizando el valor del Jitter relativo.

PPQ (PITCH PERIOD PERTURBATION QUOTIENT):

Corresponde a un tipo de medida de perturbación de la frecuencia obtenido a través del promedio de 5 periodos consecutivos.

SHIMMER ABSOLUTO(SHDB):

Evaluación de dB de la variabilidad peak to peak de la amplitud ciclo a ciclo, muy a corto plazo.

Vam (Coefficient of Amplitude Variation):

Corresponde a la desviación estándar relativa (%) de la amplitud peak to peak. Refleja tanto las variaciones a corto plazo, como las de largo plazo. Se obtiene a través del cociente entre la desviación estándar y el promedio de la variación de la amplitud ciclo a ciclo.

Shimmer relativo (Shim):

Evalúa la variabilidad peak to peak relativa (%) de la amplitud ciclo a ciclo.

2.-Análisis de la amplitud y sus medidas de perturbación

Sapq (Smoothed Amplitude Perturbation Quotient):

Corresponde a un tipo de medida de perturbación e la amplitud que también es obtenido a través de un factor suavizante, promediando un número determinado de ciclos de acuerdo a las necesidades del evaluador.

APQ (Amplitude Perturbation Quotient):

Evaluación relativa (%) de la variabilidad ciclo a ciclo de la amplitud (peak to peak) a través de un factor suavizante de 11 periodos reduce la sensibilidad del APQ frente a los errores de extracción o variaciones a corto plazo de la F0.

HNR (Harmonic to Noise Ratio):

Mide la relación entre el componente armónico de la voz que se ubica en el numerador y el ruido que se ubica en el denominador. Un HNR mayor indica mayor componente armónico en la señal, mientras que un HNR menor indica una señal con poco componente armónico.

SPI (Soft Phonation Index):

Corresponde a la relación promedio entre la energía armónica de baja frecuencia (70-1550 Hz) y la energía armónica de alta frecuencia (1600-4500 Hz). No es un indicador de ruido propiamente tal, sino que se relaciona con la pendiente espectral que es la relación entre los armónicos de baja frecuencia y alta frecuencia.

Una pendiente más abrupta se relaciona con una pérdida de energía mayor de los armónicos altos y se asocia a voces opacas y deficiencia de cierre cordal, mientras que una pendiente menor o más plana da una voz más metálica y se asocia a hiperaducción cordal.

3.-Índices de ruido glótico

NHR (Noise to Harmonic Ratio):

Es la relación inversa del HNR. En este caso, el indicador establece una relación entre el promedio de la energía espectral inarmónica o interarmónica (ruido) y la energía espectral armónica. Un NHR mayor se asocia a mayor presencia de ruido en la voz.

VTI (Voice Turbulence Index):

Corresponde a la relación entre la energía inarmónica de alta frecuencia (1800-5800 Hz) y la energía armónica en el rango frecuencial de 70 - 4200 Hz.



4.-Análisis Espectral

En el análisis espectral está el espectrograma y el espectro. Acá es importante aclarar que son cosas diferentes. El espectrograma es un gráfico tridimensional del sonido complejo que representa en el eje de abscisa (x) el tiempo, en el eje de la ordenada (y) la frecuencia y el tercer eje lo constituye la intensidad del color de las líneas que representa la amplitud. El espectro es un corte en un punto o sección del espectrograma, es un zoom, es una representación frecuencial del sonido en un determinado punto de su duración.

El análisis espectral es útil para la evaluación de la interacción entre la fuente (pliegues vocales) y el filtro (tracto vocal) (Stemple, Glaze y Klaben, 2010).

Dependiendo de la configuración que se realice en el software se pueden obtener dos tipos de espectrogramas dependiendo si queremos mejor resolución frecuencial o mejor resolución temporal.

ESPECTROGRAMA BANDA ESTRECHA:

Muestra muy bien las frecuencias (armónicos), pero mal los pulsos glóticos y deficientemente los formantes (mejor resolución del dominio de la frecuencia) (Casado y Pérez, 2009).

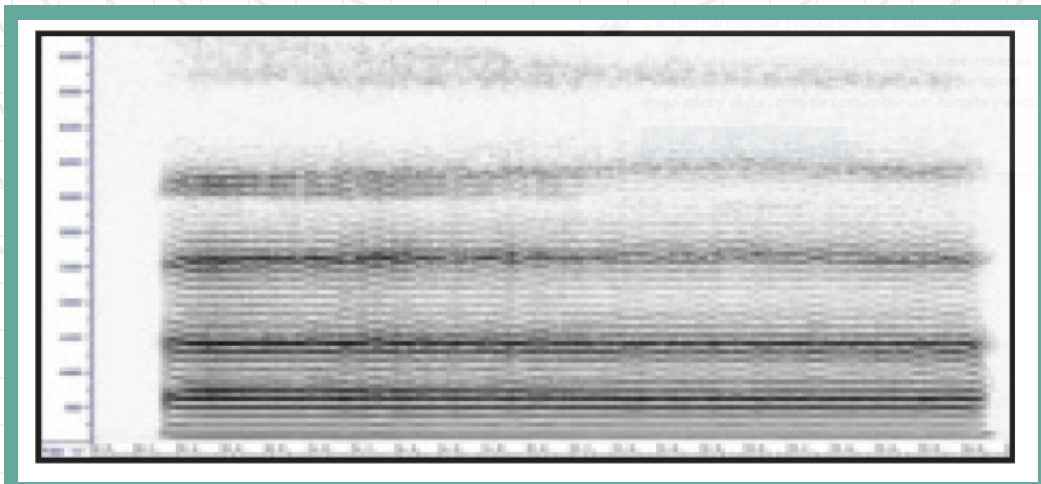


Imagen 7. Espectrograma de banda estrecha.

ESPECTROGRAMA DE BANDA ANCHA:

Muestra muy bien los pulsos glóticos y los formantes (como bandas oscuras horizontales de concentración de energía), pero mal las frecuencias (menor resolución del dominio temporal) (Casado y Pérez, 2009).

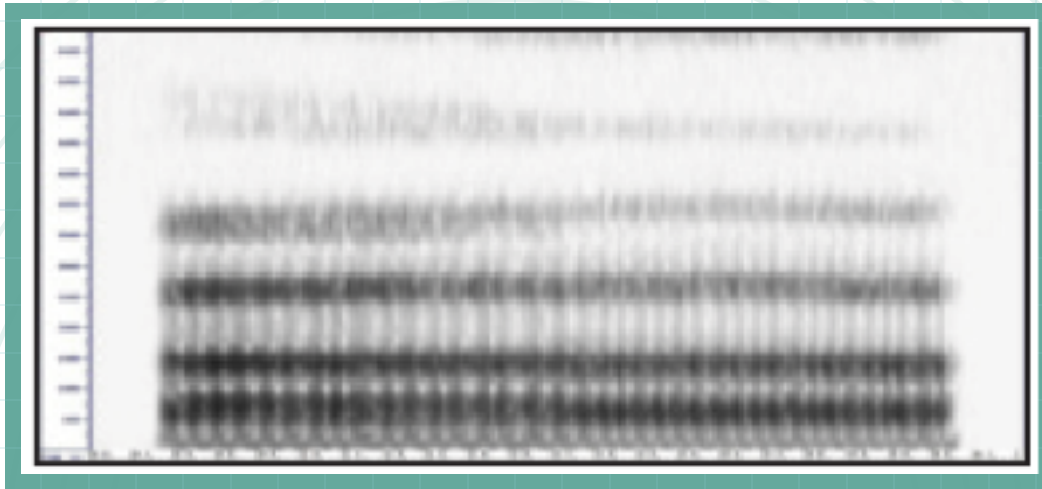


Imagen 8. Espectrograma de banda ancha

En el análisis de voces muy disfónicas o caóticas resulta de mucha utilidad la espectrografía, la cual tiene ciertas ventajas respecto del análisis acústico (Casado y Pérez, 2009):

- Crea un gráfico en el que se representan los armónicos y formantes en el espectro y en el que el ruido añadido por la patología vocal se observará de forma nítida, sin que tengamos que ayudarnos de bases de datos para su interpretación.
- Es absolutamente comparable entre los distintos programas disponibles. Se considera de mucha mayor utilidad para la evaluación de los resultados terapéuticos el empleo del espectrograma de banda estrecha que la clasificación de la voz hecha por el programa basándose en su base de datos de voces normales y patológicas.
- Es útil para el estudio de las voces sin periodicidad o creadas por patrones caóticos de vibración de las cuerdas vocales. Por muy patológica que sea la voz siempre se dispondrá de un gráfico en el que veremos representado el ruido, los armónicos, los subarmónicos y los segmentos vacíos de señal.
- Puede ser el único método para estudiar y objetivar las voces con importantes alteraciones de la estabilidad de la intensidad y el tono a largo plazo o lo que es lo mismo, el temblor vocal ya que el análisis acústico de las perturbaciones a corto plazo no documenta los fenómenos que caracterizan estas alteraciones.

VOCES TIPO I:

Existe una mezcla sutil entre los componentes armónicos de las vocales y componentes de ruido principalmente en la zona e los formantes de las vocales (F1 y F2)

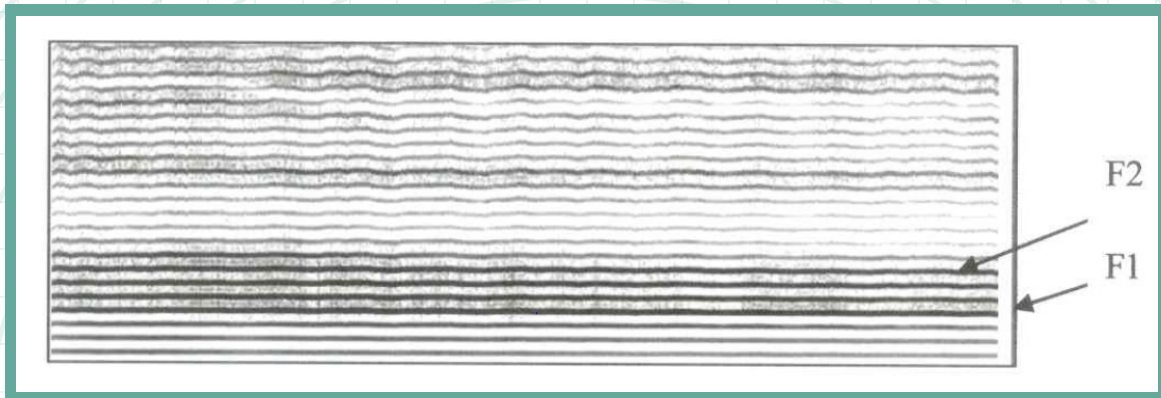


Imagen 9. Grado I de Ronquera.

VOCES TIPO II:

Los componentes de ruido en el segundo formante predominan sobre los componentes armónicos y aparecen componentes de ruido adicional en altas frecuencias alrededor de 3000 hz.

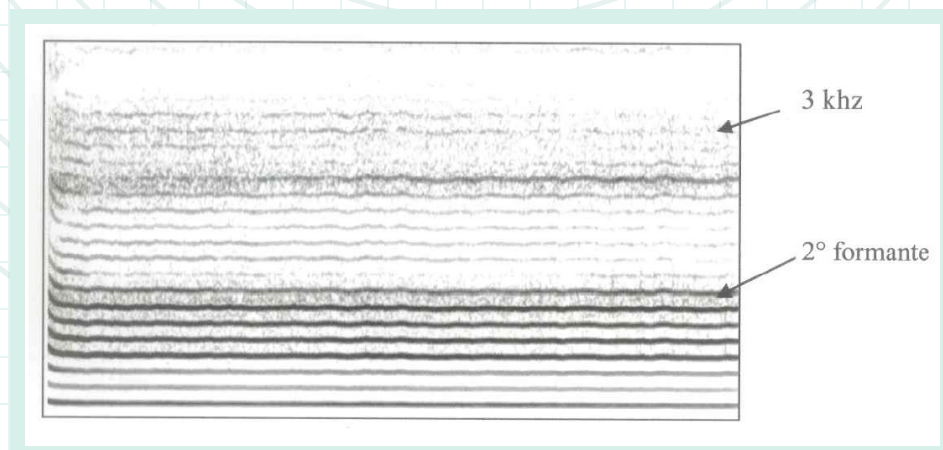


Imagen 10. Grado II de Ronquera.

VOCES TIPO III:

El segundo formante es totalmente reemplazado por componentes de ruido y se incrementan los componentes de ruido adicional alrededor de 3000 hz.

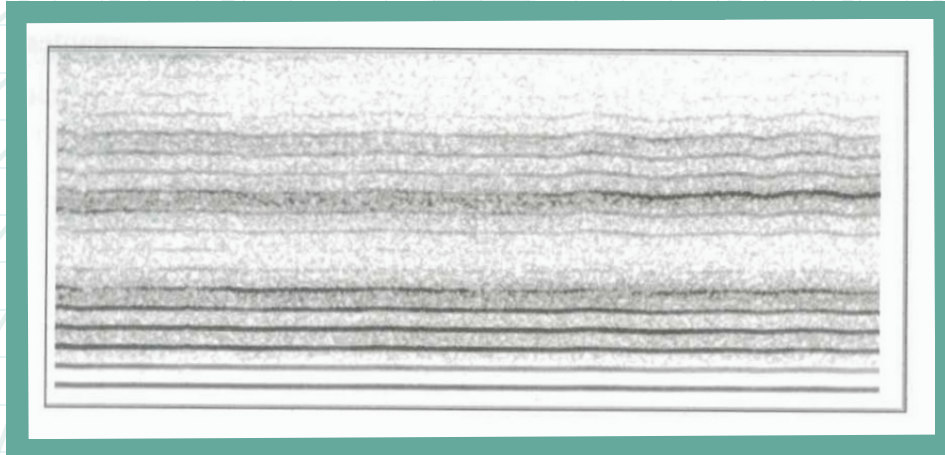


Imagen 11. Grado III de Ronquera.

VOCES TIPO IV:

El segundo formante es reemplazado por componentes de ruido, e incluso el primer formante de todas las vocales a menudo pierde sus componentes armónicos siendo reemplazado por componenteVde ruido. Se suman componentes de ruido aún mayores en altas frecuencias.

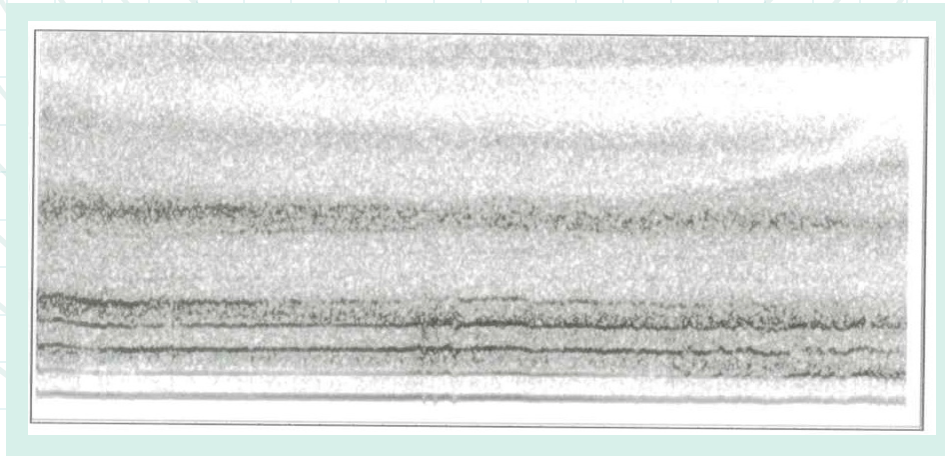


Imagen 12. Grado IV de Ronquera.

ANÁLISIS ESPECTRAL : ESPECTRO FFT (Power Spectrum)

El power spectrum se logra gracias a la aplicación de la transformada rápida de Fourier o FFT (Fourier Fast Transformation). Un espectro como el FFT, permite observar todas las frecuencias armónicas que componen la voz en un punto determinado de la señal, es decir, es un espectro a corto plazo (McCoy, 2012).

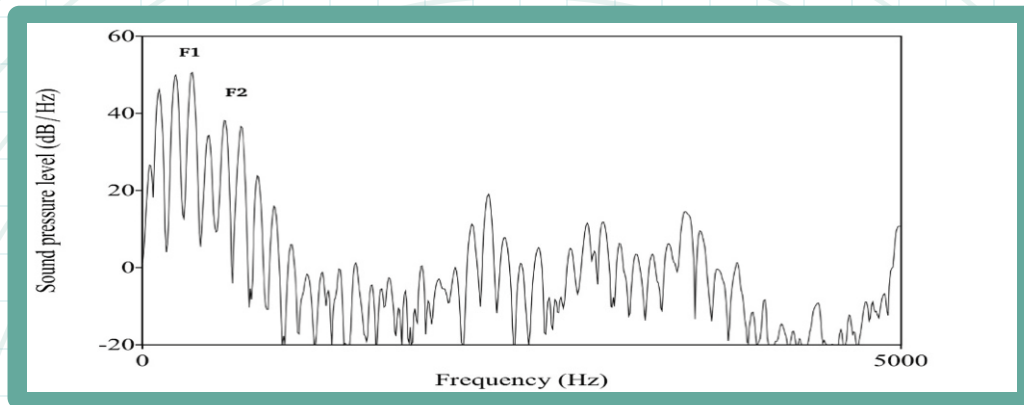


Imagen 13. Espectro FFT

ANÁLISIS ESPECTRAL : ESPECTRO LPC (Código de predicción lineal)

Este método es ampliamente utilizado debido a que (Rabiner y cols, 1993):

- * Proporciona un buen modelo de la señal de habla, ya que sus parámetros se ajustan a la envolvente espectral del tracto vocal y a las características más bien estacionarias de los sonidos sonoros del habla.
- * Los parámetros obtenidos del LPC entregan un espectro suavizado, por lo que la información es más representativa de la voz en el habla.
- * El LPC corresponde a un método preciso y muy adecuado para computación de datos.
- * Existen muy buenos resultados en las pruebas realizadas con LPC en los distintos campos relacionados con el tratamiento automático de la información de la señal de la voz y el habla.

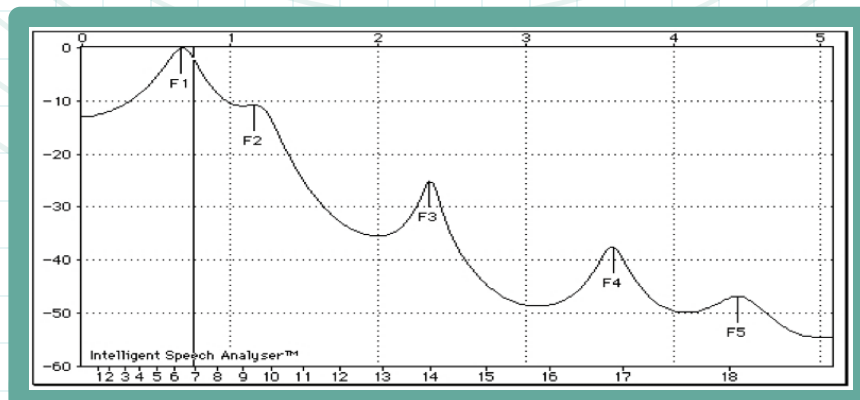


Imagen 14. Espectro LPC

Análisis Espectral: Cepstrum

El cepstrum corresponde a una segunda aplicación de la FFT sobre el espectro que resultó de la primera FFT, es decir corresponde al espectro de un espectro (Saito y cols, 2008).

- * Debe considerarse el peak cepstral (CPP) que corresponde a los valores de Q más altos.
- * El CCP representa la magnitud de la energía armónica presente en la señal de la voz en un punto específico, elevándose por sobre la señal de ruido.
- * La amplitud del CCP indicara mayor o menor estabilidad de la F0. Por tanto, mientras más nítido, delgado y alto sea el Peak, mas fuerte será la energía armónica y más estable la señal.
- * Por último, debido a que el cepstrum no se basa en algoritmos de perturbación de la señal, es posible utilizarlo para el análisis de señales tipo II y III según la National Center for Voice and Speech.

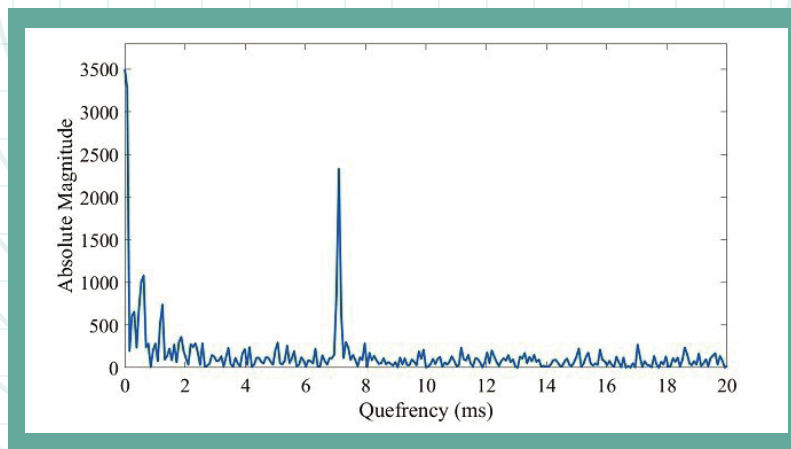
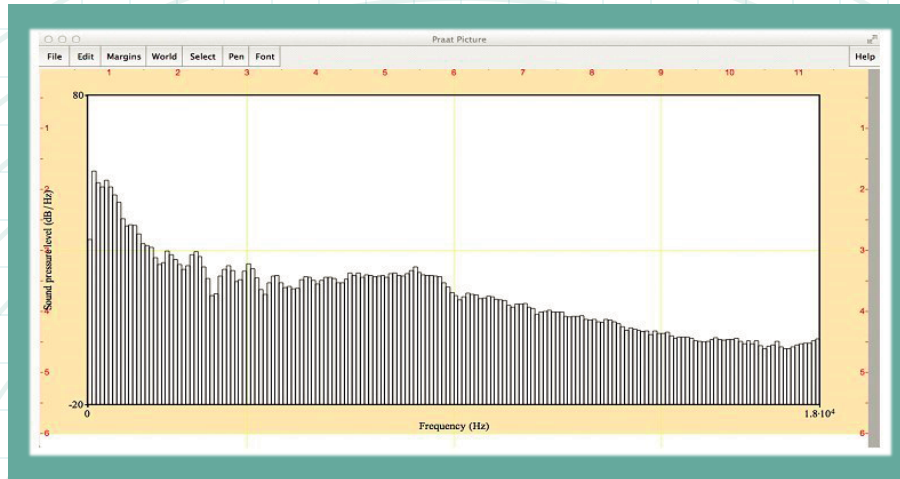


Imagen 16. Grafico del Cepstrum.
Es posible observar el peak denominado CCP alrededor de los 7 ms.

Espectro promedio a largo plazo: Long-Term-Average Spectrum (LTAS)

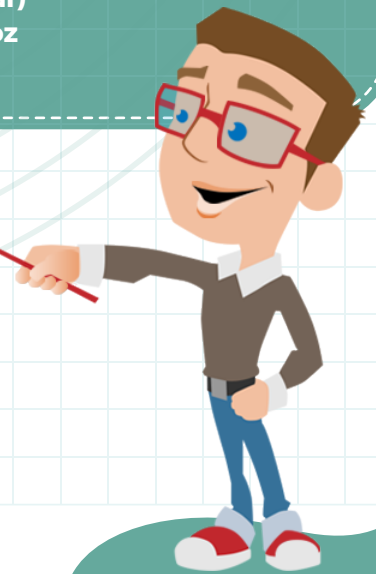


El LTAS corresponde a un promedio de muchos espectros obtenidos a través de la FFT aplicados a una muestra de habla conectada (lectura, frases, canto, etc), donde el eje “x” corresponde a la frecuencia medida en hz y el eje “y” a la amplitud o energía medida en decibeles (Saldías, 2017).

En otras palabras, nos otorga la posibilidad de observar la distribución de energía promedio de los componentes frecuenciales de un sonido proveniente de una muestra de habla conectada (Nordember y Sundberg, 2003).

Por otro lado, al no obedecer a los algoritmos aplicados en el análisis de perturbación, la ventaja del LTAS radica en que no requiere de una señal vocal periódica o casiperiódica para un análisis confiable, es decir, es posible evaluar voces de señales tipo 1, 2 o 3 (Tanner y cols, 2005).

El grafico LTAS nos permitirá observar la pendiente espectral (dependiendo de los rangos frecuenciales que se desee analizar) y así interpretar la cualidad vocal y a partir de esto, si es una voz hipofuncional o hiperfuncional (Tanner y cols, 2005).



Para poder interpretar bien el LTAS es importante consiedrar un concepto denominado “modos de fonación”. En ese sentido Sundberg (1994) describe que la fuente de la señal de la voz varia constantemente en relación a 3 dimensiones: la frecuencia, la amplitud y los modos de fonación. Sunberg describe que los modos de fonación se relacionan íntimamente con la resistencia glótica, definida como el cociente entre la presión subglótica y flujo glótico. En ese sentido describe 4 modos fonatorios:

- **Breathy o soplada:** Se caracteriza por una baja presión subglótica con alto flujo transglótico. Alto esfuerzo compensatorio y baja sonoridad. Por tanto, acá la aducción cordal esta disminuida. En el espectro se observa una caída abrupta de la amplitud en las frecuencias altas. El resultado es una voz opaca.

- **Pressed u oprimida:** Se caracteriza por una alta presión subglótica con bajo flujo transglótico. Asociada a un alto esfuerzo la sonoridad es menor que la voz “Flow o económica” pero mayor que la soplada. En el espectro los agudos conservan energía en las frecuencias altas y su resultado es una voz brillante pero estridente.

- **Fonación neutral:** Se caracteriza por una baja presión subglótica y bajo flujo transglótico. Se asocia a un bajo esfuerzo y sonoridad media.

- **Flow:** Se caracteriza por alta presión subglótica (menor que en pressed) con alto flujo transglótico (menor que en breathy). Se asocia a una voz económica (mayor sonoridad con mínimo esfuerzo).



Sundberg (1994) agrega que la fonación pressed corresponde a un modo de fonación hiperfuncional, asociado a una posición vertical alta de la laringe, mientras que la fonación flow tiende a producirse con una laringe más baja, lo que permitiría un contacto cordal óptimo. Por otro lado, la fonación breathy contaría con un cierre glótico breve e incompleto.

Verdolini (1998) hace la diferencia entre dos tipos de fonación que permitirían un desempeño óptimo, la voz flow que corresponde al tipo de fonación que presenta el mayor flujo transglótico, combinado con un cierre glótico óptimo y la voz resonante que es aquella que implica una configuración glótica óptima y se asocia a sensaciones vibratorias faciales anteriores.

Las medidas que se pueden obtener con el LTAS dependiendo de las bandas frecuenciales en las cuales quiero establecer diferencias de energía son las siguientes:

Alpha Ratio (Relación Alfa): Permite establecer la diferencia de energía existente entre las bandas frecuenciales 50-1000 Hz y 1000-5000 Hz, lo que entrega información de la pendiente espectral. Una pendiente espectral mas abrupta se correlaciona con una pérdida abrupta energía de los armónicos altos y mayor energía de los armónicos de las frecuencias bajas, lo que se asocia a voces de tipo hipofuncionales u opacas. (Guzman y cols, 2013). Por otro lado, un valor de alpha ratio muy negativo también se asocia a voces hipofuncionales. Al contrario pendientes menos abruptas con un alpha ratio mayor o muy positivo cercano a 0 se asocia a voces hiperfuncionales y estridentes.

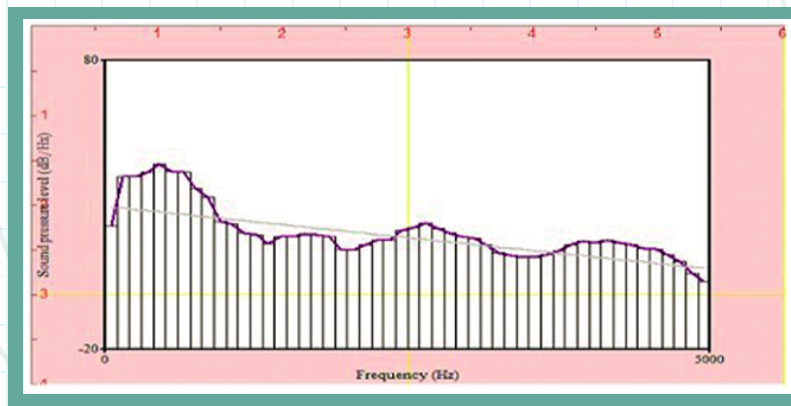


Imagen 18. Espectro LTAS obtenido mediante praat, con la pendiente espectral marcada, la cual se puede observar aun presenta bastante energía en los armónicos altos, presentando una pendiente menos abrupta.

L1-L0: Corresponde a la diferencia de energía entre las bandas frecuenciales 300-800 Hz y 50-300 Hz. Establece las diferencias de energía entre las regiones del formante 1 y la frecuencia fundamental. Esto nos permite obtener información acerca del modo fonatorio. Una diferencia con valor negativo sería indicador de voces sopladas y viceversa (Leino, 2009).

A continuación, se describen otros procedimientos de evaluación de la voz muy utilizados actualmente en la clínica e investigación de la voz como son la electroglotografía y el análisis aerodinámico de la voz.

Electroglotografía:

Es un método incruento, económico, rápido y fácil de manejar que mide las variaciones de impedancia que se producen al paso de una débil corriente alterna a través de los tejidos del cuello mediante la colocación de dos electrodos de contacto situados a cada lado del cartílago tiroides, a la altura de las cuerdas vocales.

La señal obtenida se denomina electroglotograma u onda Lx y es proporcional a la superficie de contacto entre ambas cuerdas vocales, de forma que la cima de la curva corresponde al momento de máximo contacto entre las cuerdas vocales. (Casado, J & cols 2002).

La onda electroglotográfica se puede descomponer en distintos segmentos y puntos según la etapa del ciclo vibratorio, pero por lo general los puntos y segmentos son interpretados de acuerdo al diagrama adaptado de Childers, Moore, Hicks y Alsaka en 1986.

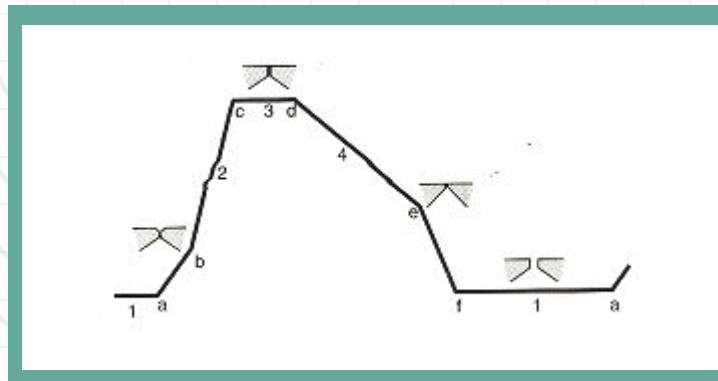
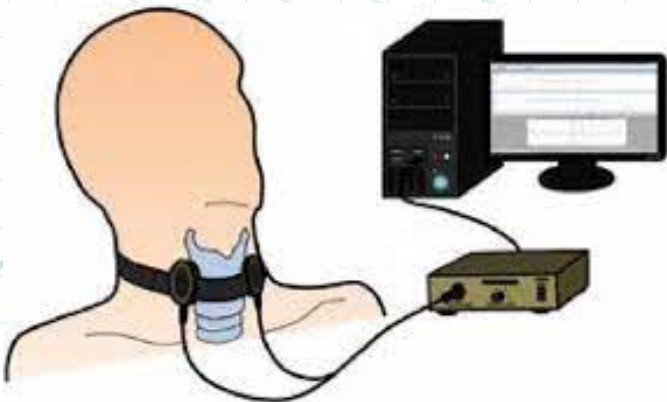


Imagen 20. Puntos y segmentos de la forma de onda de EGG (adaptado de Childers, Moore, Hicks y Alsaka. 1986).



Se pueden realizar diferentes análisis en la electroglotografía tanto a nivel de los valores objetivos que entrega, así como también un análisis cualitativo de la onda electroglotográfica. Dentro de los valores objetivos que entrega el electroglotógrafo encontramos el Cociente de cierre (CQ) el cual es el porcentaje de duración dentro del ciclo fonatorio de la fase de contacto de las cuerdas vocales. A su vez el cociente de apertura (OQ) es la porción de la etapa de apertura de la vibración de las cuerdas vocales en un ciclo glotal completo. Por otro lado, el cociente de velocidad (SQ), es la proporción de la duración de la fase de apertura de las cuerdas vocales hasta la duración de la fase de cierre. Dentro de la literatura se pueden encontrar rangos normativos.

Tipo de fonación	Cociente de cierre
Normal	0,4 a 0,6
Tensa	0,7 a 1
Soplada	0,1 a 0,3

Imagen 21. Tabla 1: valores normativos de cociente de cierre de electrografía según el manual asesor de la voz.

No obstante, la evaluación de la onda electroglotográfica es esencialmente cualitativa. Algunos aspectos que pueden ser observados son los siguientes (Fourcin, 1981; Kelman, 1981; Mac Curtain y Fourcin, 1982 y Reed, 1982):

- * Uniformidad en la amplitud de la onda es asociado a bajo nivel de perturbación.
- * Inestabilidad en la amplitud, o en el periodo, indica variabilidad anormal en el patrón de contacto de las cuerdas vocales.
- * Pequeñas oscilaciones en la región del peak de cierre indica aleteo (fluttering) o rebote de las cuerdas vocales después del contacto inicial. Alteraciones en los segmentos de apertura o cierre indican anormalidad de tensión de las cuerdas vocales.
- * Una fase de cierre rápida y abruptamente definida indica buena excitación acústica del tracto vocal y eficiente producción vocal.
- * La fase de cierre lleva menos tiempo que la fase de apertura.
- * Duración aumentada de la fase de cierre esta correlacionada con peaks de espectro no amortiguados y amplificadas.
- * La onda EGG y la frecuencia fundamental tienen estrecha relación, por lo tanto son comparables las medidas de jitter F0 y jitter-EGG; pero la amplitud de la onda EGG representa algo diferente que la amplitud de la señal, por lo tanto los valores del shimmer no son comparables.

Titze (1990) identifica cuatro formas de onda EGG características, señalando la correspondencia en su geometría glótica.

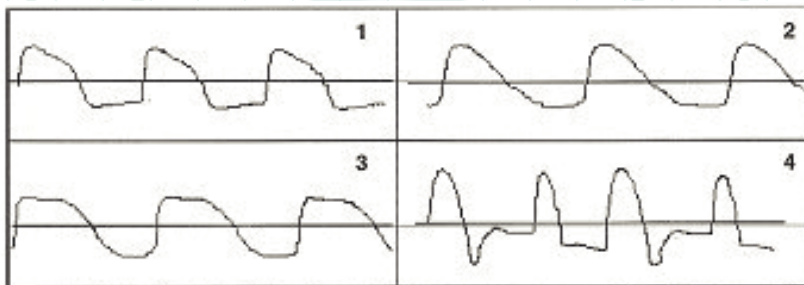


Imagen 22. Formas de onda EGG características, señalando la correspondencia en su geometría glótica. (Titze 1990).

Vaca y Cobeta (2013), describen de forma aproximada algunas formas de onda según la patología vocal específica.

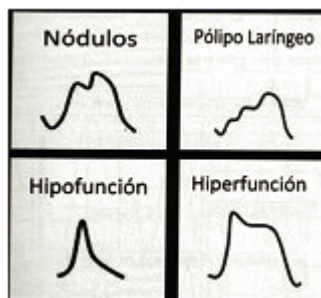


Imagen 23. Formas de onda según patología (Vaca & Cobeta, 2013)

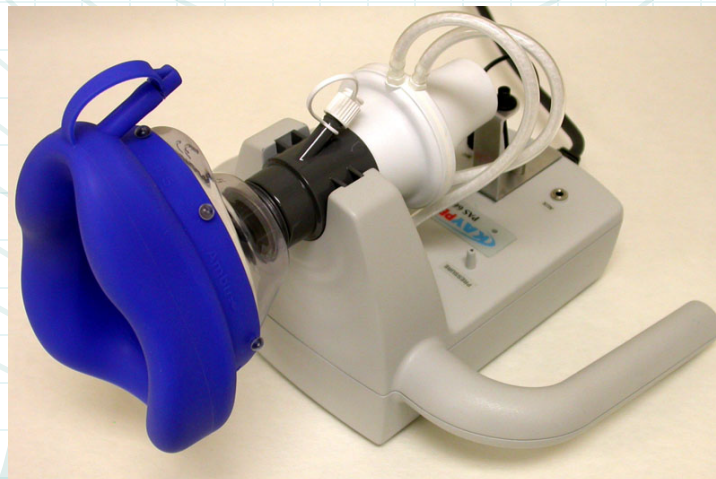
La forma de la onda electroglotográfica, también puede verse influenciada dependiendo del registro vocal utilizado, modo fonatorio e intensidad empleada en la emisión.

Medidas aerodinámicas de la fonación

Las medidas aerodinámicas de la fonación constituyen un método clínico de evaluación objetiva, funcional y no invasiva para obtener información acerca de la función vocal.

Para entender las medidas aerodinámicas de la fonación y cómo interactúan entre ellas, es importante recordar el proceso de producción de la voz a nivel glótico. La voz es producida cuando la presión subglótica de aire, generada por un volumen pulmonar de aire exhalado, excede la resistencia de los pliegues vocales que están aducidos, causando que estos sean activados y así a su vez causar un flujo de aire que los pone en oscilación. (Guzmán, 2017).

La descripción anterior muestra de manera implícita que elementos como el volumen de aire, la presión de aire, flujo de aire y la resistencia a este flujo tienen un rol primordial en la producción de la voz



Las medidas aerodinámicas utilizadas en la evaluación objetiva de la voz son (Guzmán, 2017):

Volumen de flujo:

Es la cantidad total de aire utilizada durante una tarea de habla y es medido en litros o mililitros. En adultos, el volumen fonatorio normal es de 1500-4000 ml, dependiendo del género y talla del paciente (Stemple, 2015).

Promedio de velocidad de flujo:

Se define como la velocidad con que el aire pasa entre los pliegues vocales durante la fonación y es medido en litros por segundo o mililitros por segundo. El promedio normal de flujo se encuentra entre 80 y 200 ml/seg en los adultos (Stemple, 2015).

El instrumento utilizado para medir la velocidad de flujo de aire durante la fonación es el neumotacógrafo.

Pacientes con incompetencia glótica pueden presentar valores muy altos de promedio de velocidad de flujo (400-600 ml/sg). Los pacientes con hiperfunción glótica severa o fonación vocal fry pueden presentar promedios de flujo muy bajos (10-15 ml/seg) (Guzmán, 2017).

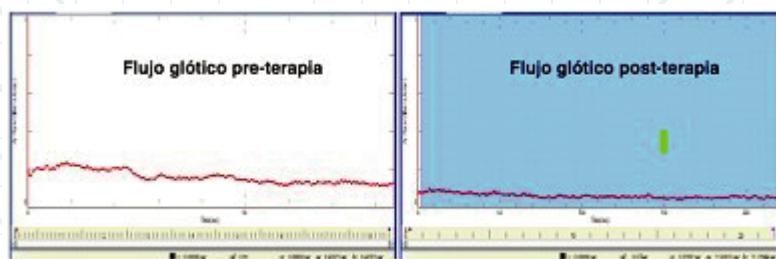


Imagen 23. Gráficos representativos de la disminución de la velocidad de flujo después de la terapia vocal. A la izquierda (pre-terapia) se ve un mayor flujo que en el gráfico de la derecha (post-terapia). Extraída de Tópicos iberoamericanos en voz cantada y hablada. Volumen I-1a ed

Presión subglótica:

Desde el punto de vista físico, la presión es definida como la fuerza por unidad de área, actuando perpendicularmente en esta área. En el caso de la fonación, la presión subglótica actúa como una fuerza debajo de los pliegues vocales, subiendo hasta que supera la resistencia que éstos oponen y dando comienzo así a la oscilación (Titze, 2010).

La presión subglótica es medida en cm de H₂O. La presión subglótica necesaria para una fonación a intensidad conversacional es entre 5 y 10 cm de H₂O en voces normales (Stemple, 2015).

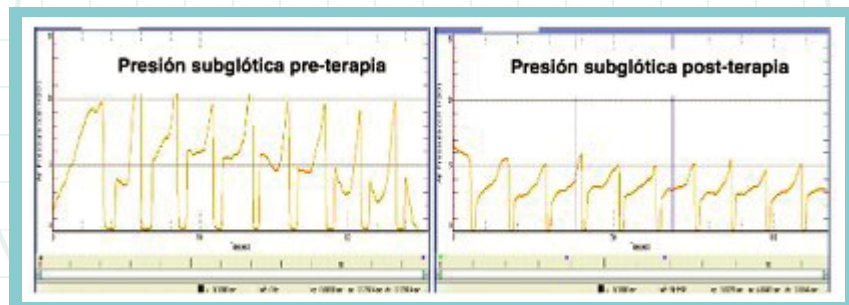
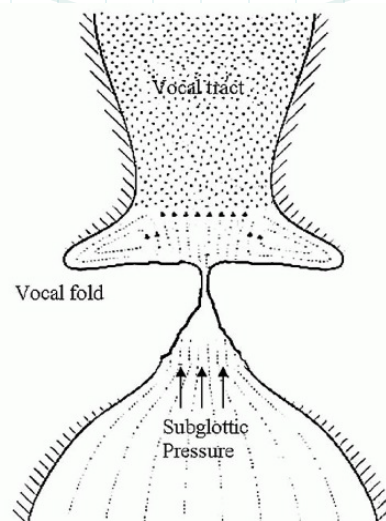
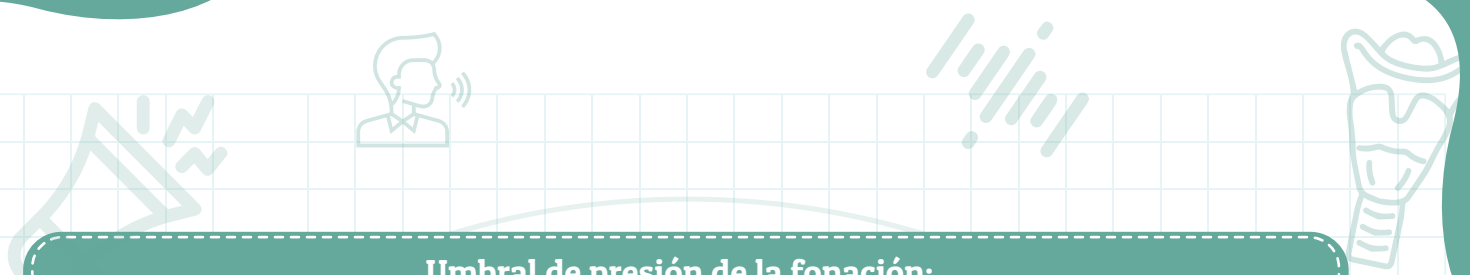


Imagen 24. Gráficos representativos de la disminución de la presión subglótica después de la terapia vocal. A la izquierda (pre-terapia) se ve una mayor presión subglótica que en el gráfico de la derecha (post-terapia). Extraída de Tópicos iberoamericanos en voz cantada y hablada. Volumen I-1a ed.





Umbral de presión de la fonación:

Este parámetro es definido como la presión subglótica mínima requerida para que los pliegues vocales inicien o mantengan la fonación. (Titze, 1992).

En sujetos con voces normales se espera que el umbral se encuentre entre 3 y 5 cm de H₂O. (Titze, 1992). En voces que refieren un alto grado de esfuerzo fonatorio, presenta normalmente un alto valor de umbral de presión de la fonación. Es decir para iniciar o mantener la fonación necesita una mayor presión subglótica.

El umbral de presión de la fonación depende de las características viscoelásticas de los pliegues vocales, de la configuración glótica, grosor del bode libre de los pliegues vocales, cantidad de amortiguación de los tejidos y la velocidad de la onda mucosa (Guzmán, 2017).

Otros elementos que afectan el umbral de presión de la fonación son la hidratación de los pliegues vocales, la frecuencia fundamental y la presencia o no de fatiga vocal. (Verdolini, titze, Druker, 1990).

- * Presión intraoral: Es la presión de aire que se encuentra en la cavidad oral y faringea.**
- * Presión intraglotica: Es la presión de aire entre los pliegues vocales.**
- * Presion transglótica: se define como la diferencia entre la diferencia entre la presión subglótica y la**
- * Presión supraglótica. La presión transglótica es fundamental en la producción de la voz (Mehta y Hillman, 2008).**

Resistencia glótica: Es la resistencia que oponen los pliegues vocales al paso del flujo desde la subglotis hacia la supraglotis. La resistencia glótica no puede ser medida directamente, sino que se calcula dividiendo la presión subglótica por el promedio de velocidad de flujo (Stemple, 2014).

La resistencia glótica depende de: Grado de aducción de los pliegues vocales, rigidez y elasticidad de los pliegues vocales, Frecuencia fundamental, registro vocal utilizado, intensidad.





PARTE II: DATOS NORMATIVOS

Introducción de la sección

¿Qué son los datos normativos y para que nos sirven?. Tomemos el ejemplo de cuando debemos hacernos un examen de sangre para medir ciertos biomarcadores como los del colesterol. ¿Como saber si estos están dentro de los rangos normales?; El resultado generalmente es sólo un valor numérico, pero ese valor nos permite comparar el resultado o rendimiento de determinada persona con los datos normativos que indican normalidad o enfermedad.

En voz sucede lo mismo con algunos parámetros. Ejemplo: en el caso de la frecuencia fundamental (F_0), de las mujeres, se sabe que los valores normativos se deben encontrar entre cierto rango de frecuencias que son diferentes al rango de frecuencias de hombres y de infantes.

En este apartado se pretende mostrar la data normativa existente y ciertas clasificaciones que se hacen en cuanto a valores normativos de la voz con el fin de orientar al estudiante a desarrollar los siguientes resultados de aprendizaje:

- * Identificar valores de normalidad de parámetros vocales en función de datos normativos de la literatura.
- * Distinguir voz normal de una patológica a través de la interpretación de los parámetros vocales.
- * Valorar el uso de datos normativos en el proceso de evaluación clínica de la voz.



1. Notas Musicales y sus equivalentes frecuencias en Hertz.

Do 0 33	Do# 0 35	Re 0 37	Re# 0 39	Mi 0 41	Fa 0 44	Fa# 0 46	Sol 0 49	Sol# 0 52	La 0 55	La# 0 58	Si 0 62
Do 1 65	Do# 1 69	Re 1 73	Re# 1 78	Mi 1 82	Fa 1 87	Fa# 1 92	Sol 1 98	Sol# 1 104	La 1 110	La# 1 117	Si 1 123
Do 2 131	Do# 2 139	Re 2 147	Re# 2 156	Mi 2 165	Fa 2 175	Fa# 2 185	Sol 2 196	Sol# 2 208	La 2 220	La# 2 233	Si 2 247
Do 3 262	Do# 3 277	Re 3 294	Re# 3 311	Mi 3 330	Fa 3 349	Fa# 3 370	Sol 3 392	Sol# 3 415	La 3 440	La# 3 466	Si 3 495
Do 4 523	Do# 4 554	Re 4 587	Re# 4 622	Mi 4 659	Fa 4 698	Fa# 4 740	Sol 4 784	Sol# 4 831	La 4 880	La# 4 932	Si 4 988
Do 5 1047	Do# 5 1109	Re 5 1175	Re# 5 1245	Mi 5 1319	Fa 5 1397	Fa# 5 1480	Sol 5 1568	Sol# 5 1661	La 5 1760	La# 5 1865	Si 5 1976
Do 6 2093	Do# 6 2218	Re 6 2349	Re# 6 2489	Mi 6 2637	Fa 6 2794	Fa# 6 2960	Sol 6 3136	Sol# 6 3322	La 6 3520	La# 6 3729	Si 6 3951
Do 7 4186	Do# 7 4435	Re 7 4699	Re# 7 4978	Mi 7 5274	Fa 7 5588	Fa# 7 5920	Sol 7 6272	Sol# 7 6645	La 7 7040	La# 7 7459	Si 7 7902

Notas musicales y sus equivalentes frecuencias en Hertz, basada en Escala temperada La 3 = 440 Hz. (Jackson Menaldi, 2005).

2. Clasificación de las voces según extensión fisiológica

Hombre Mujeres	F0 (Hz)	Extensión (Hz)	Tesitura (Hz)
Bajo	98-110	65-349 (Do 1 - Fa 3)	110-196 (La 1 - Sol 2)
Barítono	117-133	83-440 (Mi1 - La3)	147-262 (Re 2 - Do 3)
Tenor	147-165	98-523 (Sol 1 - Do 4)	173-330 (Fa 2 - Mi 3)
Contralto	196-226	131-784 (Do 2 - Sol 4)	294-523 (Do 3 - Do 4)
Mezzo	210-226	165-880 (Mi 2 - La 4)	294-523 (Mi 3 - Do 4)
Soprano	244-262	196-1175 (Sol 2 - Mi 5)	392-698 (Sol 3 - Fa 4)

Clasificación de voces según extensión fisiológica (Jackson Menaldi, 2002).

3. Frecuencia Fundamental

Tabla	Frecuencias fundamentales			
Rango etario	F0 promedio	DS	Rango	Referencia
Hombres leyendo				
7	294	2.2		1
8	297	2		1
10	270	2.4		1
11	227	1.5	192-268	2
14	242	3.4		1
19	117	2.1	85-155	3
Adulto	132	3.3		4
20-29	120			5
30-39	112			5
40-49	107			5
50-59	118			5
60-69	112			5
70-79	132			5
80-89	146			5
Mujeres leyendo				
7	281	2		6
8	288	2.8		6
11	238	1.51	98-271	2
19	217	1.71	65-255	3
20-29	224	3.8	192-275	7
30-40	196	2.5	171-222	8
40-50	189	2.8	168-208	8
60-69	200	4.3	143-235	7
70+	202	4.7	170-249	7
80-94	200	2.7	183-225	9

Frecuencias fundamentales (Colton, Casper, & Leonard, 2011).

Tabla	Media de la frecuencia fundamental en distintas edades					
Grupos de edad	20-35		40-55		65-85	
	Mean	DE	Mean	DE	Mean	DE
Mujeres						
No entrenada	192.00	2.00	195.00	3.40	175.00	2.40
Entrenada	217.00	3.60	205.00	4.10	201.00	5.20
Hombres						
No entrenado	118.00	2.60	100.00	2.60	127.00	3.10
Entrenado	131.00	3.10	126.00	3.90	125.00	4.00

Frecuencias fundamentales (Colton, Casper, & Leonard, 2011).

4. Tono Medio Hablado

Niños y jóvenes				
Edad	F0	Nota Musical	Rango	Nota Musical
1 y 2	445	La3	370-525	Fa#3-Do4
3	400	Sol3	340-460	Fa3-La#3
4	375	Fa#3	320-425	Re#3-Sol#3
5	350	Fa3	300-390	Re3-Sol3
6	325	Mi3	280-365	Do#3-Fa#3
7	295	Re3	260-330	Do3-Mi3
8	295	Re3	260-330	Do3-Mi3
9	260	Do3	220-300	La2-Re3
10	235	La#2	195-275	Sol2-Do#3
11	225	La2	185-260	Fa#2-Do3
12	210	Sol#2	170-245	Fa2-Si2
13	195	Sol2	155-230	Re#2-La#2
14	190	Fa#2	150-220	Re#2-La2
15	165	Mi2	130-195	Do2-Sol2
16	150	Re2	120-180	La#1-Fa#2
17	135	Do#2	110-170	La1-Fa2
18	125	Si1	100-155	Sol1-Re#2

Tono Medio Hablado en niños y jóvenes (Wilson,1973).

5. Valores promedio de F0 y formantes para las vocales del español Rioplatense

	Sexo	i	e	a	o	u
F0	Fem	207	205	205	204	204
	Mas	130	125	127	124	124
F1	Fem	330	454	980	546	382
	Mas	290	430	830	510	335
B1	Fem	70	80	110	97	74
	Mas	63	75	105	83	80
F2	Fem	2765	2500	1553	934	740
	Mas	2295	2120	1350	860	720
B2	Fem	130	156	160	130	150
	Mas	103	106	106	105	112
F3	Fem	3740	3130	2890	2966	2760
	Mas	2915	2628	2450	2480	2380
B3	Fem	178	190	210	185	210
	Mas	174	140	142	156	208
F4	Fem	4366	4150	3930	3854	3380
	Mas	3645	3610	3665	3485	3355
B4	Fem					
	Mas	124	180	197	170	150

Valores promedio de la frecuencia glótica F0 y de los formantes F1, F2, F3 y F4 para los dos grupos de sujetos analizados. Promedios de los anchos de banda B1, B2 y B3 para voces femeninas y B1, B2, B3 y B4 para voces masculinas. (Aronson, Ruffer y cols. 2001)

6. Rango fonacional máximo

TABLA	Rango fonacional máximo			
Rango etario	Frecuencia Baja	Frecuencia Alta	Rango	Referencia
Hombres				
17-26	80	764	39.06	1
18-36	80	675	36.92	2
35-75	80	260	20.4	3
40-65	83	443	28.99	5
68-89	85	394	26.55	4
Mujeres				
18-38	140	1,122	36.03	2
66-93	134	571	25.09	4
35-70	136	803	30.75	5

Extensión Tonal. Presentados en (Colton, Casper, & Leonard, 2011).

7. Perturbación de la Frecuencia Fundamental

TABLA	Datos de perturbación de frecuencia						
Rango etario	Medida	Jitter		Factor de perturbación direccional		Cociente de perturbación de la frecuencia	
		/ee/	/oo/	/ee/	/oo/	/ee/	/oo/
Hombres							
20-29	Media	0.78	0.72	70.73	69.48	0.65	0.57
	DE	0.4	0.36	11.25	14.87	0.3	0.26
40-49	Media	0.99	0.87	74.37	72.49	0.77	0.7
	DE	0.61	0.51	10.86	14.84	0.38	0.32
60-69	Media	0.91	0.87	67.77	69.43	0.77	0.74
	DE	0.63	0.56	15.46	14.55	0.5	0.43
Mujeres							
20-29	Media	0.55	0.56	46.3	47.83	0.56	0.57
	DE	0.41	0.41	19.67	19.68	0.39	0.37
40-49	Media	0.63	0.61	53.06	51.67	0.65	0.61
	DE	0.4	0.42	16.57	19.76	0.41	0.4
60-69	Media	0.66	0.7	49.91	48.88	0.65	0.69
	DE	0.52	0.59	19.1	18.64	0.49	0.56

Datos de perturbación de la Frecuencia. Casper,J. (1983). Presentado en (Colton, Casper, & Leonard, 2011) .

8. Intensidad vocal

TABLA Intensidad vocal en monosílabos y lectura de párrafo						
Tipo de enunciado	Nivel de sonoridad	N	Edad	dB promedio	DS (dB)	Referencia
Mujeres						
/pa/						
	Suave	10	20-30	65.35	1.84	1
	Cómoda	10	20-30	70.44	1.88	1
	Ruidosa	10	20-30	76.75	3.38	1
/æ/						
	Suave	20	18-36	83.30 <i>b</i>	3.20	2
	Cómoda	20	18-36	76.40	4.00	2
	Ruidosa	20	18-36	71.50	4.90	2
Párrafo Arcoiris			20-30	68.15		3
Hombres						
/pa/						
	Suave	10	20-30	70.42	3.19	1
	Cómoda	10	20-30	74.69	3.08	1
	Ruidosa	10	20-30	80.72	3.51	1
/æ/						
	Suave	25	17-30	75.00	2.50	2
	Cómoda	25	17-30	79.50	3.30	2
	Ruidosa	25	17-30	86.00	4.30	2
Párrafo Arcoiris			20-30	70.42		3

Datos de perturbación de la Frecuencia. Casper, J. (1983). Presentado en (Colton, Casper, & Leonard, 2011).

9. Niveles de Intensidad máxima

TABLA	Niveles máximos de intensidad			
Edad	Media	DE	Rango	Referencia <i>a</i>
Hombres				
18-39	106	5.1	92-116	1
45-65	110	7.1	99-129	2
68-89	101	5.9	88-110	1
Mujeres				
18-38	106	3	99-112	1
40-70	101	18.2	93-115	2
66-93	99	4.5	90-104	1

Niveles Máximos de Intensidad presentados en (Colton, Casper, & Leonard, 2011).

10. Presión Sonora

Presión	Decibeles (dB)
5 cm de H2O	20 dB Susurro
10 cm de H2O	30 dB Conversación tranquila
60 cm de H2O	60 dB Canto de salón
100 cm de H2O	70 dB Voces de advertencia
160 cm de H2O	80 dB Pick de intensidad de un orador en una reunión
360 cm de H2O	120 dB Primer tenor al máximo de su potencia

Niveles de presión sonora en diversas exigencias vocales (Husson, 1962).

Intensidad habitual (moderada)	Lectura (F0 habitual de habla)= 70 dB SPL
	Vocal sostenida (F0 habitual de habla)= 75 dB SPL
Intensidad mínima	Vocal sostenida (al 10% del rango de frecuencias, promedio de los valores masculinos y femeninos) = 49 dB SPL
Intensidad máxima	Vocal sostenida (al 90% del rango de frecuencias, promedio de los valores masculinos y femeninos) = 102 dB SPL
Rango de intensidad dinámico	Vocal Sostenida (Promediado a través del rango completo de frecuencias)
	Mujeres = 51 – 104 dB SPL
	Hombres = 47 – 100 dB SPL

Niveles de presión sonora en diversas actividades vocales (Berhman, 2012).

11. Perturbación de la amplitud

Tabla	Datos de perturbación de la amplitud	
Vocal	Media	DE
Hombres		
/ah/	0.47	0.34
/ee/	0.37	0.28
/oo/	0.33	0.31
Mean	0.33	0.31
Mujeres		
/ah/	0.33	0.22
/ee/	0.23	0.08
/oo/	0.19	0.04
Promedio	0.25	0.11

Datos perturbación de la amplitud (Colton, Casper, & Leonard, 2011).

12. Presión subglótica

Presión subglótica a diferentes intensidades en adultos

Volumen	Promedio PS (cm de H2O)	Rango (cm de H2O)
Suave		
Hombres	3.8	2.3 – 5.7
Mujeres	4.1	2.6 – 6.3
Normal		
Hombres	7.6	4.5 – 12.8
Mujeres	7.6	3.8 – 12.6
Fuerte		
Hombres	14.9	7.3 – 24.4
Mujeres	19.8	9.1 – 27.7

Datos normativos de presión subglótica en adultos. Wilson & Leeper, (1992).
Presentados en (Berhman, 2012).

13. Presión sonora a nivel intraoral

Presión de aire basadas en estimación de presión intraoral (cm de H2O)

Rango de edad	N	Condición	Presión	DS
Hombres				
17 - 30	25	Normal	5.91	
	25	Suave	4.79	
	25	Fuerte	8.39	
21 – 30	8	Normal	4.12	1
20 – 31	10	Cómodo	5.81	1.4
69 +	10	Cómodo	7.99	2.73
Presión de aire basadas en estimación de presión intraoral (cm de H2O)				
Rango de edad	N	Condición	Presión	DS
Mujeres				
17 - 30	25	Normal	6.09	
	25	Suave	4.79	
	25	Fuerte	8.46	
21 – 30	8	Normal	4.28	0.88
20 – 31	10	Cómodo	6.51	0.8
69 +	10	Cómodo	6.35	1.71

Niveles de presión de aire (Colton, Casper, & Leonard, 2011).

14. Tiempos máximos de fonación

TABLA		Duración máxima de fonación	
Grupo	Edades	Grupo	Edades
Hombres			
Infante	3-4	Infante	3-4
Niño	5-12	Niño	5-12
Adulto	13-65	Adulto	13-65
Anciano	65+	Anciano	65+
Mujeres			
Infante	3-4	Infante	3-4
Niño	5-12	Niño	5-12
Adulto	13-65	Adulto	13-65
Anciano	65+	Anciano	65+

Tiempo máximo de fonación (Colton, Casper, & Leonard, 2011).

Tabla	Tiempo máximo de fonación	
Edad	Media (segundos)	Desviación típica (segundos)
Hombres		
< 4 años	8.9	2.1
Entre 4 y 12 años	17.1	4.1
Adultos	25.9	7.4
> 65 años	14.7	6.2
Mujeres		
< 4 años	7.5	1.8
Entre 4 y 12 años	14.9	3.8
Adultas	21.3	5.6
> 65 años	13.5	5.7

Valores medios normales de tiempos máximos de fonación según edad y sexo (Casado & Adrián, 2002).

Tabla	Tiempo máximo de fonación	
Edad	Media (segundos)	DS(segundos)
Hombres		
Niños pequeños	8.9	2.1
Niños mayores	17.7	4.1
Adultos	25.9	7.4
Personas mayores	14.7	6.2
Mujeres		
Niñas pequeñas	7.5	1.8
Niñas mayores	14.9	3.8
Adultas	21.3	5.6
Personas mayores	13.5	5.7

Tiempos máximos de fonación según edad y sexo, Kent (1987) presentados en (Tapia & Cobeta, 1996).

15. Índice s/z

TABLA	Proporción S/Z		
Edad	Promedio	Rango	Referencia
Hombres			
5	0.92	0.82-1.08	1
7	0.7	0.52-0.97	1
9	0.92	0.66-1.5	1
Mujeres			
5	0.83	0.50-1-14	1
7	0.78	0.51-1.10	1
9	0.91	0.75-1.26	1
Adultos	0.99	0.41-2.67	2
Ancianos	0.76		3
	0.82		3

Valores promedio índice s/z, (Colton, Casper, & Leonard, 2011).

1. Valores límites de normalidad índice s/a o s/e según autores

16. Valores índice s/a y su implicancia clínica

Condición	Índice	Implicancia clínica
Normal	1.5	Condición normal o flujo aéreo limitado
Valores mayores	> 1.5	Defecto de cierre por incompetencia glótica. Lesiones de borde libre
Valores menores	< 1.5	Sin significación clínica para los autores citados

Tiempos máximos de fonación según edad y sexo, Kent (1987) presentados en (Tapia & Cobeta, 1996).

17. Flujo medio aéreo

Fonación /pa/	Flujo medio aéreo
F0 normal e intensidad comfortable	120 ml/s
F0 baja e intensidad comfortable	105 ml/s
F0 alta e intensidad comfortable	92 ml/s
F0 normal e intensidad baja	145 ml/s
F0 normal e intensidad alta	200 ml/s

Flujo medio aéreo, (Fernández & Núñez, 2013)

2. Resistencia Glótica

Fonación	Resistencia glótica
F0 normal e intensidad comfortable	38.83 cm H2O x ml/s
F0 baja e intensidad comfortable	55.29 cm H2O x ml/s
F0 alta e intensidad comfortable	34.12 cm H2O x ml/s
F0 normal e intensidad baja	57.35 cm H2O x ml/s
F0 normal e intensidad alta	93.67 cm H2O x ml/s

Resistencia glótica (Fernández & Núñez, 2013)

18. Promedio de velocidad de flujo e implicancias clínicas

Condición	Valor	Implicancia
Normal	80 – 200 ml/s	
Aumentado	400 – 600 ml/s	Incompetencia glótica
Disminuido	10 – 15 ml/s	Voz apretada, registro fry
Irregular		Alteraciones neurológicas ^{as}

Promedio de velocidad de flujo (Guzmán, 2017), presentado en apunte de aprendizaje “Medidas aerodinámicas de la fonación”, Comunikrte.

19. Presión subglótica e implicancia clínica

Condición	Presión subglótica
Rango normalidad	5-10 cm H ₂ O
Voz soplada	< 10 cm de H ₂ O
Voz apretada	> 10 cm de H ₂ O

Presión subglótica (Guzmán, 2017), presentado en apunte de aprendizaje “Medidas aerodinámicas de la fonación”, Curso Evaluación objetiva de la voz, Comunikrte.

Cabe mencionar que la presión necesaria para una fonación a intensidad conversacional es entre 5 y 10 cm de H₂O en voces normales (Stemple, 2015; Behrman, 2013; Titze, 2010; Hixon, Weismer, Hoit, 2008).

20. Umbral de presión de la fonación

En sujetos portadores de voces normales se espera que el umbral se encuentra entre 3 y 5 cm de H₂O (Titze, 1992; Fisher, Swank, 1997; Titze, 2010; Stemple, 2014).

Registro	Umbral de presión de la fonación
Fry	3 cm de H ₂ O
Falsete	8 cm de H ₂ O
Modal	3-5 cm de H ₂ O

Umbral de presión de la fonación (Guzmán, 2017), presentado en apunte de aprendizaje “Medidas aerodinámicas de la fonación”, Curso Evaluación objetiva de la voz, Comunikrte.

21. Cociente de contacto

Se considera el CQ en porcentajes de tiempo de contacto de cada ciclo vibratorio de los pliegues vocales de 0% a 100%, considerando valores cercanos a 0% cuando existe un contacto escaso y a medida que se acrecientan los porcentajes se produce un mayor tiempo de aducción. (Jeréz y Cols, 2016).

Vaca y Cobeta (2013), consideran para emisiones en registro modal valores aproximados a 40%. Por otro lado, Howard (2009) considera valores de normalidad entre un 40% y 60%, indicando que valores mayores se pueden atribuir a una fonación apretada.

22. Valores normativos del PRAAT

Vocal /a/	Jitter local	Jitter local absol	Jitter rap	Jitter ppq5	Shim local	Shimm local dB	Shimm apq11	NHR	HNR
Valores limites	1.04	83.200	0.68	0.84	3.81	0.35	3.07	0.19	12
Normal	0.63	27.000	0.35	0.36	1.99	0.17	1.39	0.11	16.5

Boersma & Weenink (2001)

23. Escala de pérdida de armónicos

Afonía	0 a 5000 hz
Alteración severa	500 a 5000 hz
Alteración moderada a severa	1000 a 5000 hz
Alteración moderada a severa	1500 a 5000 hz
Alteración moderada	2000 a 5000 hz
Alteración moderada	2500 a 5000 hz
Alteración leve a moderada	3000 a 5000 hz
Alteración leve a moderada	3000 a 5000 hz

Escala de pérdida de armónicos (Cecconello, 2012).

24. Clasificación de astenia

Tipo I	Ausencia de armónicos en altas frecuencias, por encima de 4000
Tipo II	Tipo I + ausencia de armónicos entre 1° y 2° formante
Tipo III	Tipo II + desaparición del 2° formante
Tipo IV	Tipo III + desaparición del 1° formante

Clasificación de astenia (falta de armónicos) según Núñez Batalla et al (2004).

25. Grados de severidad según valoración perceptual

	Leve	Moderado	Severo
TMF	De 10 a 15 segundos	Alrededor de 10 segundos	Menor de 10 segundos
Eficiencia de cierre glótico y respiratoria	Levemente ineficiente	Mayor falta de cierre y escape de aire	Ineficiencia absoluta de cierre glótico y de control respiratorio
Relación s/z	Cociente 1 o ligeramente aumentado (1.1)	Cociente aumentado, 1.2	Cociente elevado, más de 1.2 o menor a 1
Efecto Kayser	Alteración inestable de ataque y/o filatura	Ataque y filatura alteradas, cuerpo sin alteración o con leve inicio de alteración	Ataque, cuerpo y filatura muy alteradas (golpe glótico, temblor, cortes, cese brusco de emisión, etc)
Frecuencia fundamental	Adecuada o desplazada por momentos	Desplazada 1, 1 1/2 tonos	Desplazada más de dos tonos
Presencia de registros	Conservados, con posible soplo en agudos, o reemplazo (modificable) de tonos de modal por falsete y viceversa	Conservados pero con cortes, agujeros tonales, dificultades de pasaje	Pérdida de alguno o más registros de frito y de falsete, o de falsete, o de frito y modal
Intensidad	Regulación conservada	Se regula con dificultad, se dificulta la emisión en pianos o en fortes	Se pierde en forma permanente la emisión en pianos o en fortes
Timbre	Adecuado o levemente fuera de foco (ej: nasal)	Alteración tímbrica modificable con estimulación	Alteración tímbrica permanente que influye en la falta de inteligibilidad
Ligados en escala	Conservado	Aparecen dificultades intermitentes	Pérdida de ligado
Ruidos en la voz	Ninguno	Aparición intermitente de aspereza, diplofonía, tensión, cortes	Presencia de ruidos permanente, voz áspera y ronca con pérdida de musicalidad
Fatigabilidad al examen	Ninguna o muy leve	Aparece después de un tiempo considerable	Aparece enseguida

Grados de severidad según valoración perceptual (Farías, 2007).



MANUAL BÁSICO DE PRAAT PARA ANÁLISIS DE VOZ

Introducción de la sección

El PRAAT es un programa desarrollado en la Universidad de Amsterdam por Paul Boersma y David Weenink a partir del año 1992. Dentro de las ventajas en relación a otros software se encuentra que es un programa de libre distribución, código abierto, multiplataforma y gratuito. Fue diseñado especialmente para investigaciones en fonética, sin embargo puede cumplir múltiples propósitos como son análisis acústico, síntesis articulatoria, procesamiento estadístico de datos, edición y manipulación de señales de audio, entre otros. El programa se actualiza constantemente. En cuanto a los inconvenientes o dificultades podría contarse que al ser un programa especializado requiere adiestramiento y manejo de mucha información, por otro lado, no tiene un formato amigable a primera vista ni muy intuitivo para los que se están iniciando en la utilización de dicho programa.

Una vez revisados ya los principales parámetros acústicos de la voz y su normativa, ahora veremos como obtener dichos parámetros. El propósito de este apartado del manual es ser una guía básica para los estudiantes que se están iniciando en el estudio y análisis de la voz. Se describe paso a paso y detalladamente el uso de diferentes opciones que nos brinda el programa y se sugiere ir leyendo esta parte con el programa abierto en un computador. Cada apartado irá con un link asociado en donde se mostrará el video del procedimiento a ejecutar.

El objetivo de aprendizaje que se pretende desarrollar en este apartado del manual es:

- * Aplicar instrumentos de evaluación objetivos relativos al proceso evaluativo de la voz.

1. Grabación de la señal

Antes de comenzar a utilizar el PRAAT se sugiere al lector remitirse al apartado anterior de acústica de la voz para entender algunos conceptos básicos del análisis acústico y poder comprender e interpretar los procedimientos y resultados entregados por el programa.

Para comenzar a grabar, el primer paso es descargar el programa y para ello podemos ir al siguiente link <https://www.fon.hum.uva.nl/praat/>.



Imagen 1. Página para descargar el Praat.

A continuación en Download Praat descargamos la versión compatible con nuestro sistema operativo (Macintosh, windows, etc). Para windows se sugiere la versión 64 bits. Una vez instalado en nuestro computador, debería aparecer el siguiente ícono



Imagen 2. Icono programa Praat.

Luego hacemos doble click en el ícono y debería desplegarse la siguiente ventana:

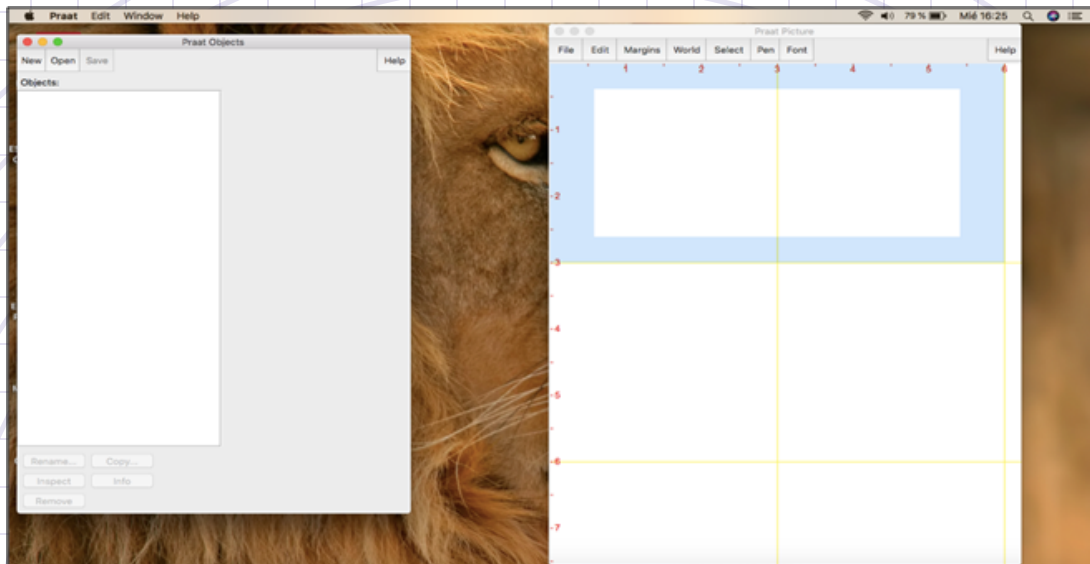


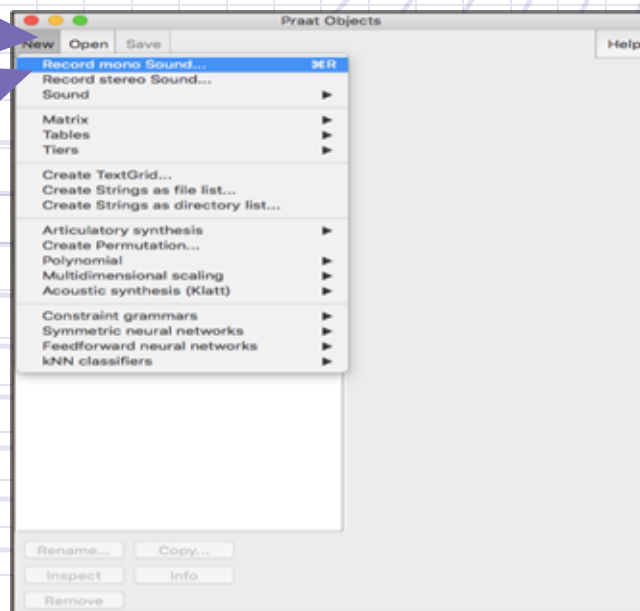
Imagen 3. Pantalla inicial al abrir programa Praat.

Como se puede observar, se abren 2 cuadros uno denominado Praat Picture a la derecha y otro denominado Praat Objects a la izquierda. Cada cuadro tiene su respectivo panel de menú en la parte superior. Para grabar una muestra de voz trabajaremos con Praat objects. Mientras tanto, minimizamos el cuadro correspondiente a Praat Pictures ya que lo ocuparemos más adelante para graficar.

Los pasos para grabar una nueva señal de audio son los siguientes:

En Praat Objects seleccionamos New y luego la opción Record Mono Sound.

Imagen 4. Selección opción de Record Mono Sound.



Seleccionada la opción se desplegará la siguiente ventana:

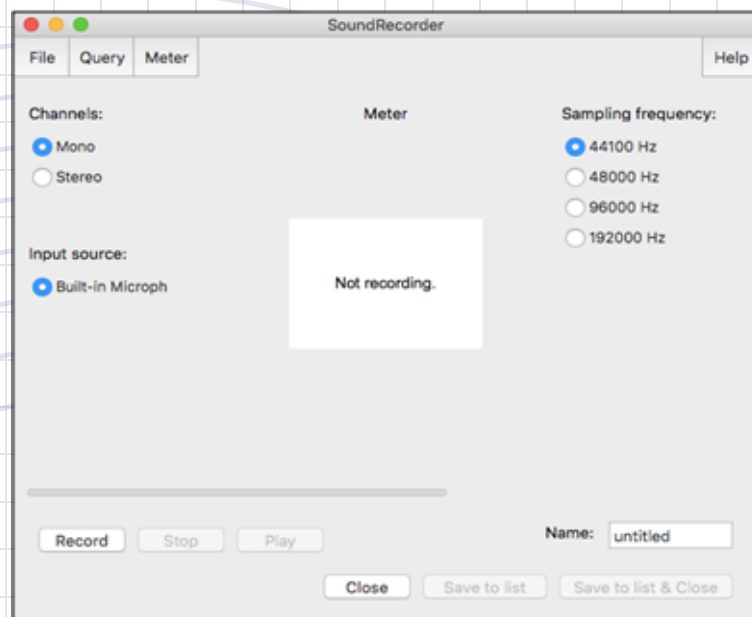


Imagen 5. Panel de grabación de Praat.

En esta ventana podemos realizar las siguientes acciones:

- En **Channels** seleccionamos el canal, que para este efecto se debe dejar en **A**.
- En **Input Source** se selecciona el dispositivo de entrada del sonido. En este caso aparece el micrófono incorporado del computador. Si conectamos un micrófono externo acá debe salir la opción para seleccionarlo.
- En **Sampling frequency**(frecuencia de muestreo) seleccionamos 44100 Hz lo cual es suficiente para voz hablada y cantada.
- Al centro **Meter** nos indica el nivel de la señal. Al comenzar la grabación debemos asegurarnos que al captar la señal el **Meter** permanezca en color verde lo cual indica que el nivel es correcto. Si cambia a color amarillo o rojo es porque la señal está saturada entregando un análisis de mala calidad, por tanto se aconseja tomar nuevamente la señal y nivelar el volumen del micrófono.
- En **Name** le colocamos nombre a la grabación ya que por defecto el programa le asigna el nombre **untitled**.
- Para comenzar a grabar presionamos **Record**.
- Para finalizar la grabación seleccionamos **Stop**.

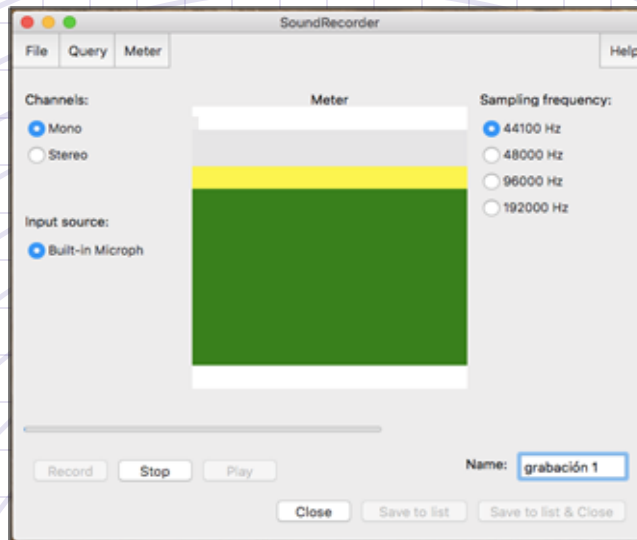


Imagen 6. Panel de grabación de Praat en este caso indicando una señal saturada.

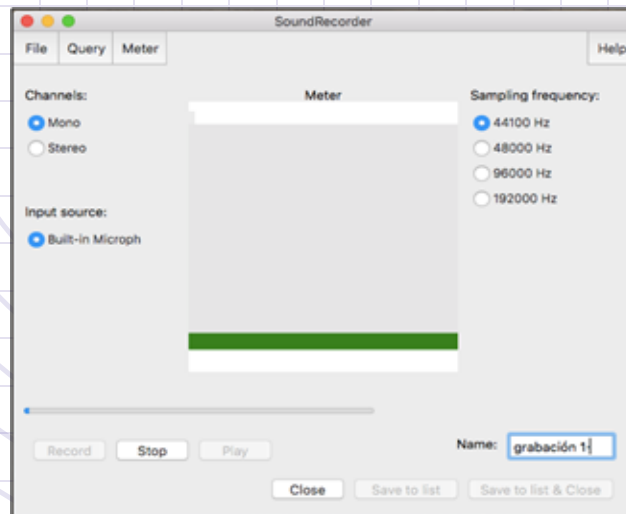


Imagen 7. Panel de grabación de Praat con un nivel de señal adecuado.

Una vez finalizada la grabación tenemos las siguientes opciones:

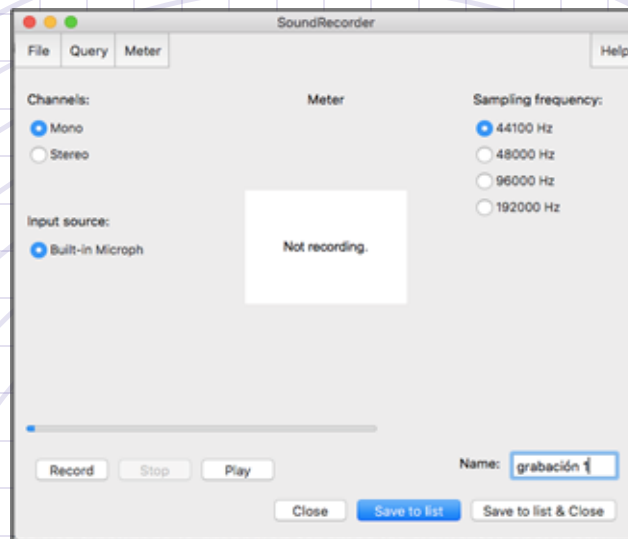


Imagen 8. Panel de grabación de PRAAT con una señal de voz tomada.

- Seleccionamos la opción **Save to list & Close** el cual guardará y creará el objeto en la ventana **Praat Objects** cerrando posteriormente la ventana de grabación.

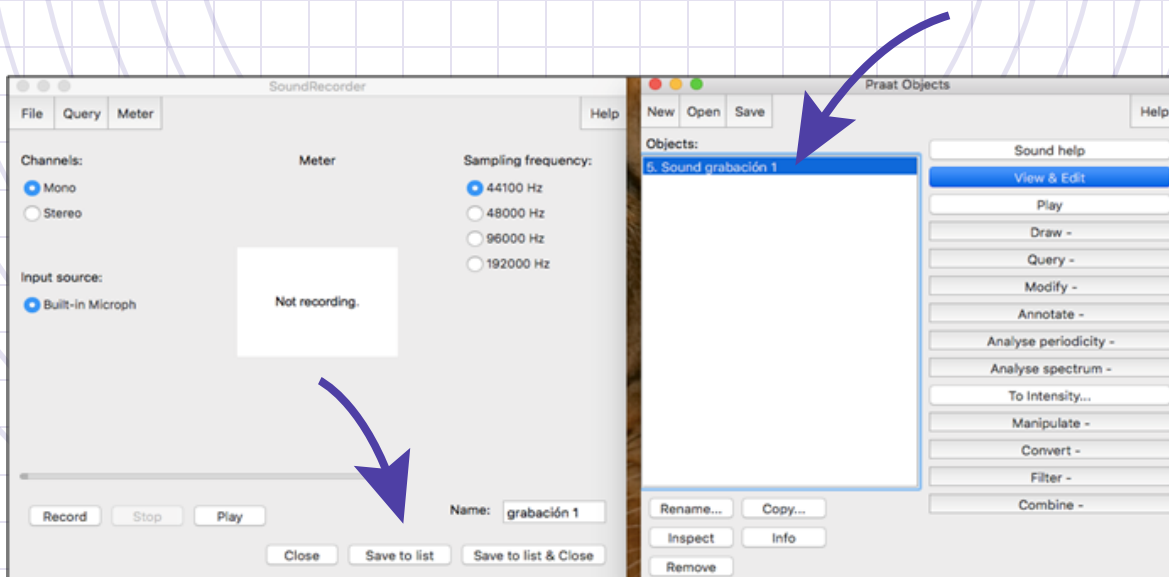


Imagen 9. Creación de objeto en ventana de trabajo Praat Objects

- Seleccionamos la opción **Save to list & Close** el cual guardará y creará el objeto en la ventana **Praat Objects** cerrando posteriormente la ventana de grabación.
- Si deseamos guardar la señal de audio como un archivo para posteriormente trabajar en él, seleccionamos la opción **File** y elegimos el formato del archivo.

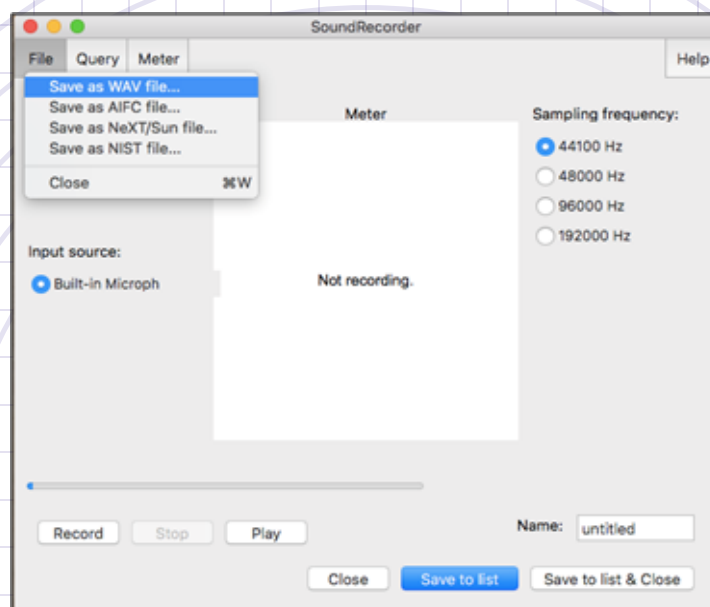


Imagen 10. Creación de un archivo de audio en panel de grabación Praat.

NOTA

Al momento de grabar recordar las recomendaciones sugeridas en cuanto a tipo de micrófono, distancia, ruido ambiente, etc. Para eso remitimos al lector a la sección de análisis acústico de la voz en donde se indican las sugerencias dadas por la National Center for voice and Speech (NCVS).

- Para analizar una señal de audio ya tomada con anterioridad y guardada como archivo, en **Praat Objects** debemos seleccionar la opción **Open** luego **Read from file**, seleccionar el archivo abrirlo y se creará el objeto de sonido en la ventana **Praat Objects** para posteriormente trabajar con él.

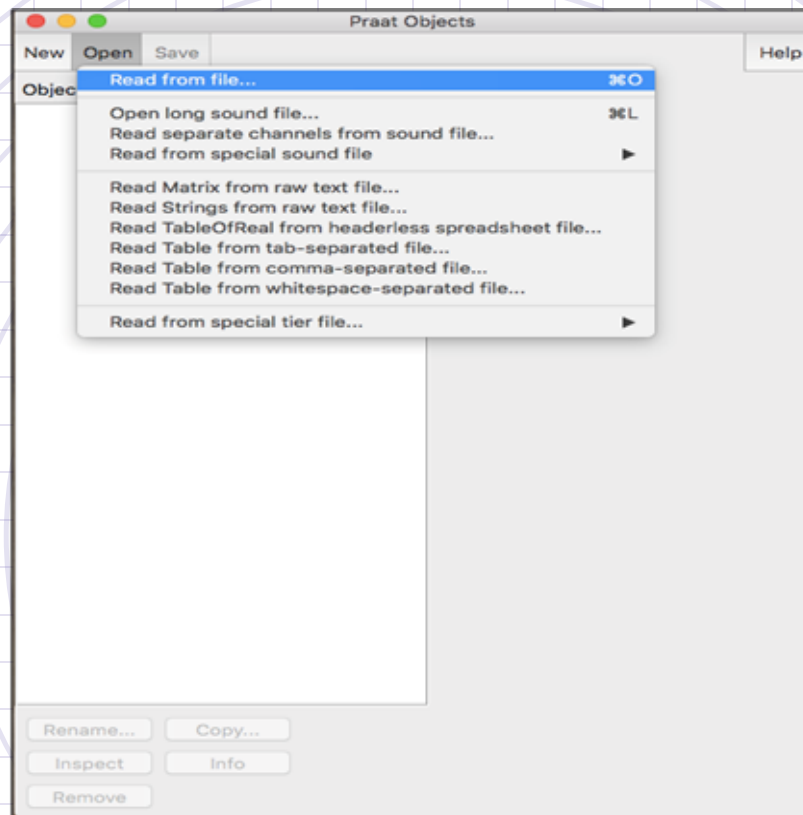


Imagen 11. Creación de un objeto en ventana de trabajo Praat Objects a partir de un archivo de audio.

2. Análisis de la señal

Para comenzar el análisis con la señal de audio ya en nuestra ventana de trabajo **Praat Objects** presionamos la opción **View & Edit** se desplegará la siguiente ventana:

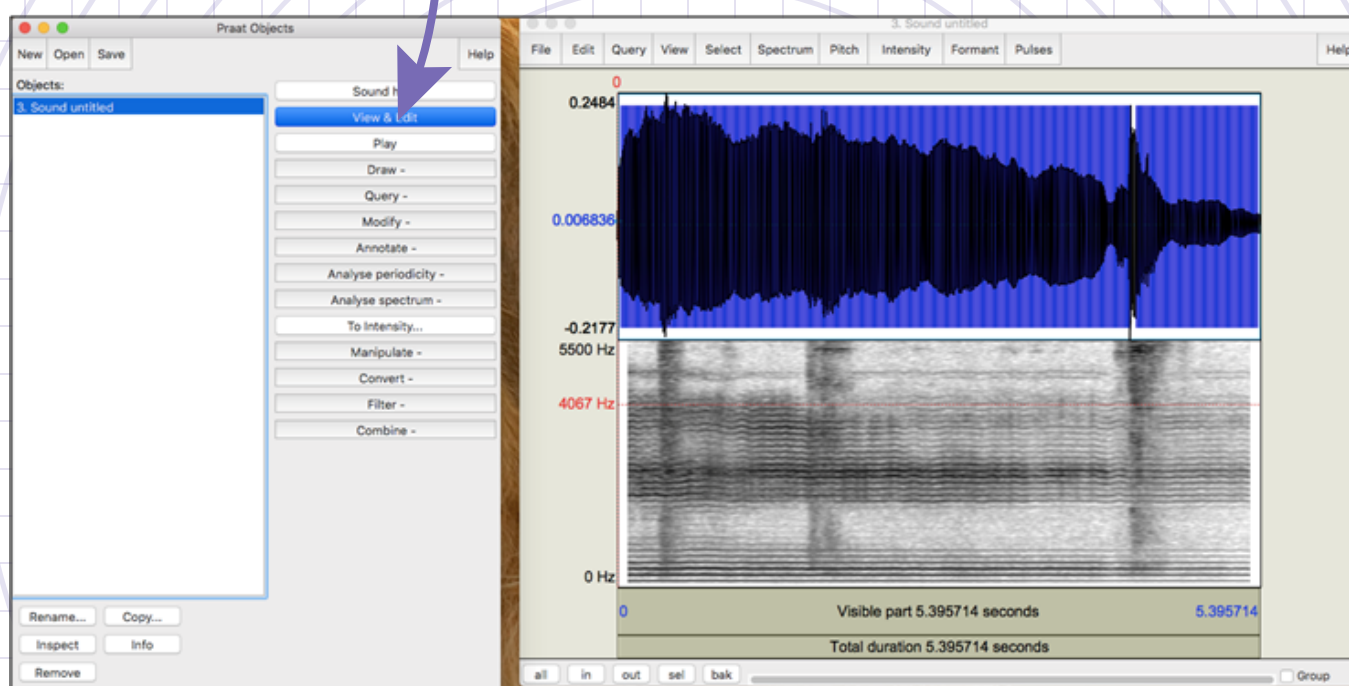


Imagen 12. Análisis de señal de audio

NOTA: El programa no analiza señales que duren más de 10 segundos, por tanto para visualizar el espectrograma debemos seleccionar una parte de la grabación. (o podemos modificar el largo máximo de análisis).

Como podemos ver en la figura 12 en el extremo inferior izquierdo tenemos los siguientes botones:

- All:** Nos muestra toda la grabación.
- In:** Permite hacer acercamientos de la señal.
- Out:** Permite volver atrás.
- Sel:** Muestra sólo la sección seleccionada.
- Bak:** Retroceder a la sección anterior.

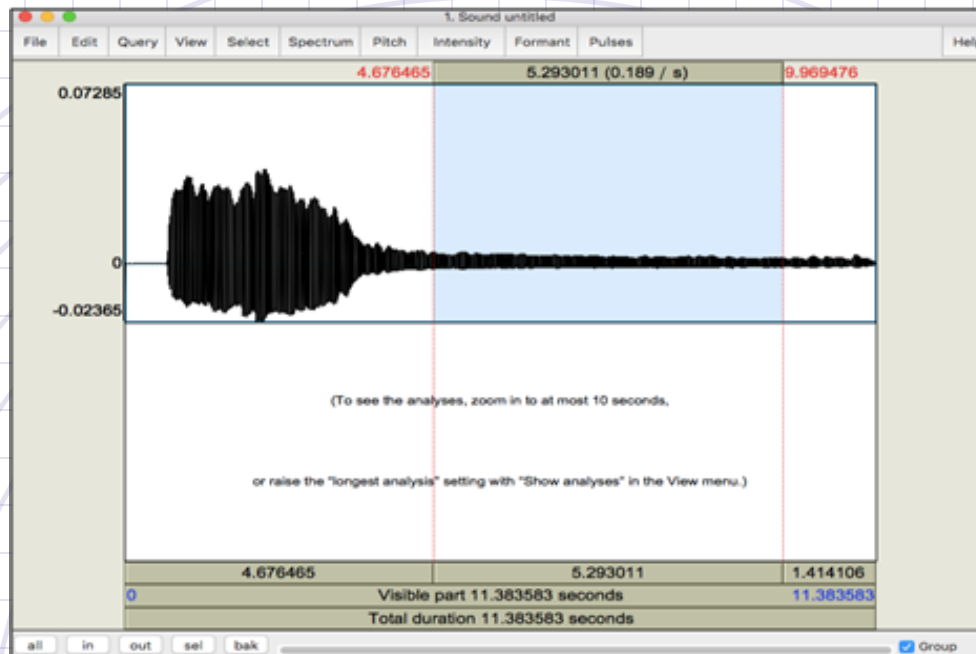


Imagen. 13. Análisis de una señal mayor a 10 segundos.

Por otra parte, cuando analizamos la señal se despliegan dos gráficos. En la parte superior un oscilograma que nos permite visualizar la forma de la onda y en la parte inferior un espectrograma.

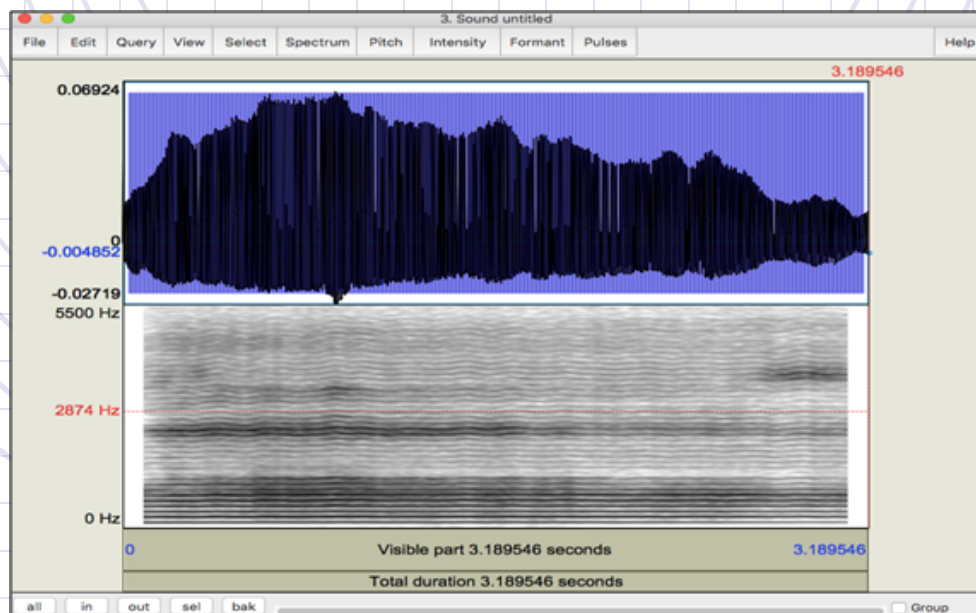


Imagen.14. Oscilograma en la parte superior y espectrograma en la parte inferior en la ventana de análisis.

Para graficar el oscilograma y adjuntarlo a un informe de análisis tenemos dos opciones, sacar una foto de la pantalla de trabajo o graficarlo en la ventana Praat Picture. Para la segunda opción debemos realizar los siguientes pasos:

- Presionamos el boton **File** (pestaña superior izquierda) y luego seleccionamos la opción **Draw visible sound**.
- Se despliega otra ventana en donde mantenemos los valores que vienen por defecto en el programas y presionamos **OK**.

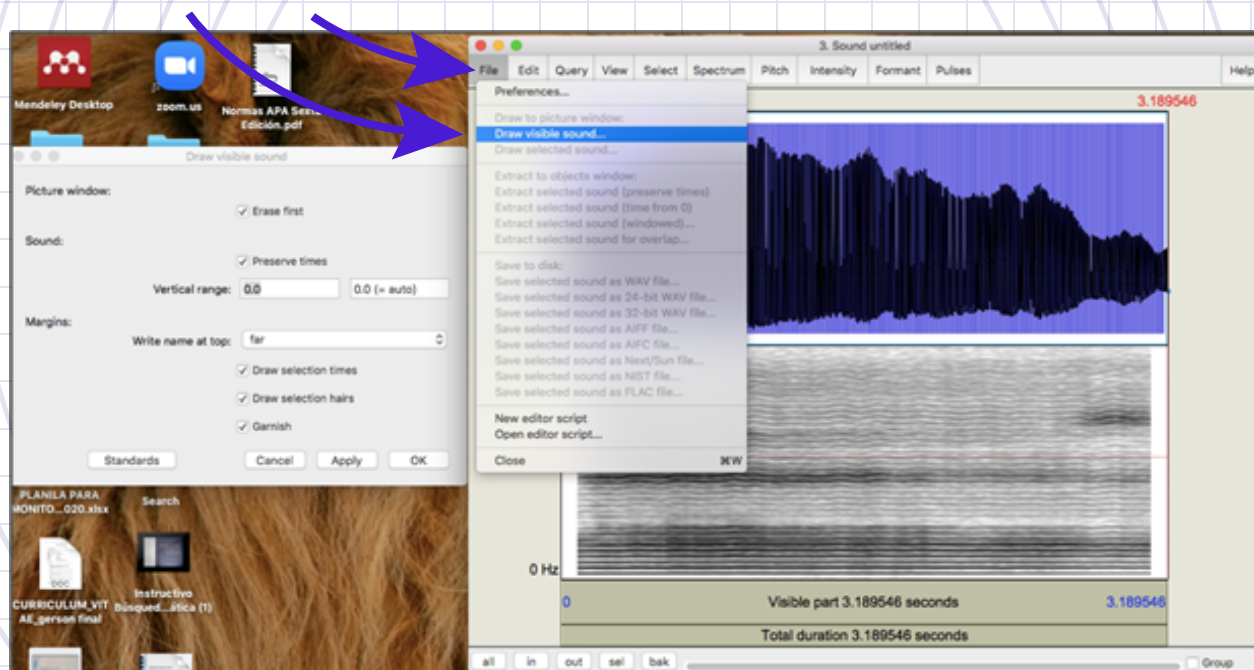


Imagen. 15. Graficar oscilograma en ventana Praat Picture.

- Una vez seleccionado **OK** se despliega la ventana **Praat Picture** con el oscilograma.

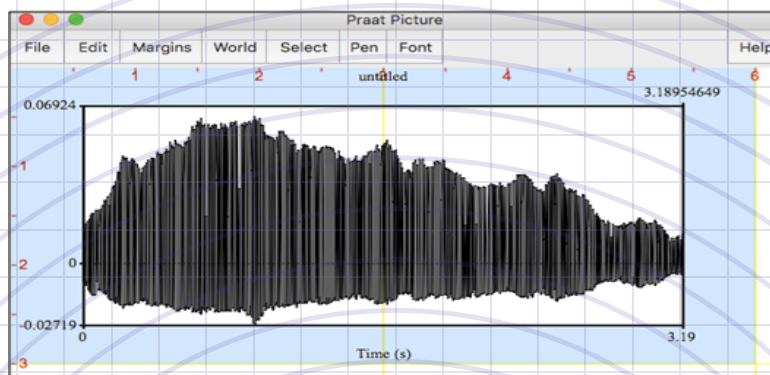


Imagen. 16. Gráfico de oscilograma en Praat Picture

- Para guardar el gráfico, seleccionamos **File** elegimos el formato en que queremos guardarlo (PDF, PNG, etc).

NOTA: Recordar que para que aparezca el gráfico en la ventana Praat Picture debemos tener seleccionado con anterioridad donde queremos graficarlo. De lo contrario no aparecerá el gráfico.

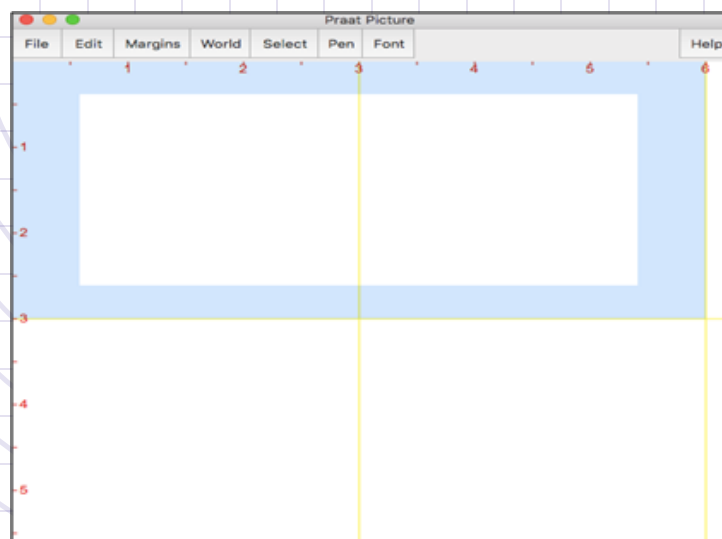


Imagen. 17. Selección de área para graficar en Praat Picture.

NOTA: Es importante saber que no todas las voces son candidatas a ser analizadas mediante análisis acústico. En voces muy caóticas los valores cuantitativos pierden validez. En ese sentido la National Center for Voice and Speech (Titze, 1995) recomienda que la evaluación debe comenzar con un análisis espectrográfico para clasificarla con el fin de determinar los métodos más idóneos de evaluación y en base a eso podemos determinar qué tipo de medidas utilizar.

3. Espectrograma

En el análisis espectrográfico tenemos dos opciones. Configurar el programa para obtener un espectrograma de banda ancha o espectrograma de banda estrecha (Respecto de las diferencias y ventajas de uno u otro remitimos al lector a la sección de análisis acústico de la voz).

- Para comenzar el análisis espectrográfico en la ventana de análisis seleccionamos la opción **Spectrum** y luego **Spectrogram settings..** y se desplegará una nueva ventana de configuración.
- En **View range(Hz)** modificamos el rango de frecuencias y colocamos de 0 a 5000 Hz en el caso de que sea hombre o de 0 a 5500 Hz en el caso de que sea mujer. Si queremos analizar frecuencias más agudas como en el caso de una cantante podemos aumentar el rango.
- En **Window length (s)** determinamos el tipo de espectrograma que queremos obtener. Para espectrograma banda ancha colocamos 0.005. Para espectrograma banda estrecha 0.09.
- En la opción **Dynamic range (dB)** colocamos entre 50 y 70 dB.

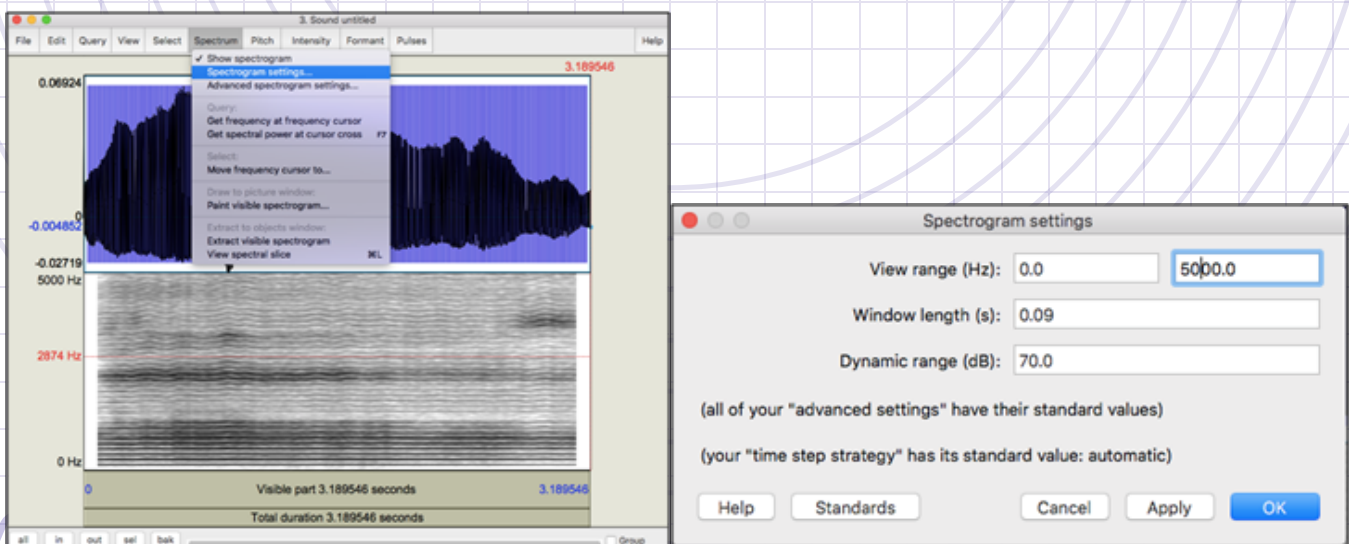


Imagen. 18. Configuración Espectrograma

Para graficar seleccionamos **Spectrum** luego **Paint visible spectrogram** y se genera otra ventana de configuración. Dejamos la configuración que viene por defecto como se muestra en la siguiente figura

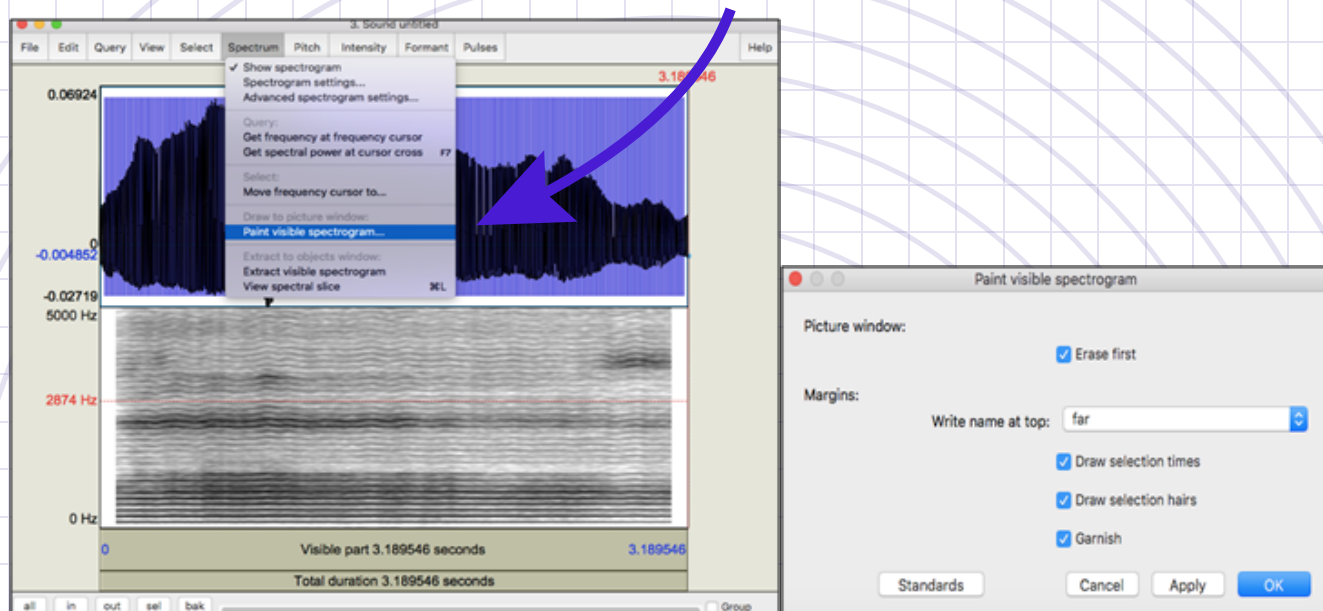


Imagen. 19. Configuración para graficar espectrograma en ventana Praat Picture.

- Presionamos **OK** y obtenemos nuestro espectrograma en la ventana **Praat Picture**. Luego en **File** seleccionamos el formato en que queremos guardarlo

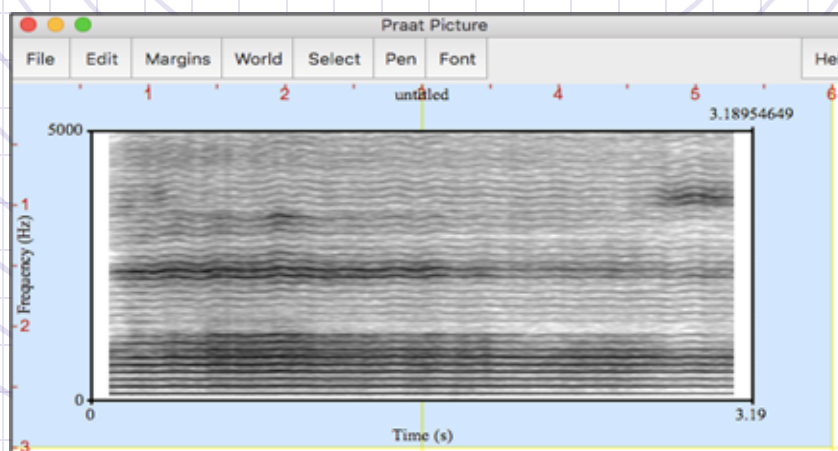


Imagen. 20. Espectrograma banda estrecha graficado en Praat Picture.

4. Análisis de Formantes

Para analizar los formantes debemos trabajar sobre un espectrograma de banda ancha, en donde las zonas más oscuras corresponderán a los formantes.

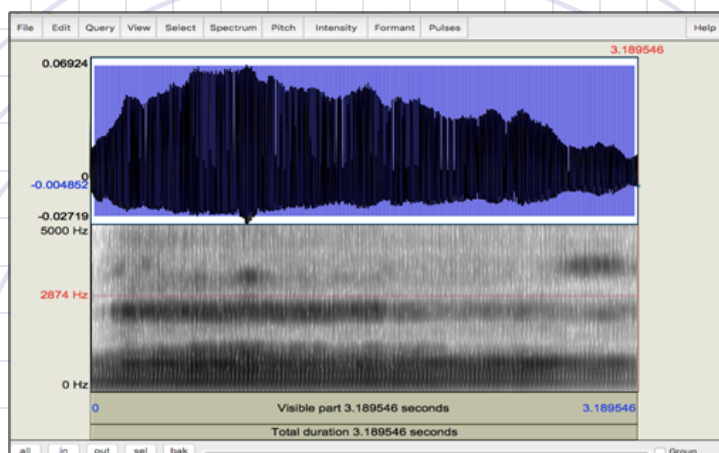


Imagen. 21. Espectrograma banda ancha

Si queremos visualizarlos y marcarlos en el gráfico seleccionamos la opción **Formant** luego pinchamos **Show formants** y los formantes aparecerán como líneas rojas sobre el espectrograma.

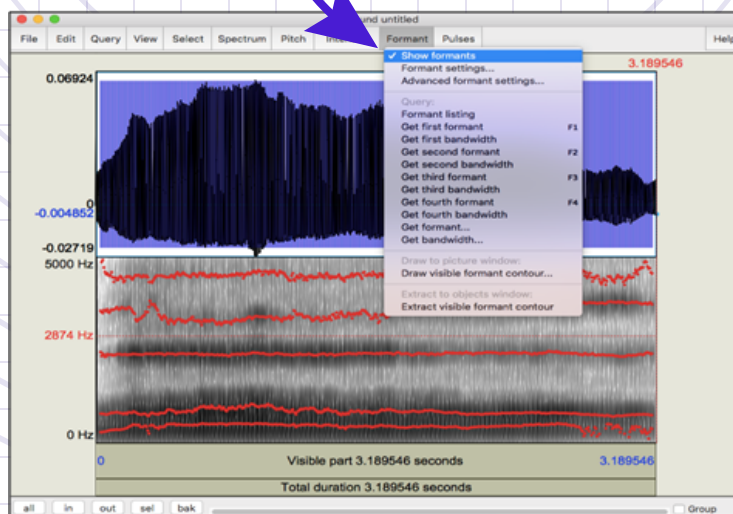


Imagen. 22. Visualización de formantes en espectrograma banda ancha.

Luego podemos realizar un análisis cuantitativo de los formantes en cuanto a su valor y ancho de banda. Para obtener el valor de un formante seleccionar toda la señal y obtener un promedio o colocar el cursor en un solo punto del espectrograma y obtener el valor del formante en ese punto. Luego seleccionamos la opción **Formant** y **Get first formant** en el caso del primer formante **Get second formant**, etc, dependiendo del formante que queremos analizar y se despliega un cuadro que nos indica el valor. Una vez obtenido el valor podemos compararlo con los valores normativos.

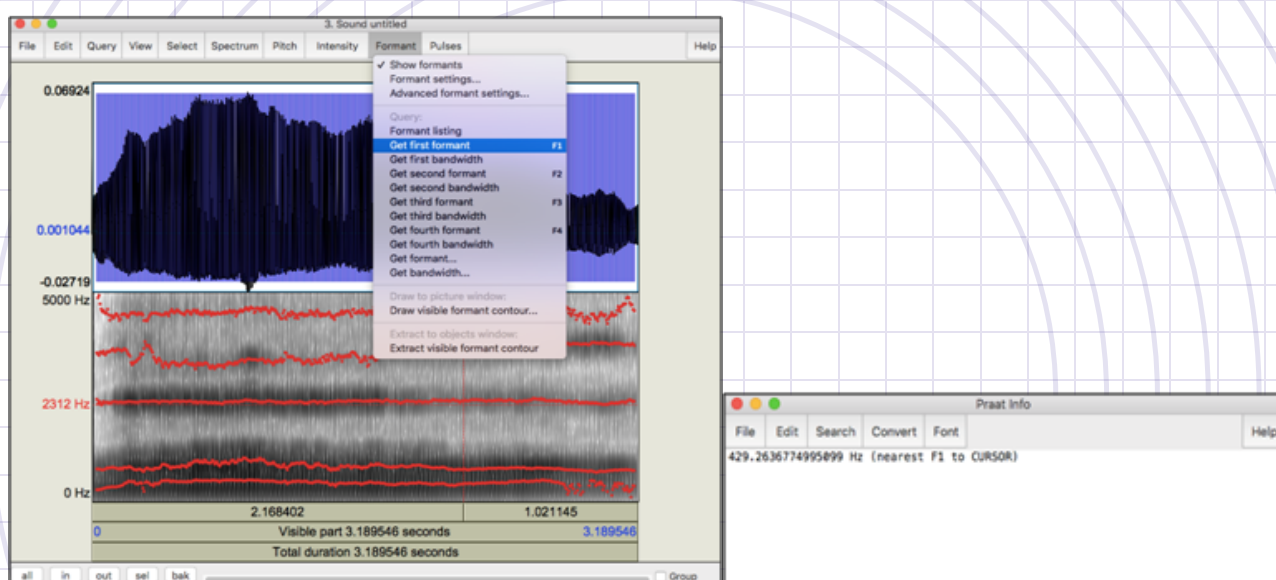


Imagen. 23. Obtención del valor de los formantes.

Para obtener el valor del ancho de banda seleccionamos **Formanty** luego la opción **Get first Bandwidth** en el caso del primer formante o dependiendo el que queramos analizar. Luego se despliega un cuadro con el valor y podemos compararlo también con la data normativa existente.

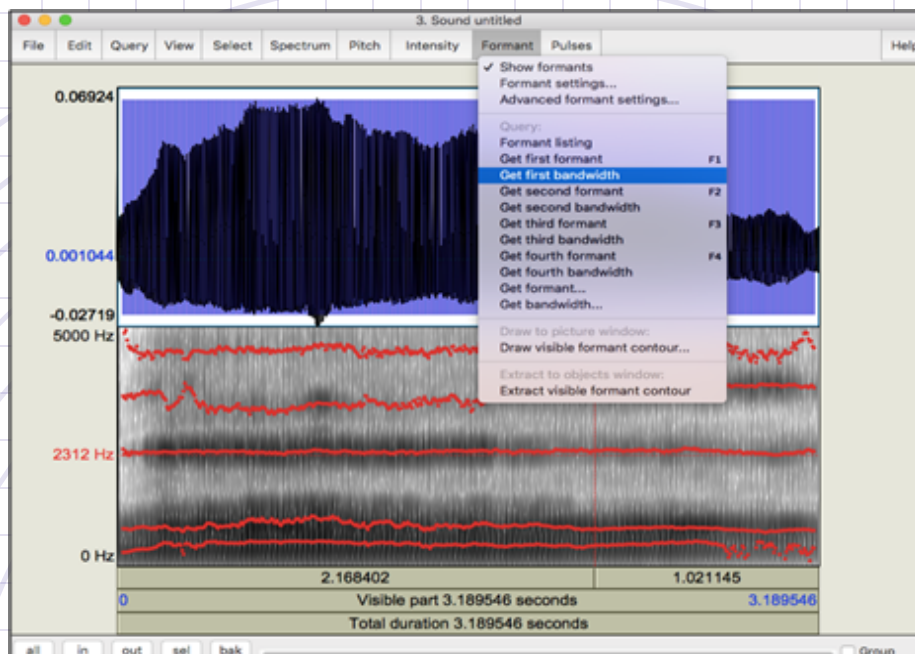


Imagen. 24. Obtención del ancho de banda de los formantes.

Respecto del análisis cualitativo podemos describir si estos se encuentran bien definidos o dispersos indicando mayor dispersión de la energía.

Para graficar los formantes en la ventana **Praat Picture**, seleccionamos **Formant**, luego **Draw visible formant contour**. Se despliega una ventana de configuración y presionamos **OK**.

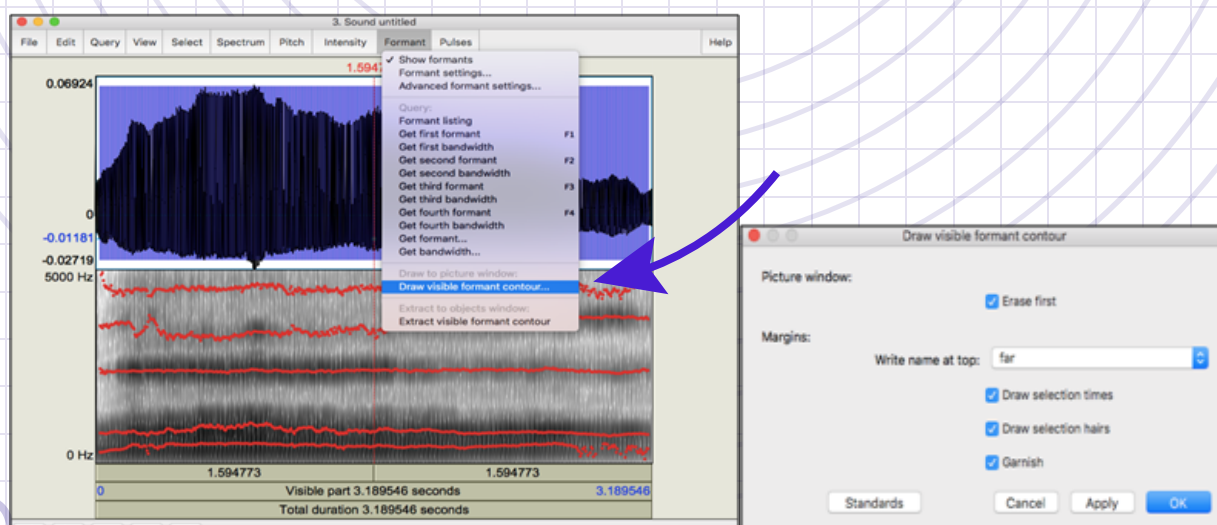


Imagen. 25. Como graficar los formantes en la ventana Praat Picture

Una vez presionado **OK** obtenemos nuestro gráfico de los formantes en la ventana **Praat Picture**. Seleccionamos **File** y guardamos según el formato deseado.

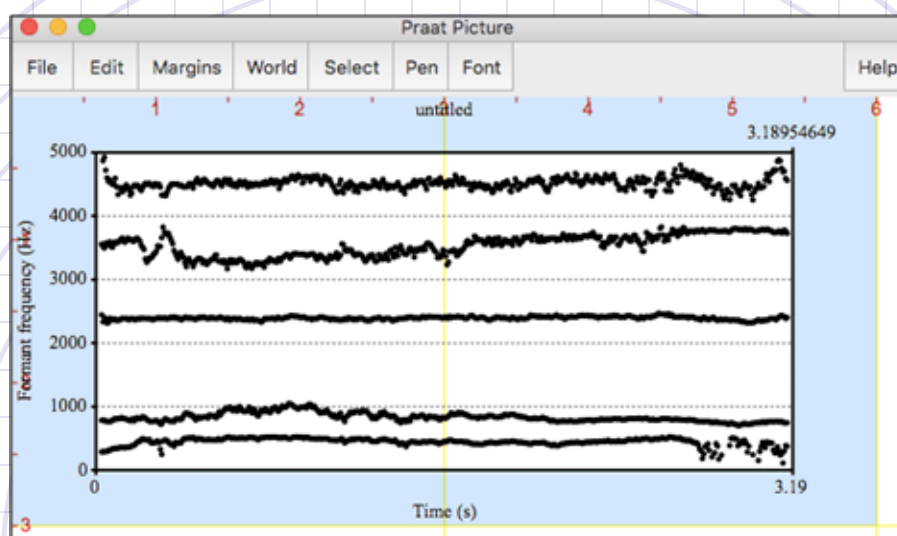


Imagen. 26. Gráfica de los formantes en ventana Praat Picture.

5. Análisis espectral: Espectro FFT (Power Spectrum)

Para obtener el espectro FFT debemos trabajar con la emisión de una vocalización sostenida (una /a/ en este caso) y con un espectrograma banda estrecha, ya que en el espectro FFT queremos estudiar los armónicos.

- Primero debemos seleccionar un punto del espectrograma (colocar cursor en un punto del espectrograma), ir a **Spectrums** luego seleccionamos **View spectral slice** y se nos despliega otra ventana con el espectro FFT.

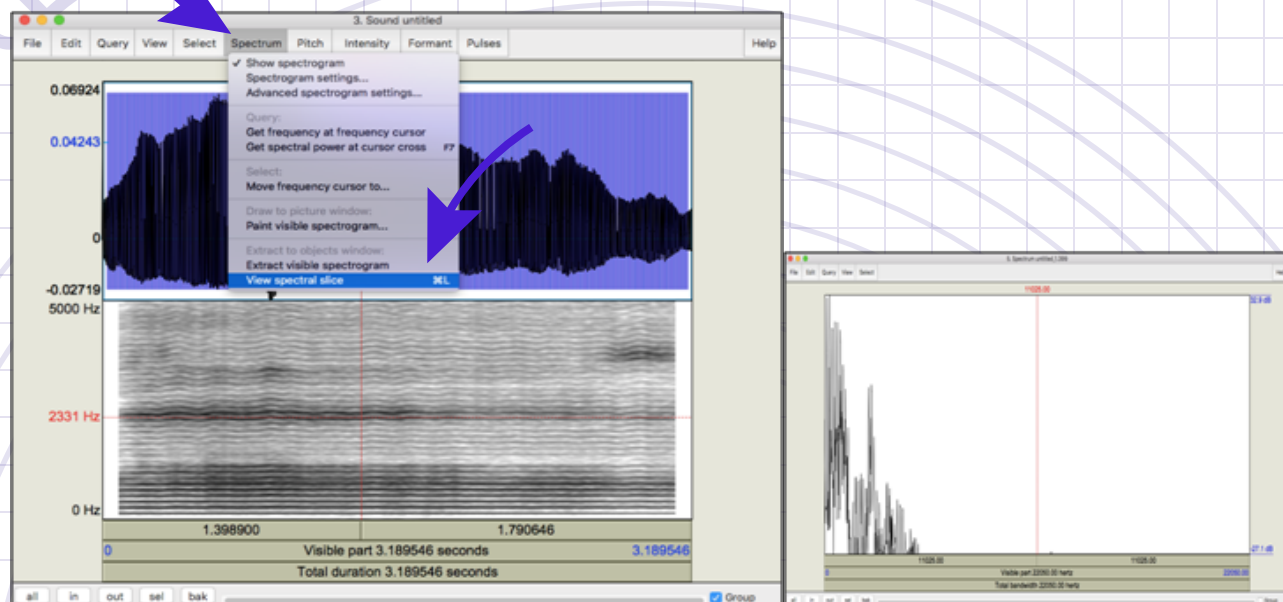


Imagen.27. Obtención espectro FFT

- Luego se debe cambiar la configuración del rango de frecuencias. Ir a **View**, **Zoom**, configurar hasta 5000 Hz y presionamos **OK**.

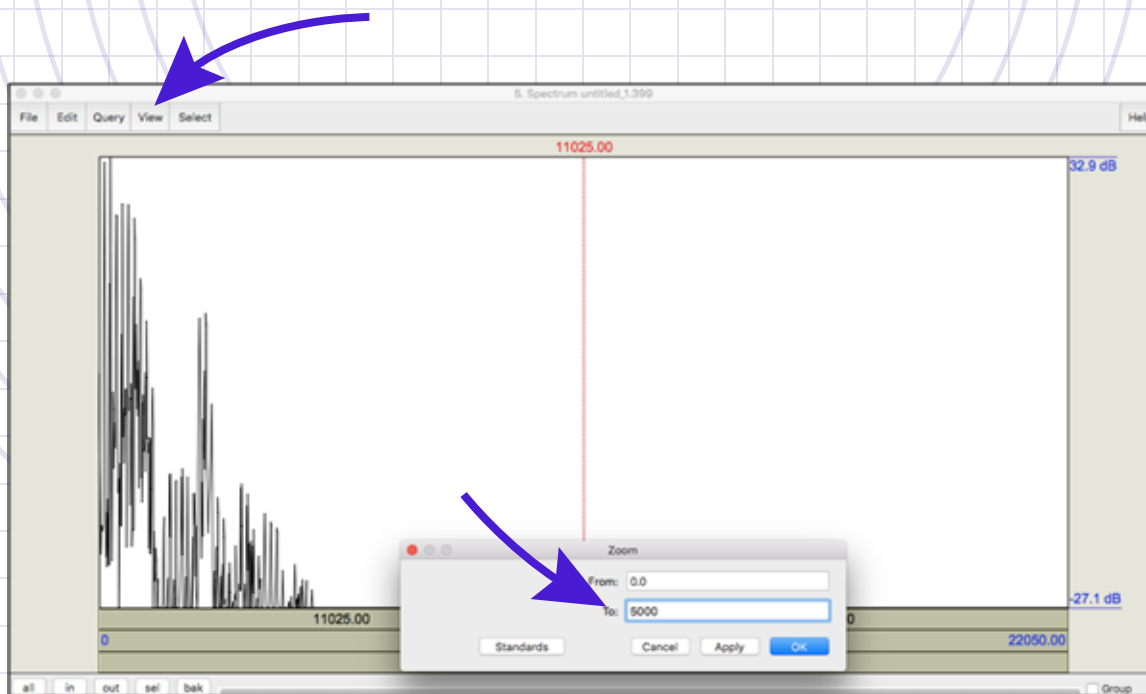


Imagen. 28. Configuración rango de frecuencias

- Una vez presionado **OK**, obtenemos nuestro espectro FFT.

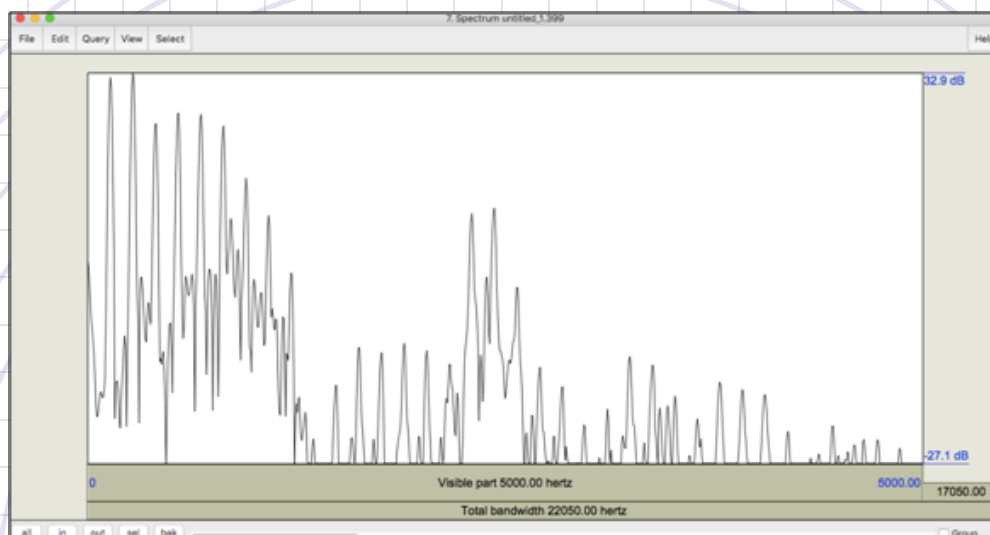


Imagen. 29. Espectro FFT

- Una vez obtenido, se nos genera en la ventana **Praat Objects** nuestro espectro y para graficarlo en nuestra ventana **Praat Picture**, debemos ir a Draw.

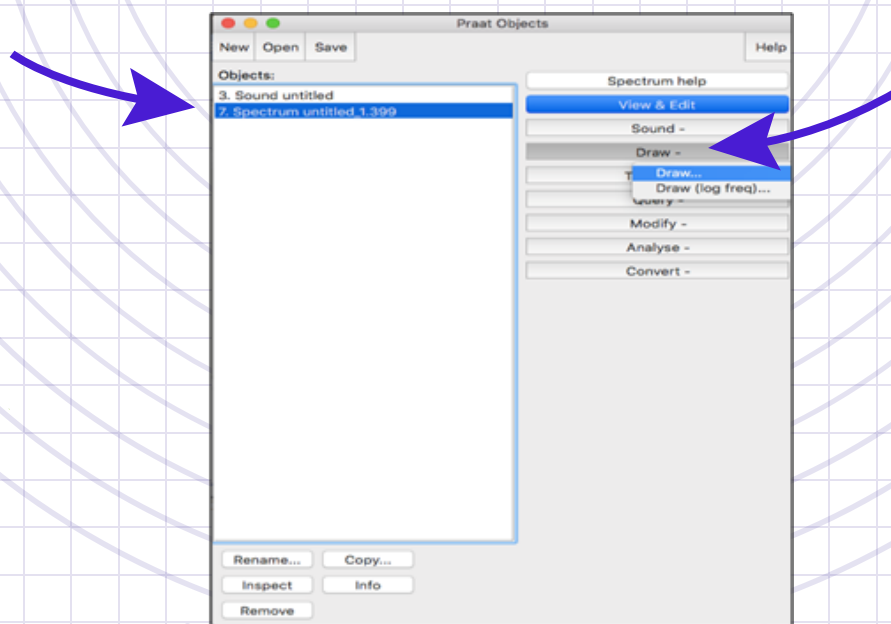


Imagen. 30. Como graficar el espectro FFT

- Se nos despliega una ventana de configuración y debemos colocar nuestro rango frecuencial hasta 5000 Hz.

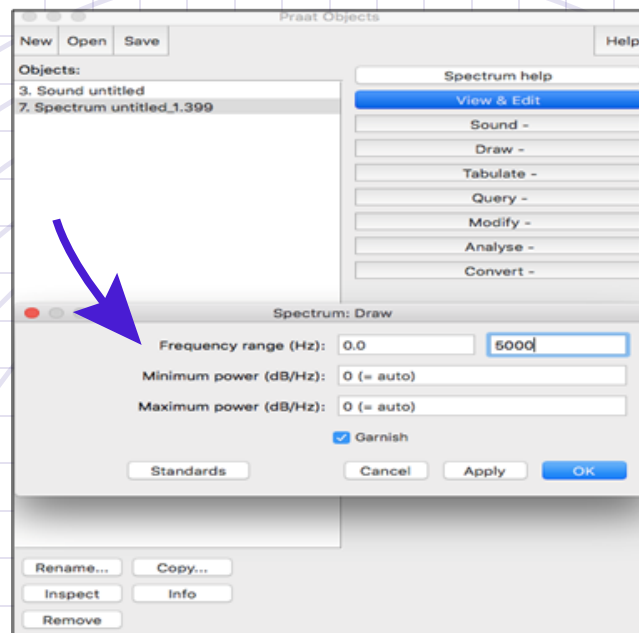


Imagen. 31. Configuración rango frecuencial

- Presionamos OK y aparecerá nuestro gráfico del espectro FFT en la ventana Praat Picture. Posteriormente presionamos File guardamos según el formato deseado.

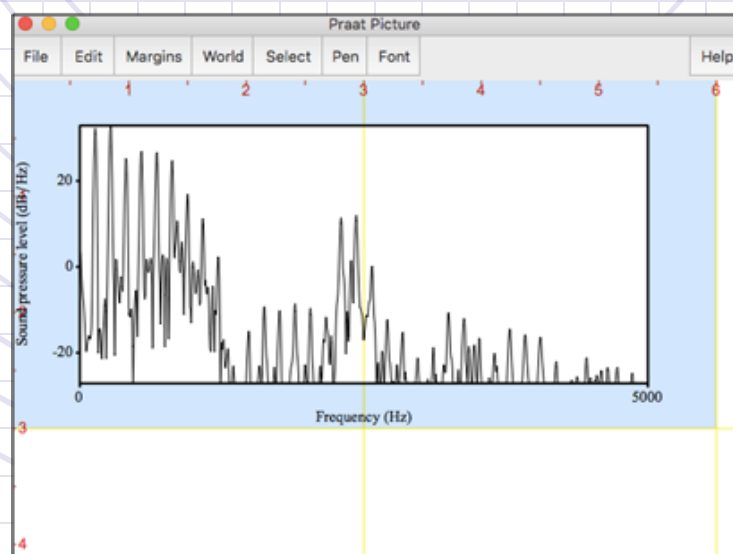


Imagen 32. Gráfica del espectro FFT en Praat Picture.

6. Análisis Espectral: Espectro LPC (Código de predicción lineal)

- Para obtener nuestro Espectro LPC, debemos dejar en primer lugar nuestro espectro FFT en la ventana Praat Picture, para que el espectro LPC se grafique por sobre los armónicos de nuestro espectro FFT.

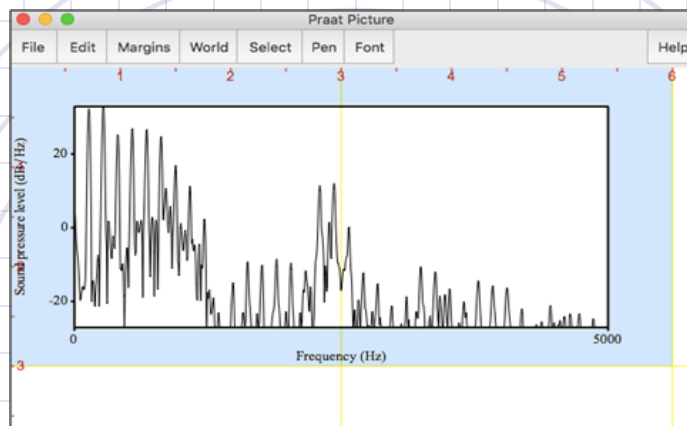


Imagen 33. Espectro FFT en Praat Picture

- Luego volvemos a nuestra ventana **Praat Objects** y con la misma muestra de /a/ sostenida, seleccionamos **Analyse spectrum** y presionamos la opción **to LPC** (autocorrelation).

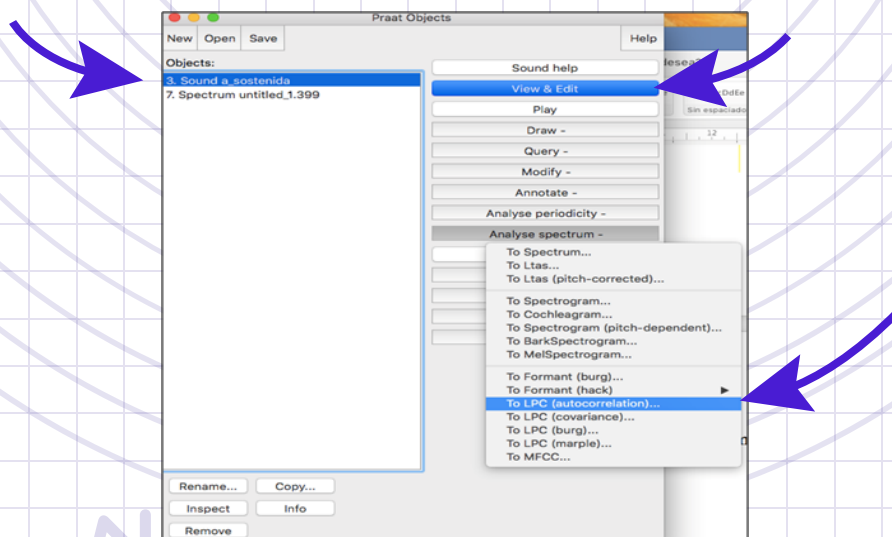


Imagen 34. Selección LPC autocorrelation

- Luego nos pide una configuración. En la opción **Prediction order** (número de coeficientes). Por defecto viene en 46 (dejar en 46 para frecuencia Niquist de 22.050 Hz).

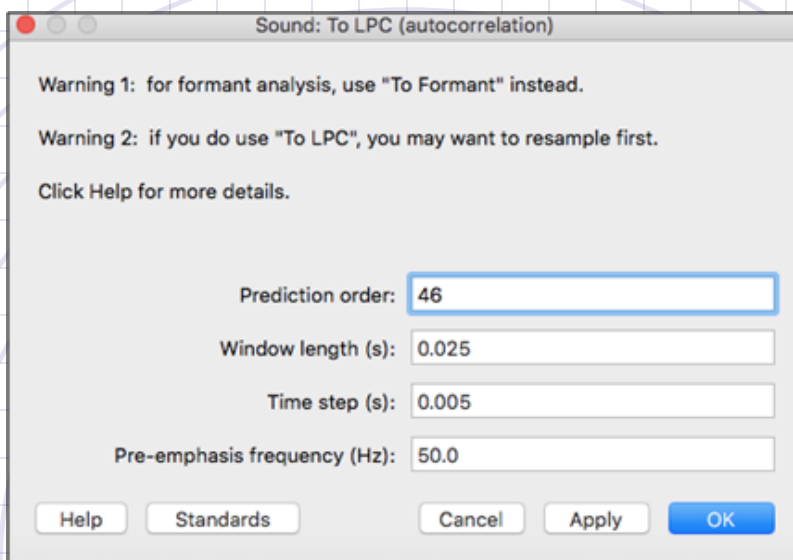


Imagen 35. Configuración LPC autocorrelation

- Presionamos **OK** y se generará un archivo LPC en nuestra ventana **Praat Objects**.

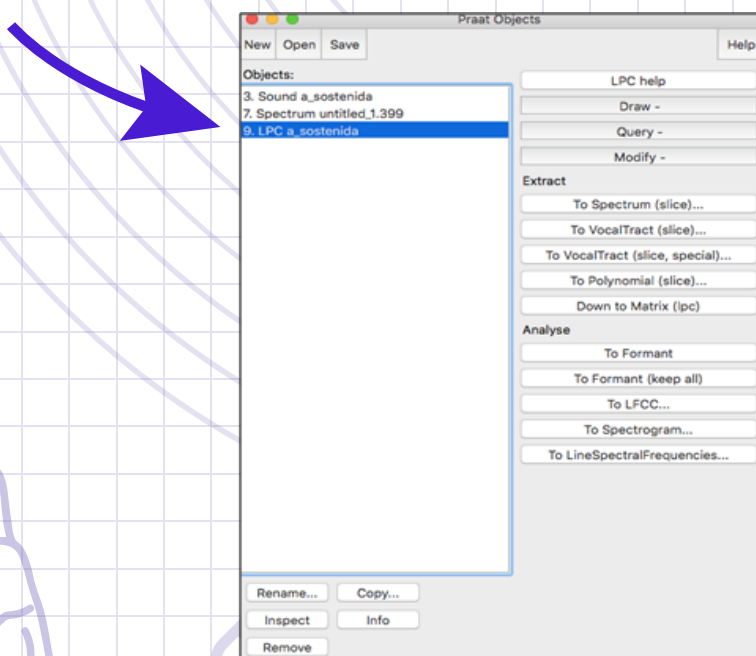


Imagen 36. Archivo LPC en ventana Praat Objects

- Luego para graficar seleccionamos **To Spectrum (slice)** y se abre una ventana de configuración en donde en la opción **Time (seconds)** debemos colocar el mismo punto donde obtuvimos nuestro espectro FFT.

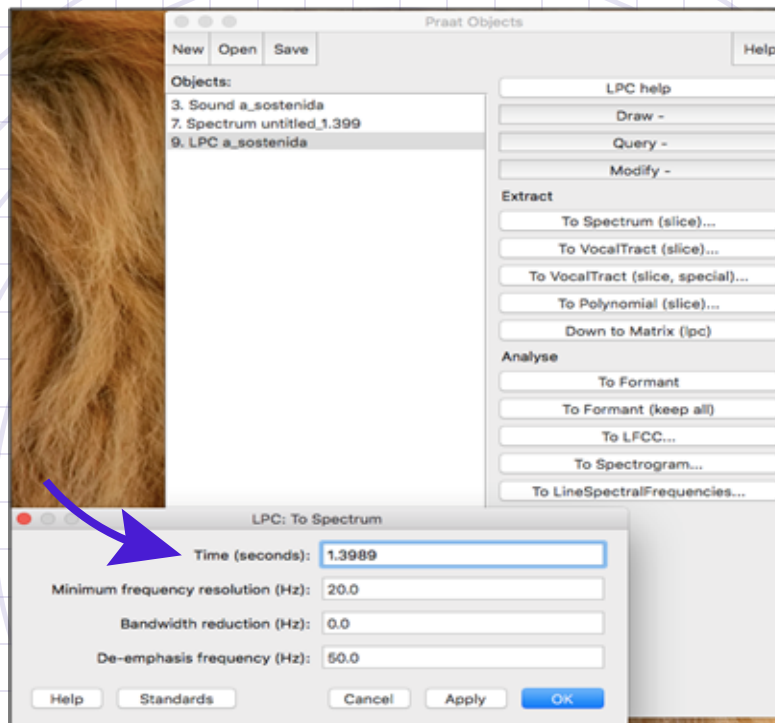


Imagen 37. Configuración del punto donde queremos obtener nuestro espectro LPC

Si no recuerda el punto exacto, vuelva al espectrograma de banda estrecha y lo anota. (aparece en rojo en la parte superior del oscilograma). Como sugerencia para que no vayamos a perderlo, es mejor que cuando este calculando el espectro FFT lo anote aparte, ya que como vemos luego se necesitará para calcular nuestro espectro LPC. Es sumamente importante que el tiempo sea el mismo para que de esa forma nuestra envolvente espectral (LPC) coincida al momento de graficarla con nuestro espectro FFT.

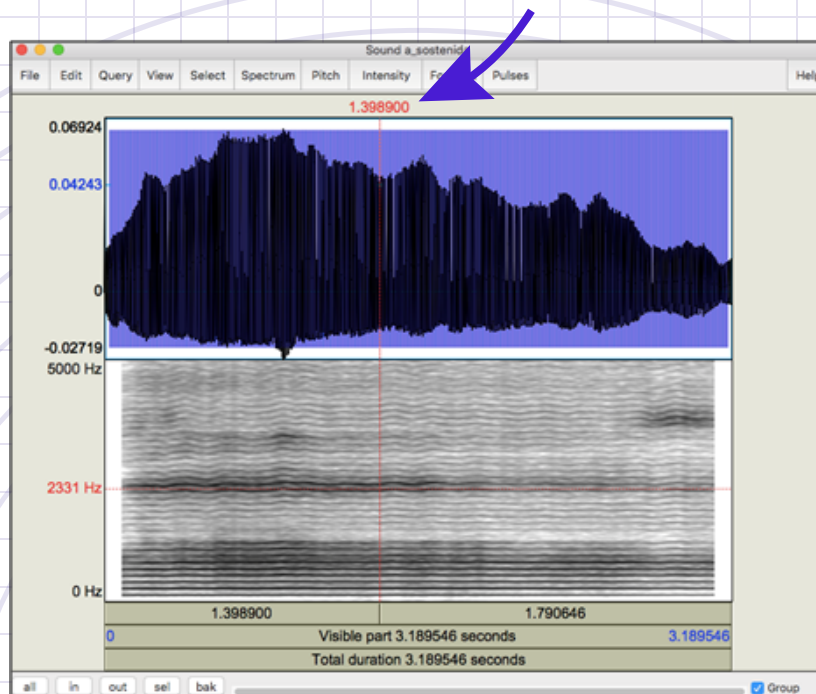


Imagen 38. Punto (tiempo) en donde se obtuvo el espectro FFT

- El resto de la configuración la dejamos por defecto, presionamos **OK** y se generará un nuevo archivo en nuestra ventana **Praat Objects**.

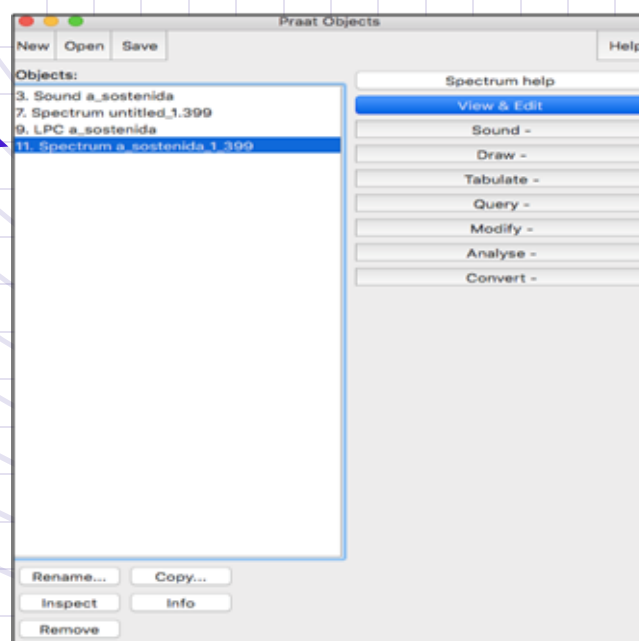


Imagen 39. Espectro LPC en ventana Praat Objects.

- Por último, para dibujarlo en la ventana **Praat Picture**, con este nuevo archivo,
- seleccionamos **Draw**, se abre una ventana de configuración la cual dejamos tal cual y presionamos **OK**.

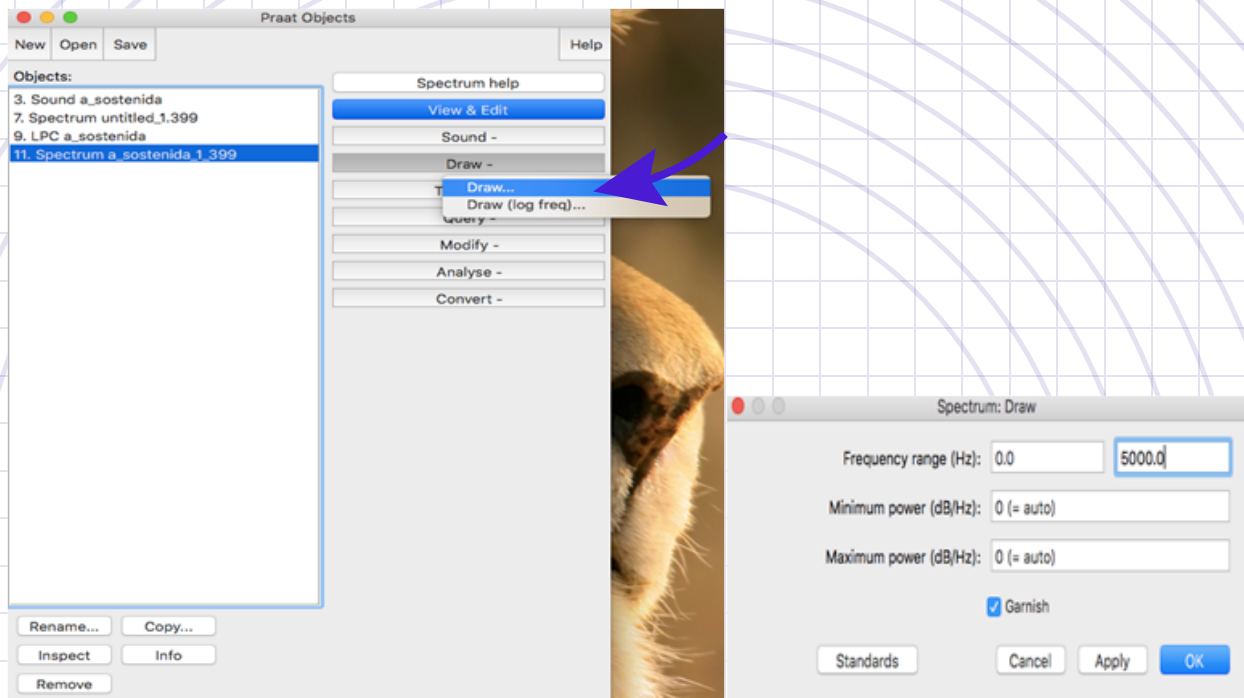


Imagen 40. Como graficar espectro LPC en Praat Picture

- Por último, para dibujarlo en la ventana **Praat Picture**, con este nuevo archivo,
- seleccionamos **Draw**, se abre una ventana de configuración la cual dejamos tal cual y presionamos **OK**.

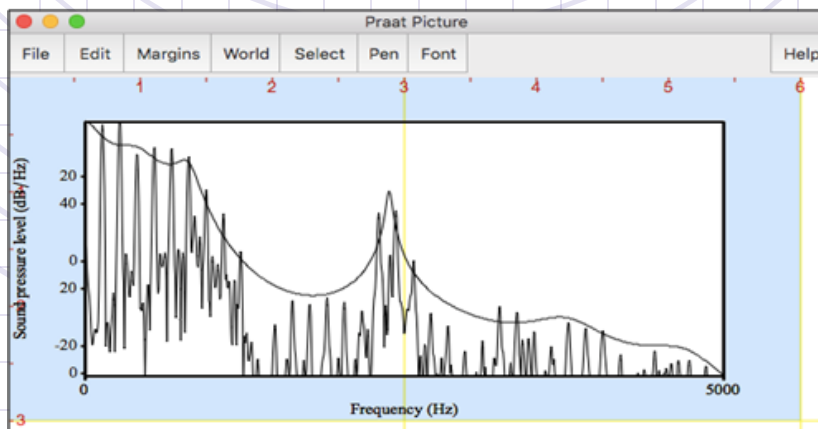


Imagen 41. Envolvente espectral graficada en sobre espectro FFT en Praat Picture.

7. Análisis Espectral: Cepstrum

- Para calcular el **Cepstrum** debemos tener un archivo de /a/ sostenida (o vocalización sostenida) en la ventana **Praat Objects**.
- Se selecciona la opción **Analyse spectrum**, luego **To Spectrum** y se desplegará un pequeño cuadro **Sound: To Spectrum**, en donde sólo presionamos **OK**.

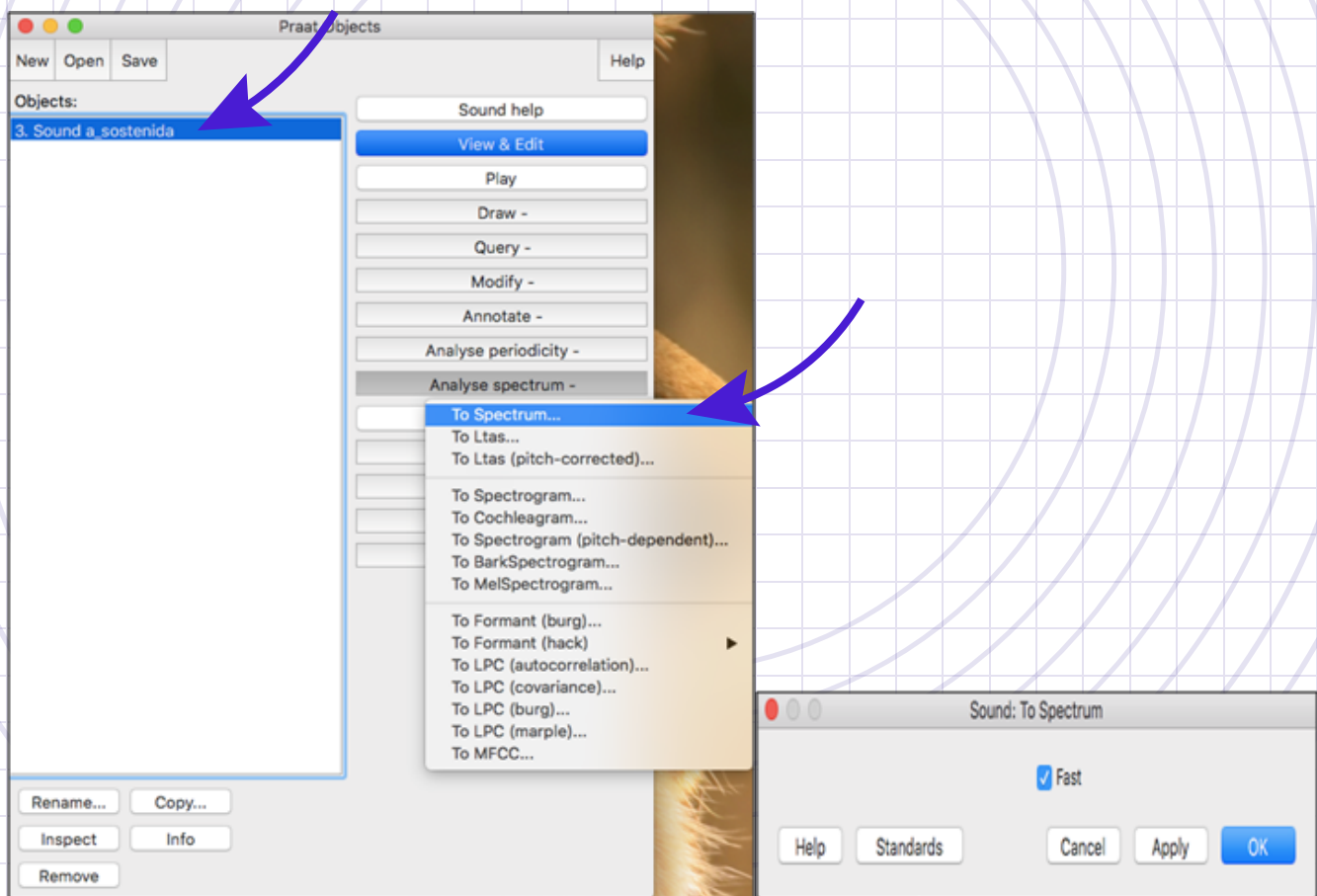


Imagen 42. Generando el primer espectro para calcular el Cepstrum

- Posterior a esto, se generará el primer espectro en la ventana de objetos.

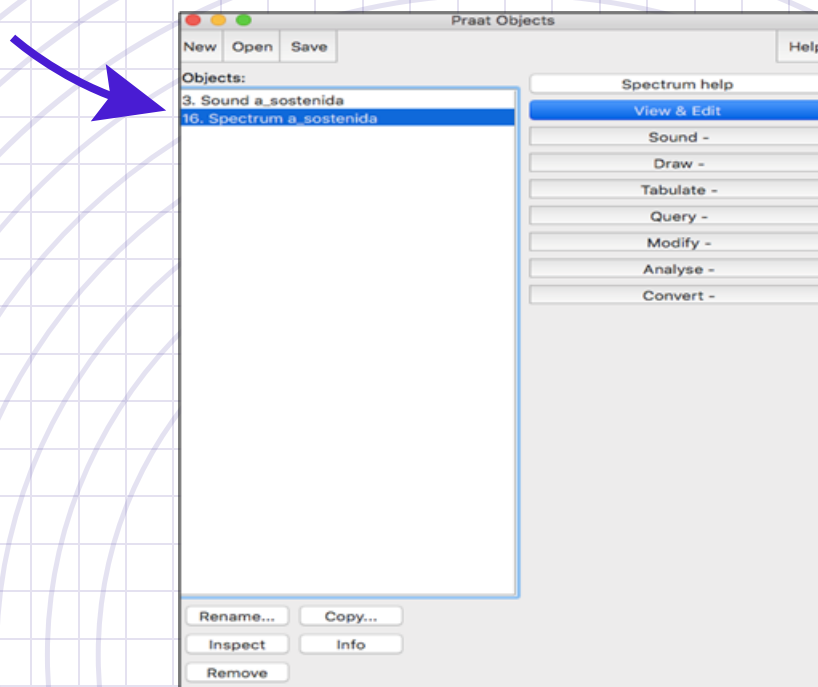


Imagen 43. Primer espectro en ventana Praat Objects.

- Ahora con ese **spectrum** nos vamos a **Analyse** y luego a **To Power Cepstrum**.

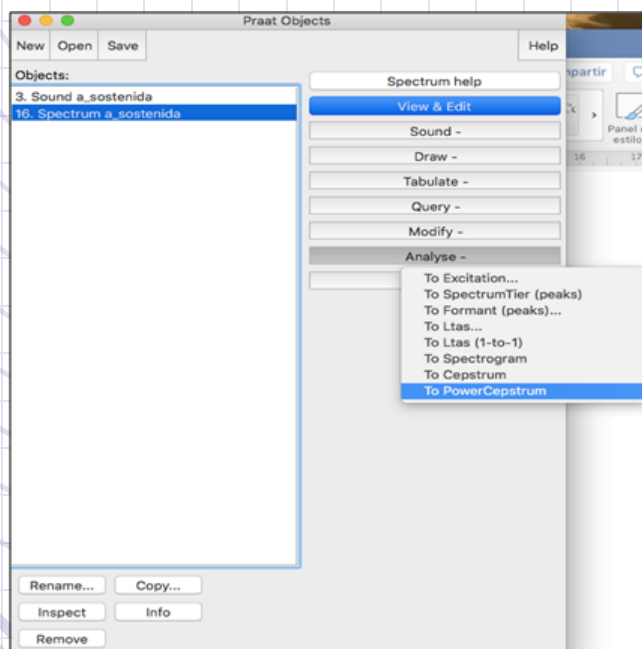


Imagen 44. Generando Cepstrum en ventana Praat Objects

- Ahora con ese **spectrum** nos vamos a **Analyse** y luego a **To Power Cepstrum**.

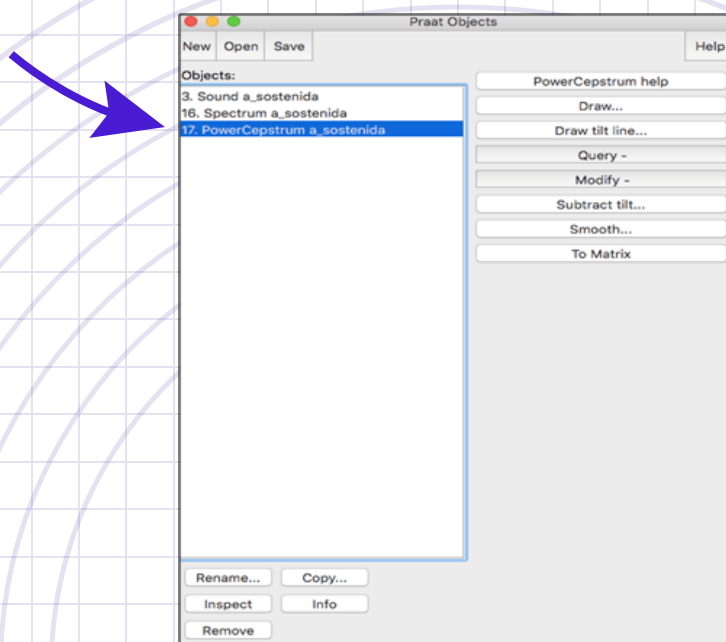


Imagen 45. Power Cepstrum en ventana Praat Objects

- Luego con ese **Power Cepstrum** generado, seleccionamos **Query**, a continuación
- **Get Peak Prominence**. Se abre una ventana de configuración la cual dejamos tal cual y presionamos **OK**.

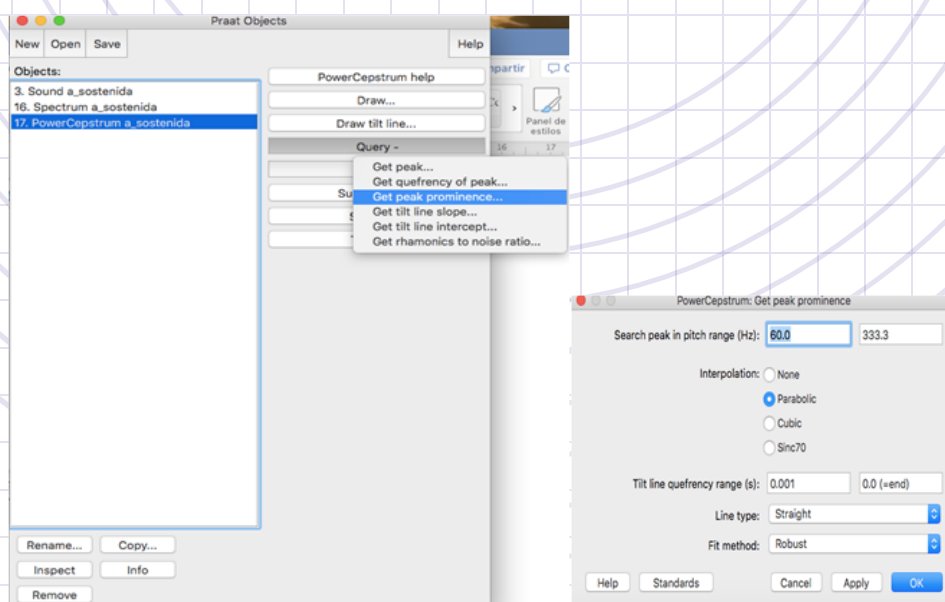


Imagen 46. Calculando el valor del pico cepstral

Luego de presionado **OK**, el programa nos entregará el valor del pico **cepstral** en decibeles y la quefrecy.

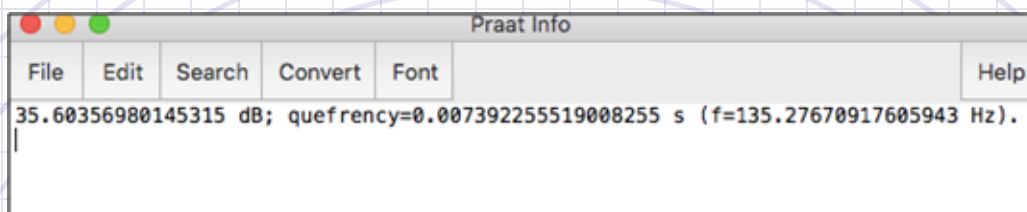


Imagen 47. Valor pico Cepstral en decibeles.

Para graficar en la ventana **Praat Picture**, seleccionamos **Draw**. Luego nos solicita una configuración en la cual dejamos el **Quefrecy rangeen** máximo 0.015 (aunque se puede ir variando el valor para ver los cambios).

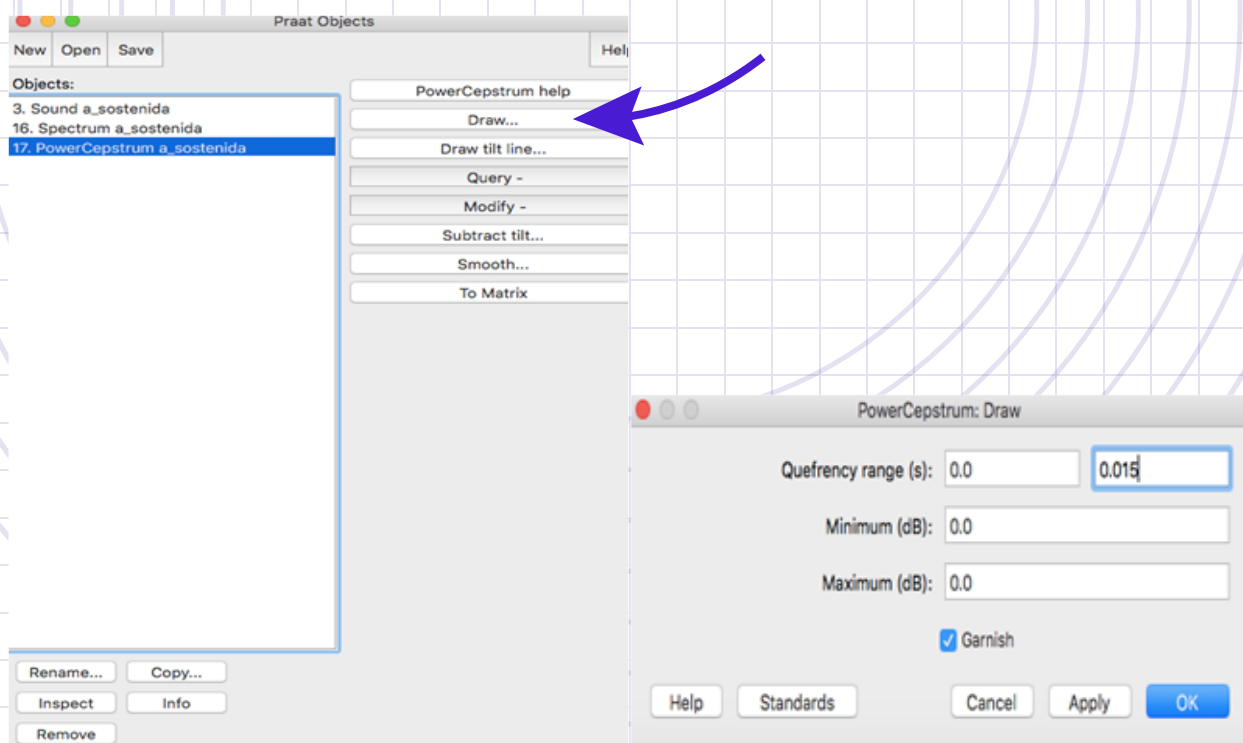


Imagen 48. Configuración para graficar pico Cepstral

- Luego presionamos **OK** y se obtiene el grafico en la ventana **Praat Picture**. Seleccionamos **File** y guardamos según el formato deseado.

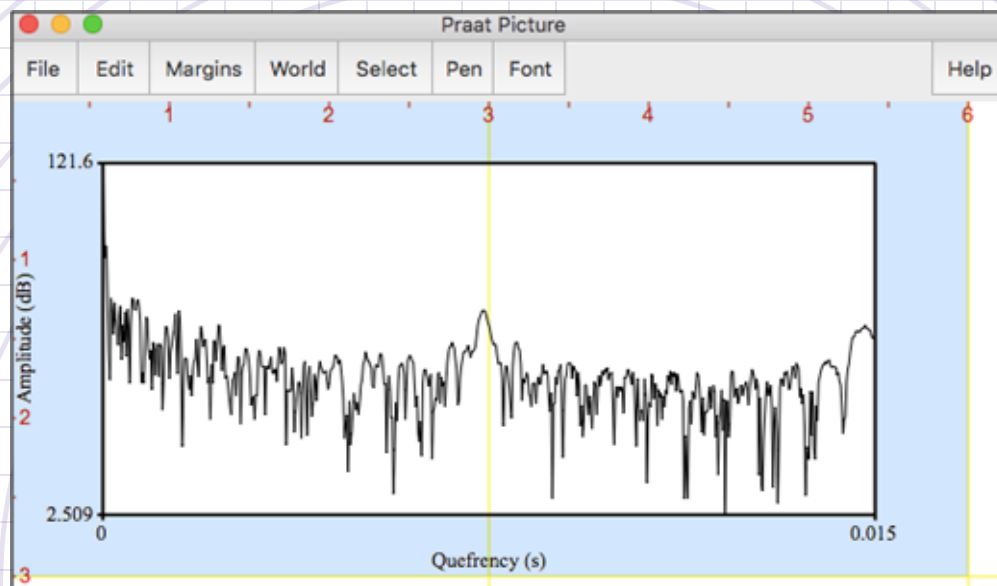


Imagen 49. Gráfico de pico cepstral en Praat Picture

8. Análisis de la intensidad

Para analizar la intensidad debemos ir a la pestaña **Intensity** y se despliegan varias opciones. Seleccionamos **Show Intensity** y se grafica la intensidad en el espectrograma como una línea horizontal de color amarillo. Debemos tener siempre graficada la intensidad para cualquier otro análisis.

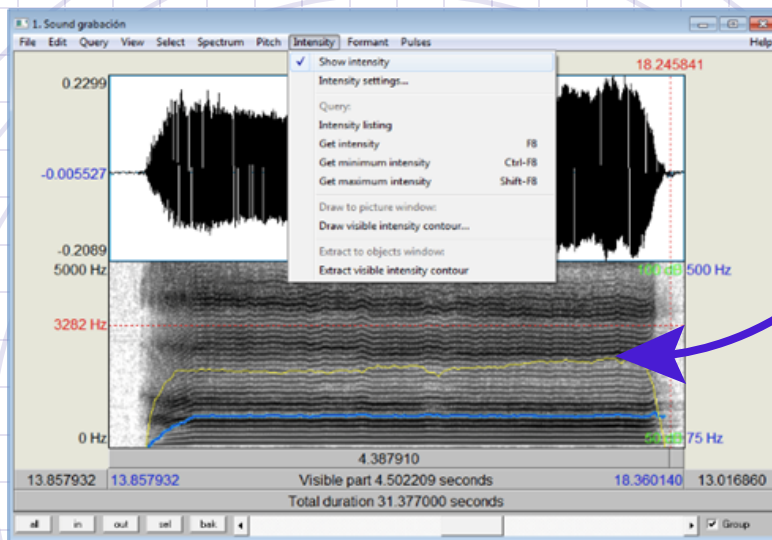


Imagen 50. Gráfica de la intensidad (color amarillo) en el espectrograma.

Para analizar la intensidad debemos ir a la pestaña **Intensity** y se despliegan varias opciones. Seleccionamos **Show Intensity** y se grafica la intensidad en el espectrograma como una línea horizontal de color amarillo. Debemos tener siempre graficada la intensidad para cualquier otro análisis.

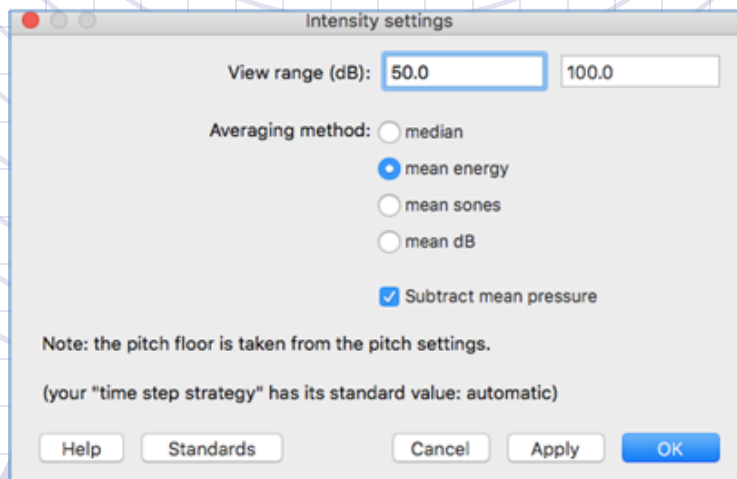


Imagen 51. Configuración del rango de intensidad a analizar

En la opción **Intensity listing**, si seleccionamos un segmento de la grabación, nos entregará un listado de todas las intensidades que hizo la persona en ese segmento de la emisión.

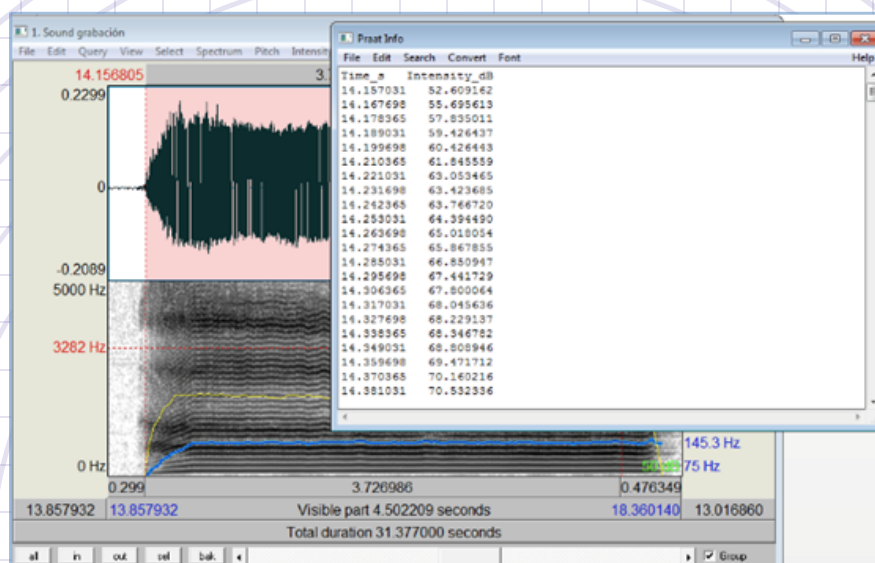


Imagen 52. Listado de intensidades de un segmento de la grabación.

En la opción **Get intensity** nos entrega un promedio de la intensidad de un segmento seleccionado. Se debe tener cuidado con no seleccionar el ataque ni la filatura para evitar errores.

La opción **Get minimum intensity** nos entrega el valor mínimo de la intensidad de un segmento seleccionado de la grabación.

La opción **Get maximum intensity** nos entrega el valor máximo de la intensidad de un segmento seleccionado de la grabación.

La opción **Draw visible intensity contour** nos permite graficar el contorno de la intensidad en nuestra ventana **Praat Picture**.

9. Análisis de la frecuencia

Para analizar la intensidad debemos ir a la pestaña **Pitch** y se despliegan varias opciones. Seleccionamos **Show Pitch** y se grafica la frecuencia en el espectrograma como una línea horizontal de color azul. Debemos tener siempre graficada la frecuencia para cualquier otro análisis.

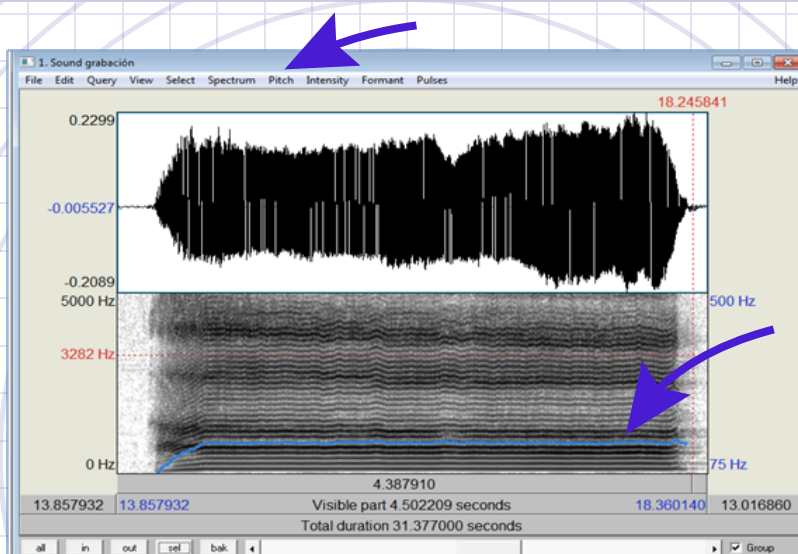


Imagen 53. Gráfica de la frecuencia (color azul) en el espectrograma

En la opción **Pitch settings**, luego **Pitch range (Hz)** podemos configurar el rango de frecuencias a analizar el cual se encuentra en Hertz. Por defecto el programa lo trae entre 75 y 500 Hz, pero si queremos analizar voces más graves o agudas, podemos modificarlo.

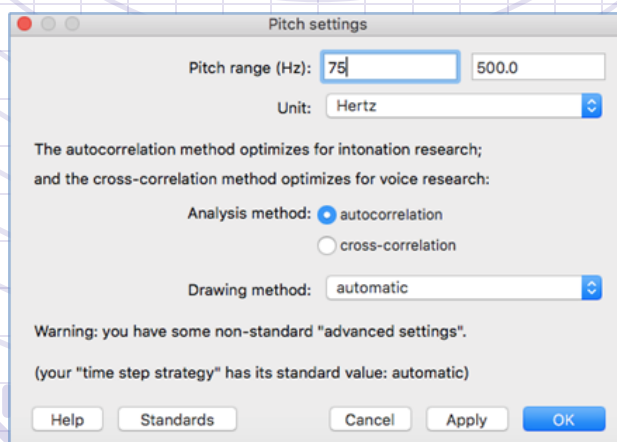


Imagen 54. Configuración del rango de frecuencias a analizar

En la opción **Pitch listing**, si seleccionamos un segmento de la grabación, nos entregará un listado de todas las frecuencias que hizo la persona en ese segmento de la emisión.

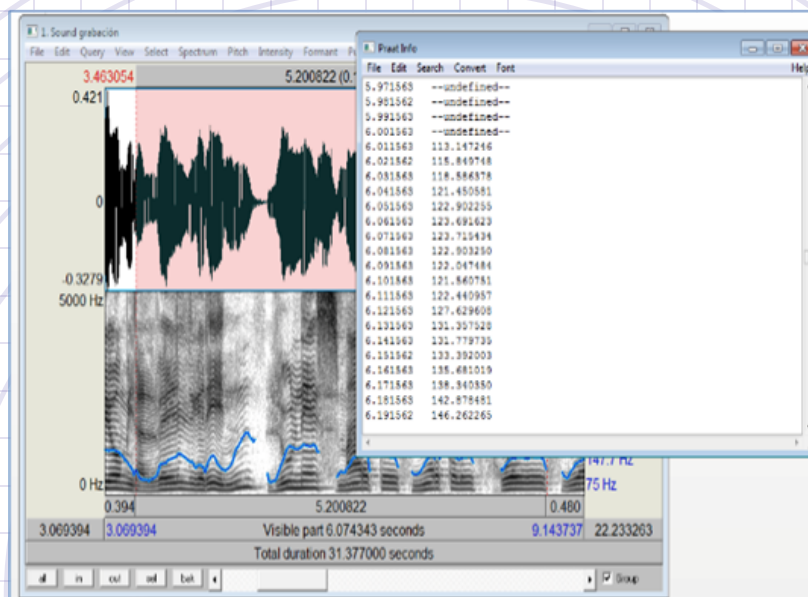


Imagen 55. Listado de frecuencias de un segmento de la grabación

- En la opción **Get pitch** nos entrega un promedio de la frecuencia de un segmento seleccionado. Esto serviría para calcular el tono medio hablado.
- La opción **Get minimum pitch** nos entrega el valor mínimo de la frecuencia de un segmento seleccionado de la grabación.
- La opción **Get maximum pitch** y nos entrega el valor máximo de la frecuencia de un segmento seleccionado de la grabación.
- La opción **Draw visible pitch contour** nos permite graficar el contorno de la frecuencia en nuestra ventana Praat Picture.

10. Medidas de perturbación de la voz

Para obtener los parámetros de perturbación de la voz, debemos ir a la opción **Pulses** y luego marcar la opción **Show pulses**. Una vez realizado esto podemos

- apreciar que en el oscilograma se despliegan barras azules seguidas una tras otra, que representan los pulsos glóticos. Es muy importante dejar marcada esta opción, ya que de lo contrario no podremos obtener los valores.

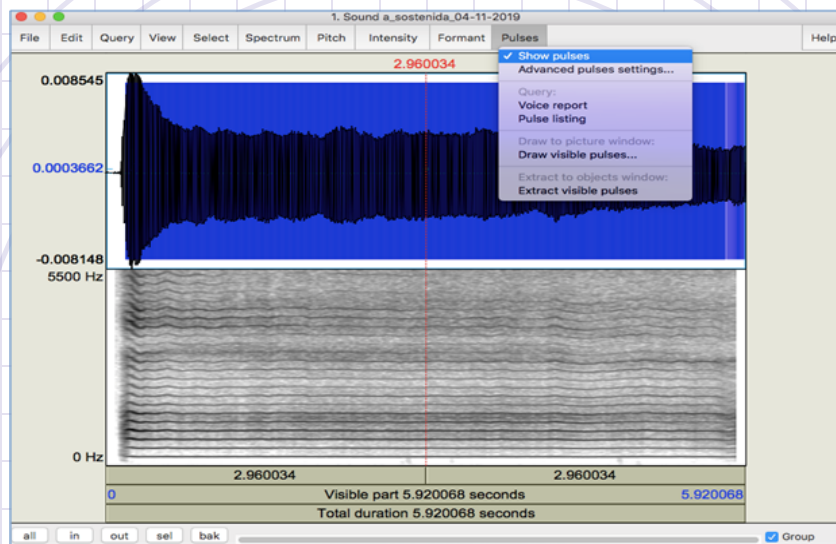


Imagen 56. Demarcación de los pulsos glóticos en el oscilograma

- Marcada ya la opción **Show pulses**, seleccionamos un segmento de la grabación y nos vamos a la opción **Voice report**.

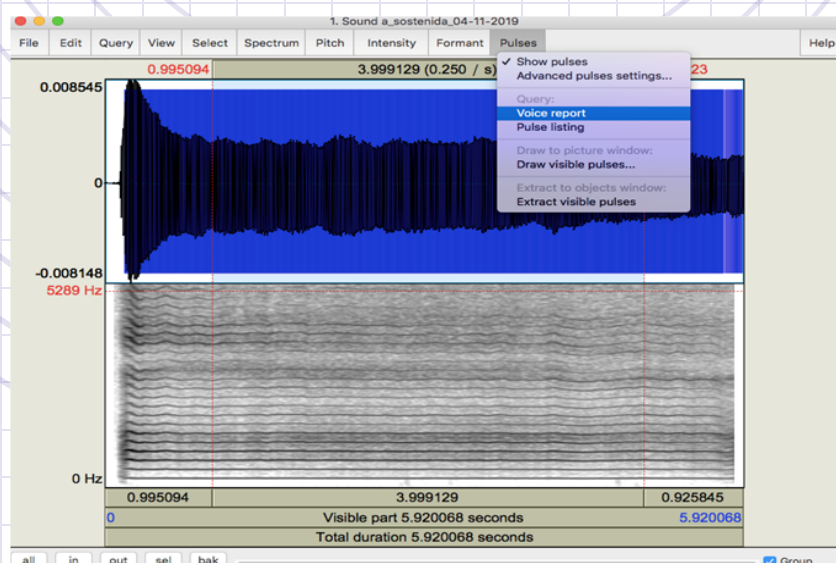
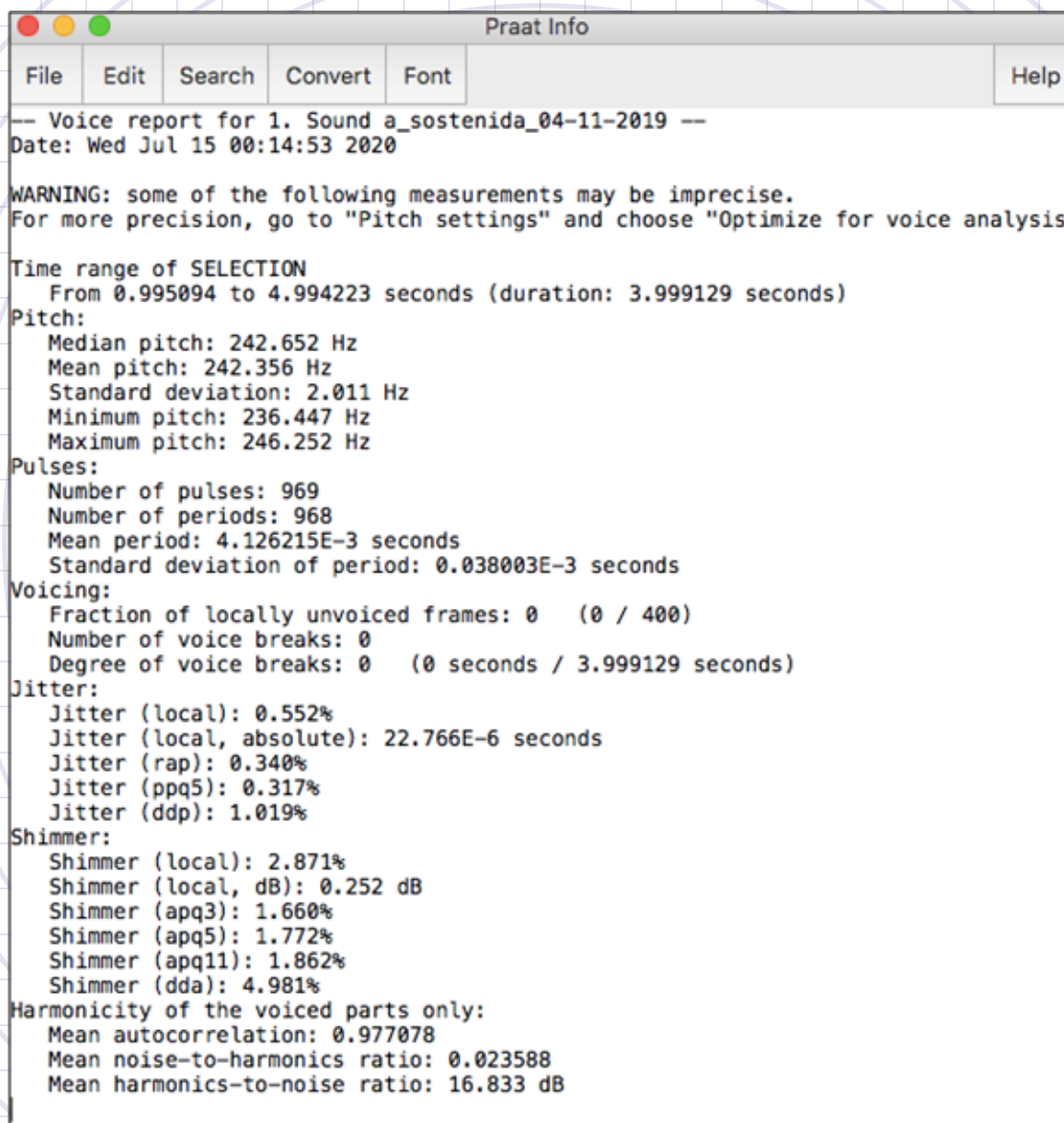


Imagen 57. Como obtener un reporte de voz

Seleccionada la opción **Voice report**, se desplegará una ventana en donde se muestra múltiples parámetros entre ellos el **Jitter**, **Shimmer** e **Índice armónico** versus ruido, los cuales podemos comparar con los valores normativos existentes.



```
Praat Info
File Edit Search Convert Font Help
-- Voice report for 1. Sound a_sostenida_04-11-2019 --
Date: Wed Jul 15 00:14:53 2020

WARNING: some of the following measurements may be imprecise.
For more precision, go to "Pitch settings" and choose "Optimize for voice analysis"

Time range of SELECTION
  From 0.995094 to 4.994223 seconds (duration: 3.999129 seconds)
Pitch:
  Median pitch: 242.652 Hz
  Mean pitch: 242.356 Hz
  Standard deviation: 2.011 Hz
  Minimum pitch: 236.447 Hz
  Maximum pitch: 246.252 Hz
Pulses:
  Number of pulses: 969
  Number of periods: 968
  Mean period: 4.126215E-3 seconds
  Standard deviation of period: 0.038003E-3 seconds
Voicing:
  Fraction of locally unvoiced frames: 0 (0 / 400)
  Number of voice breaks: 0
  Degree of voice breaks: 0 (0 seconds / 3.999129 seconds)
Jitter:
  Jitter (local): 0.552%
  Jitter (local, absolute): 22.766E-6 seconds
  Jitter (rap): 0.340%
  Jitter (ppq5): 0.317%
  Jitter (ddp): 1.019%
Shimmer:
  Shimmer (local): 2.871%
  Shimmer (local, dB): 0.252 dB
  Shimmer (apq3): 1.660%
  Shimmer (apq5): 1.772%
  Shimmer (apq11): 1.862%
  Shimmer (dda): 4.981%
Harmonicity of the voiced parts only:
  Mean autocorrelation: 0.977078
  Mean noise-to-harmonics ratio: 0.023588
  Mean harmonics-to-noise ratio: 16.833 dB
```

Imagen 58. Reporte de voz

11. Espectro promedio a largo plazo: Long-Term-Average Spectrum (LTAS)

Gráfico LTAS

- En primer lugar debemos tener un archivo de audio grabado correspondiente a habla conectada (texto o canto si es voz cantada lo que se desea analizar).
- Para investigación se sugiere un texto fonéticamente balanceado y si es voz cantada debiese ser la misma canción para todos.
- Para comenzar el análisis debe estar el archivo de la grabación en la ventana **Praat Object**, nos vamos a **Analyse spectrum** y seleccionamos **To Ltas (Pitch corrected)**.

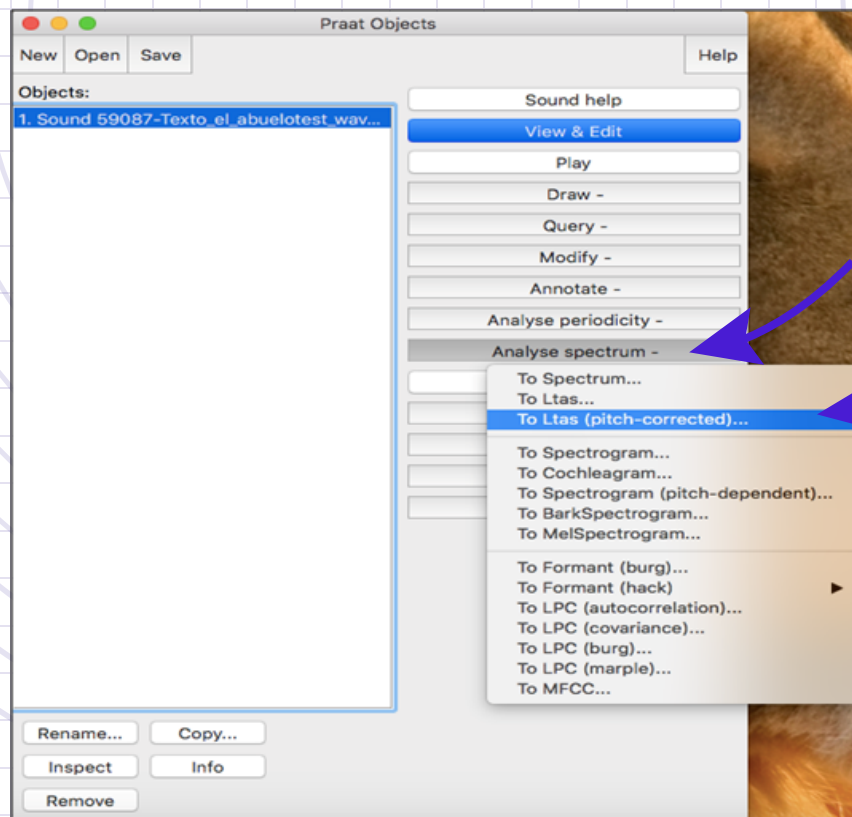


Imagen 59. Selección Ltas (Pitch-corrected)

Luego se abre una ventana para calibrar **Sound: To Ltas (Pitch-corrected)**. En **Minimum pitch** se deja 75 Hz (nunca 0, para eliminar ruidos de baja frecuencia). En **Maximum pitch** se deja 600 Hz lo cual es adecuado para voz hablada, si es voz cantada podemos subirlo. En **Maximum frequency** dejamos 5000 Hz si se trata de la voz de un hombre o 5500 Hz si se trata de la voz de una mujer. El resto se deja tal cual.

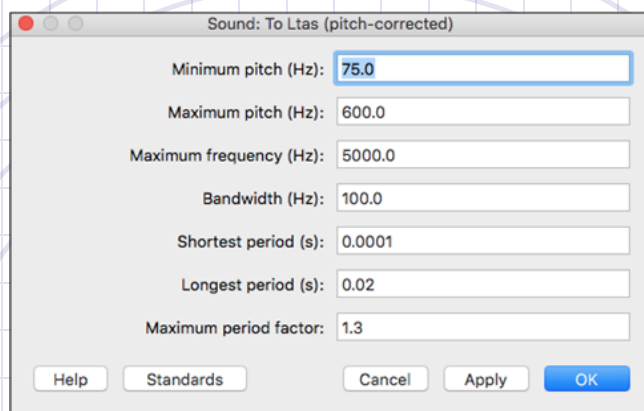


Imagen 60. Configuración Sound To Ltas (pitch-corrected)

Luego de apretar **OK** el software eliminó todos los sonidos afonos y segmentos sin voz, quedando esta nueva señal la ventana **Praat Objects**.

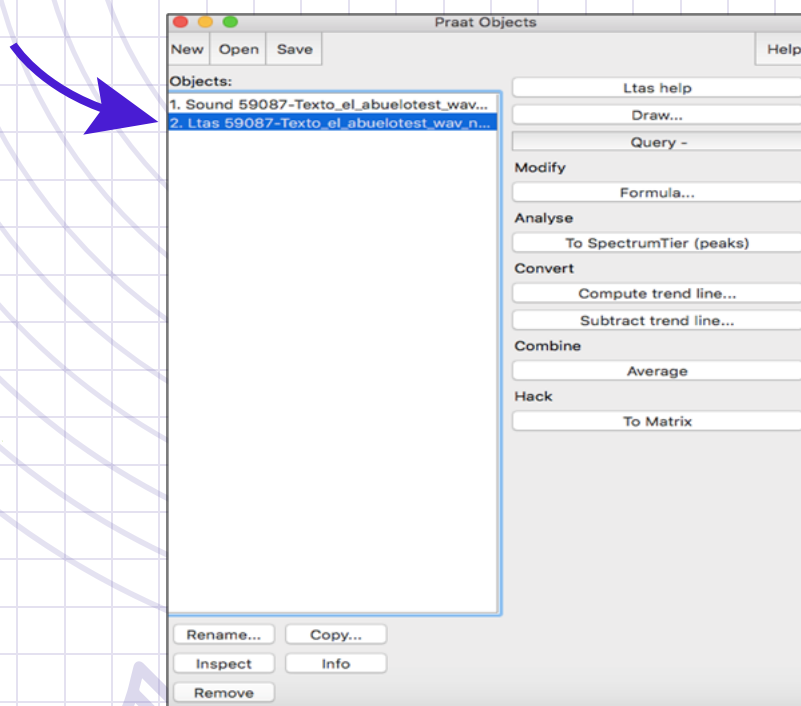


Imagen 61. Archivo creado mediante Ltas (pitch-corrected)

Para graficar, con el archivo seleccionado vamos a la opción **Draw** y se abre una ventana llamada **Ltas: Draw**, la cual debemos configurar. La **frequency range** se deja entre 0 y 5000 Hz. **Power range** se deja tal cual la configuración que viene dada. Se selecciona la opción **Garnishy** en **Drawing method** seleccionamos si queremos gráfico de curvas, barras, etc.

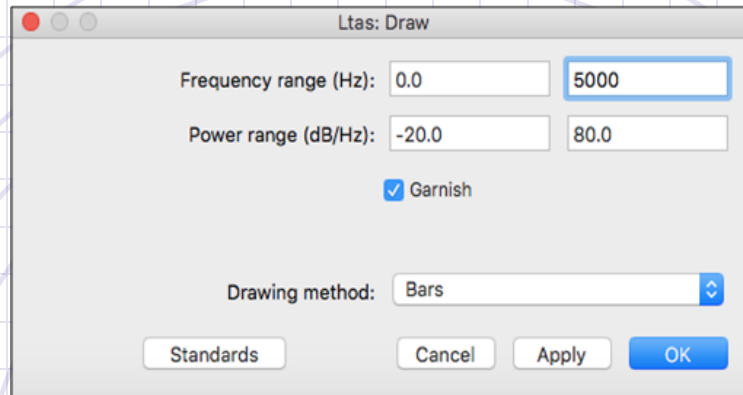


Imagen 62. Configuración para graficar LTAS

Se presiona **OK** y se genera el gráfico en la ventana **Praat Picture**. El gráfico nos muestra las bandas frecuenciales (bin) hasta los 5000 Hz. Nos muestra el nivel de energía en dB que tiene cada uno y las mayores concentraciones de energía (formantes) y orientarnos respecto del color de la voz del usuario que se está analizando.

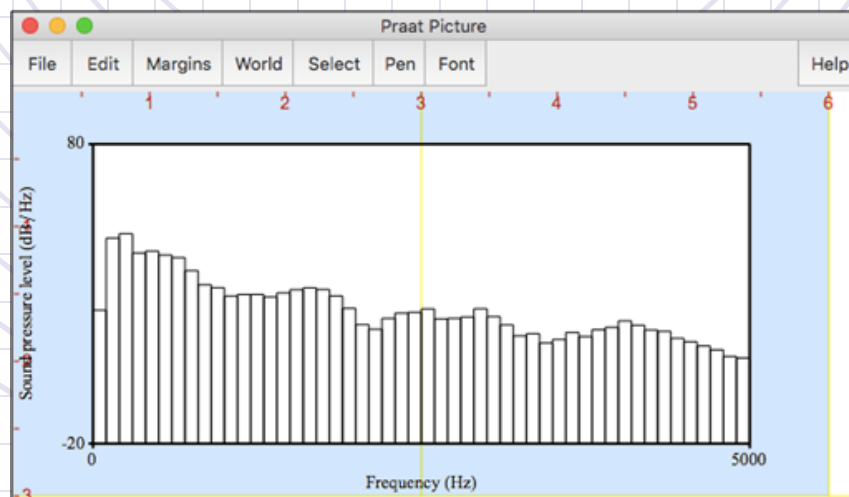


Imagen 63. Gráfico Ltas en Praat Picture

- Obtenido el gráfico podemos hacer un análisis descriptivo y luego obtener los valores numéricos.
- Para graficar la **Spectral Tilt** (pendiente espectral) la cual muestra la inclinación de la distribución de energía a lo largo del rango **frecuencial**, en el gráfico obtenido en **Praat Picture** seleccionamos **Peny** elegimos otro color.

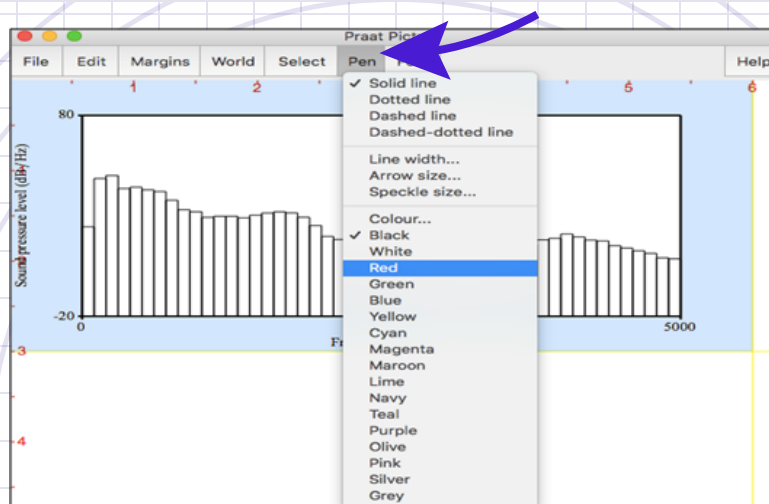


Imagen 64. Selección de color para dibujar la Spectral Tilt

- Luego se vuelve a la ventana **Praat Objects** y seleccionamos **Compute trend line**. Se abre una pequeña ventana de configuración en la cual mantenemos los valores formativos.

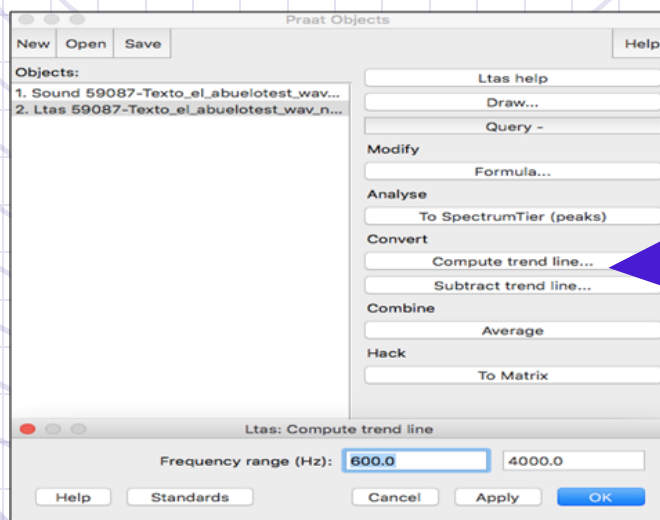


Imagen 65. Configuración para generar archivo Spectral Tilt

- Presionamos **OK** y se generará un nuevo archivo llamado **Ltas__trend** en la ventana **Praat Objetcs**. Lo seleccionamos y vamos a la opción **Draw**.

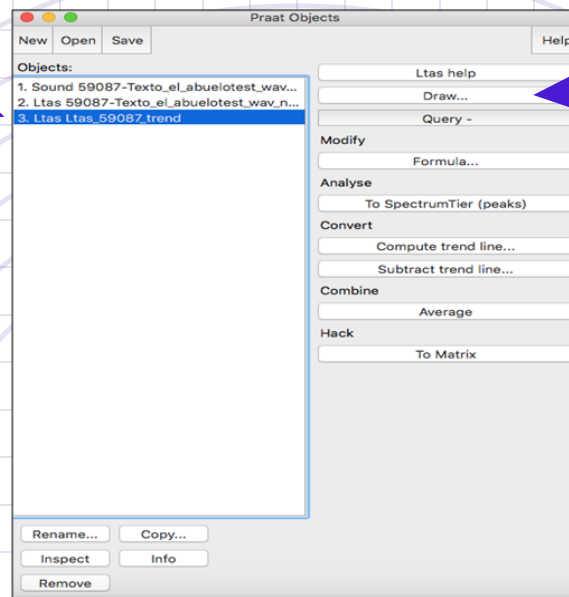


Imagen 66. Archivo generado para graficar la Spectral Tilt

- Se abre una ventana de configuración la cual dejamos tal cual.

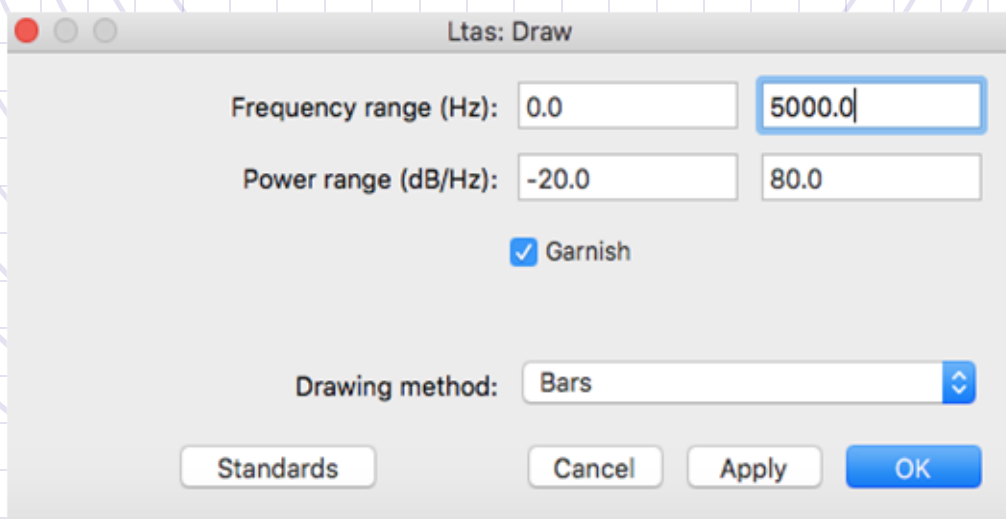


Imagen 66. Archivo generado para graficar la Spectral Tilt

Presionamos **OK** y se abrirá nuevamente nuestro **gráfico Ltas** en la ventana **Praat** **Picture**, pero esta vez con la línea de inclinación de la pendiente espectral (Spectral tilt).

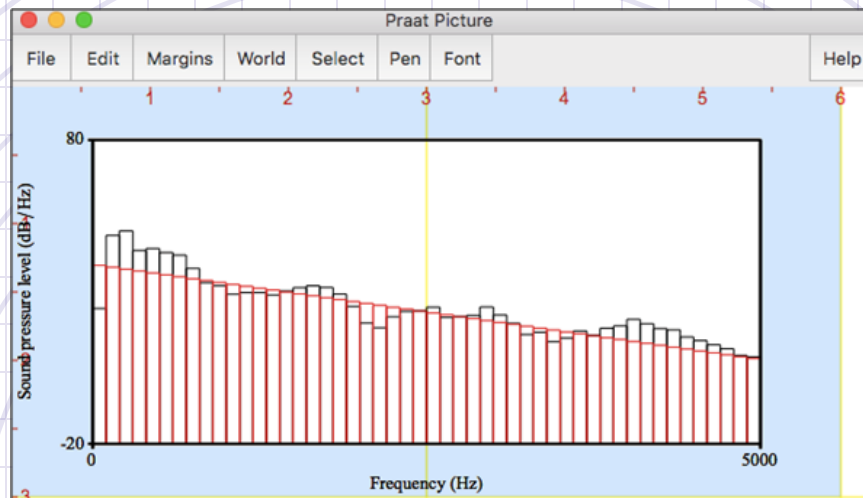


Imagen 68. Pendiente Espectral demarcada en gráfico.

12. Obtención de valores numéricos

Para obtener los valores numéricos, con el archivo Ltas creado en nuestra ventana **Praat Objects**, vamos **Query** y luego la opción **Get slope** y se abrirá una ventana para configurar (la configuración variará según la medida a largo plazo que queremos obtener) llamada **Ltas: Get slope**.

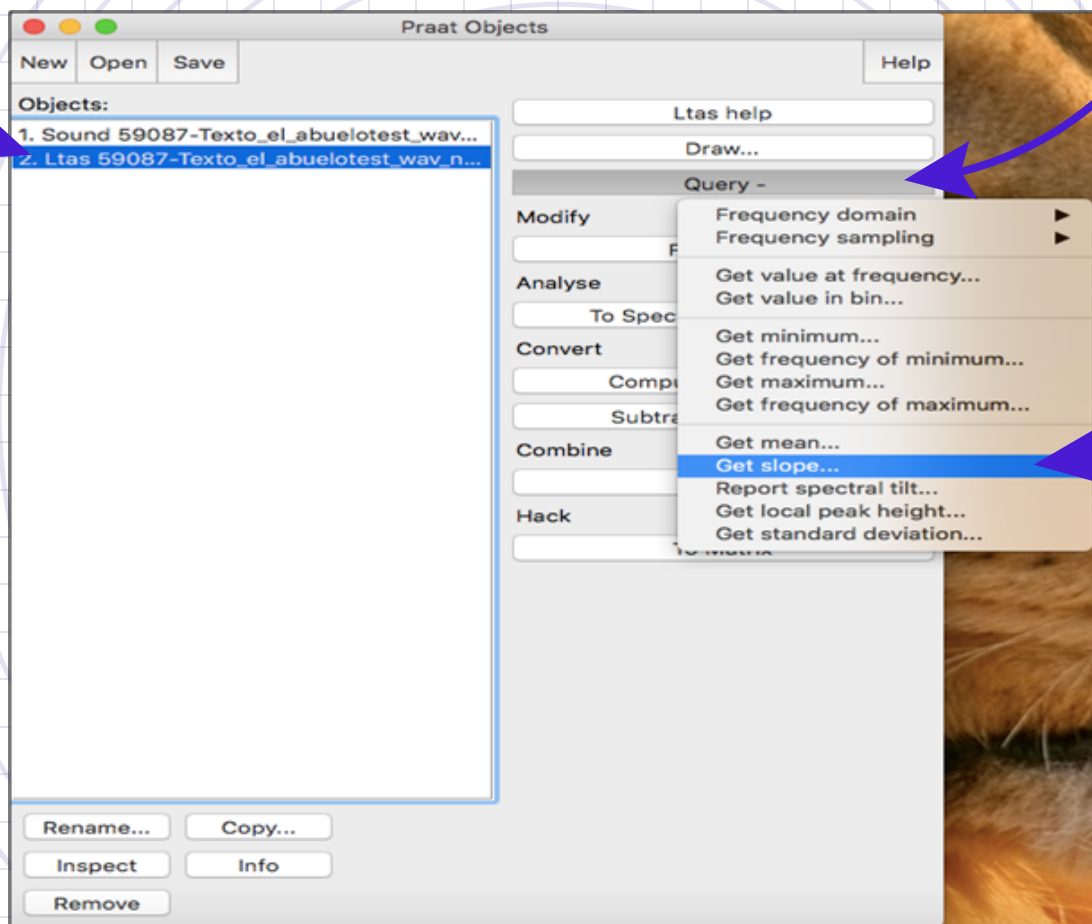


Imagen 69. Selección Get slope

Alpha Ratio

- Con la ventana de configuración **Ltas: Get slope** generada, debemos colocar en
 - **Low band (Hz)** 50-1000. Luego en **High band (Hz)** 1000-5000. Por último, en la opción **Averaging method** se selecciona **energy**.

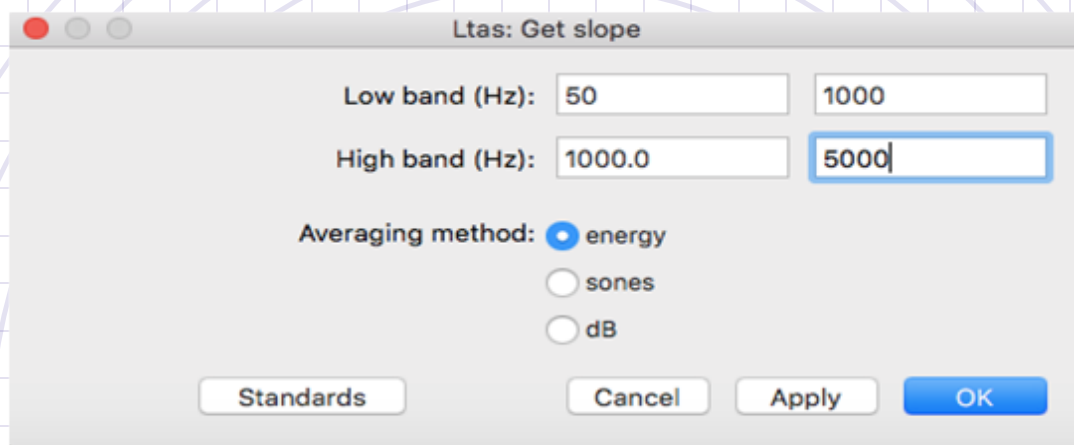


Imagen 70. Configuración Alpha Ratio

- Luego presionamos **OK** y el programa nos entrega el valor del Alpha Ratio en dB.

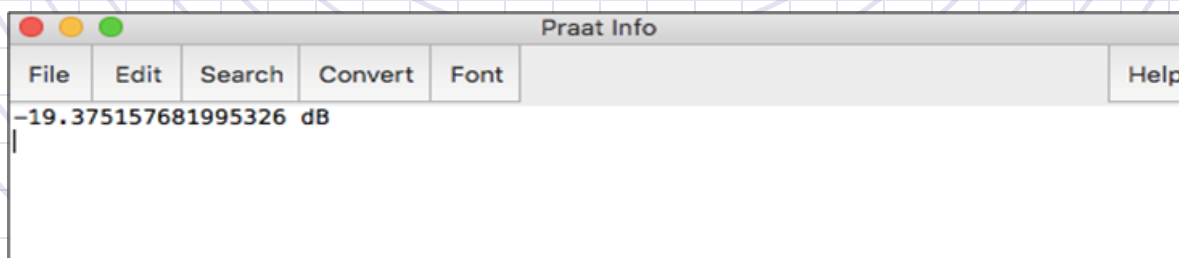


Imagen 71. Valor Alpha Ratio.

L1 - L0

- Con la ventana de configuración **Ltas: Get slope** generada, debemos colocar en
- **Low band (Hz)** 50-300. Luego en **High band (Hz)** 300-800. Por último, en la opción **Averaging method** se selecciona **energy**.

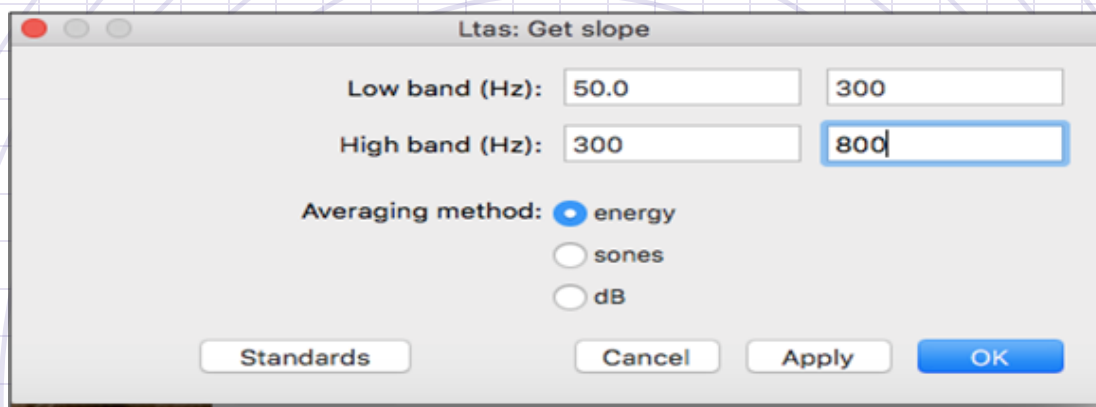


Imagen 72. Configuración L1-L0

- Luego presionamos **OK** y el programa nos entrega el valor de L1-L0 en dB.

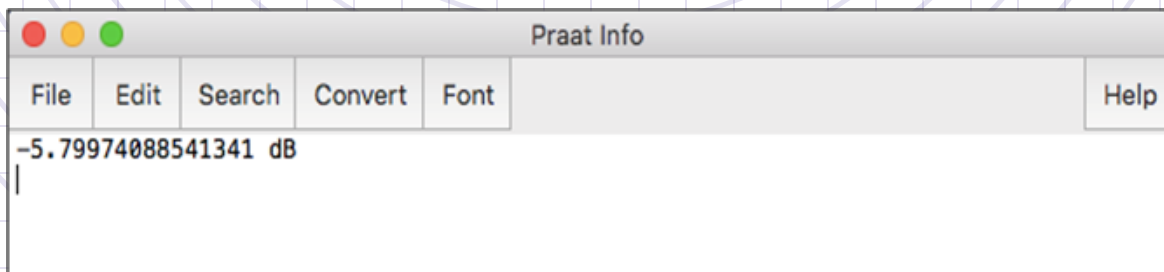


Imagen 73. Valor L1-L0

Singer Power Ratio (SPR)

Para obtener este valor, con el archivo **Ltas** en nuestra ventana **Praat** objects seleccionamos la opción **Query** y luego **Get maximum**.

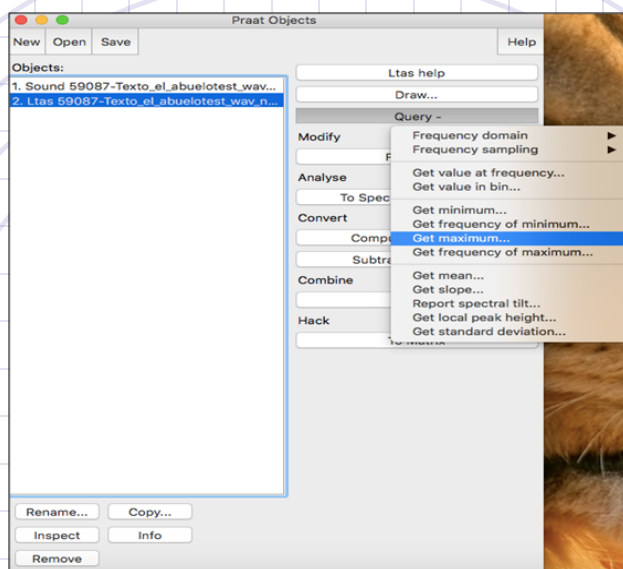


Imagen 74. Selección Get maximum

Luego se abre una ventana de configuración. Para obtener el primer peak colocamos en **From frequency (Hz)**: 0 y en **To frequency (Hz)**: 2000. En la opción **Interpolation** seleccionamos **Cubic**.

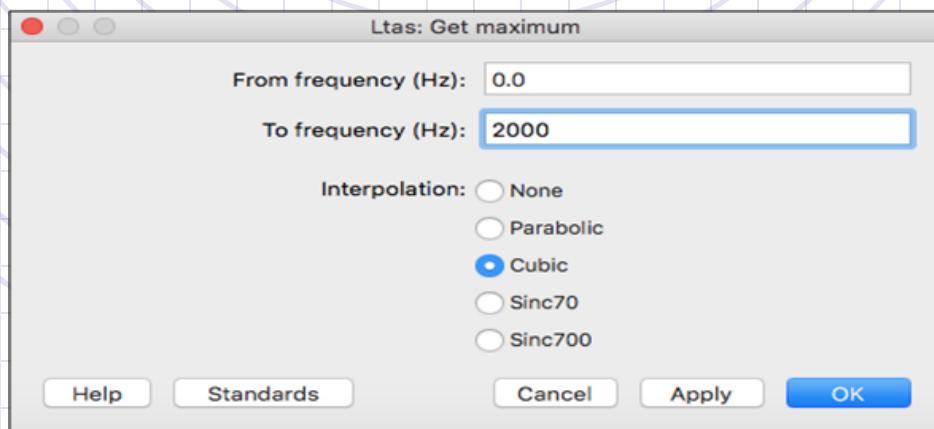


Imagen 75. Configuración para obtener valor del primer peak

- Presionamos **OK** y se obtiene el valor del **primer peak**, el cual lo anotamos.

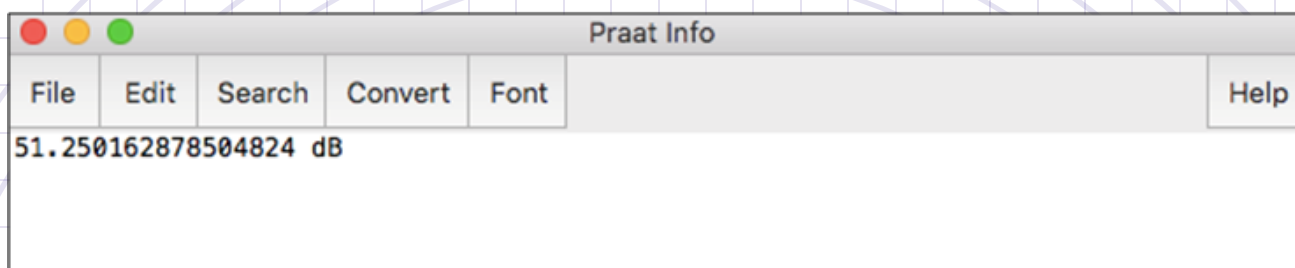


Imagen 76. Valor del primer peak

- Para obtener el valor del **segundo peak** hacemos lo mismo, pero en la ventana de configuración colocamos en **From frequency (Hz): 2000** y en **To frequency (Hz): 4000**.
-

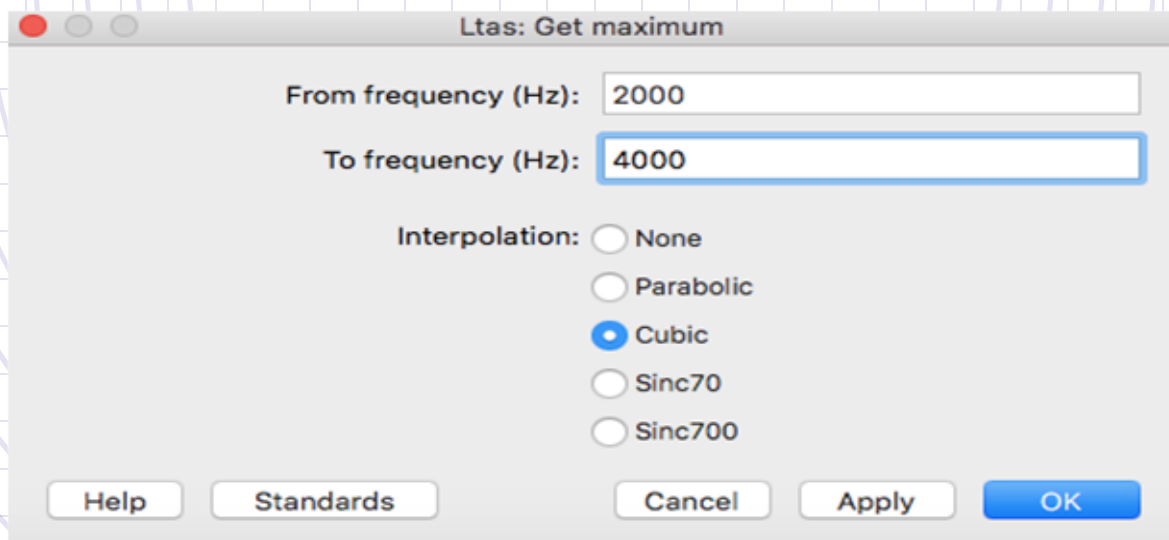


Imagen 77. Configuración para obtener el valor del segundo peak

- Luego presionamos **OK** y obtenemos el valor del **segundo peak**.

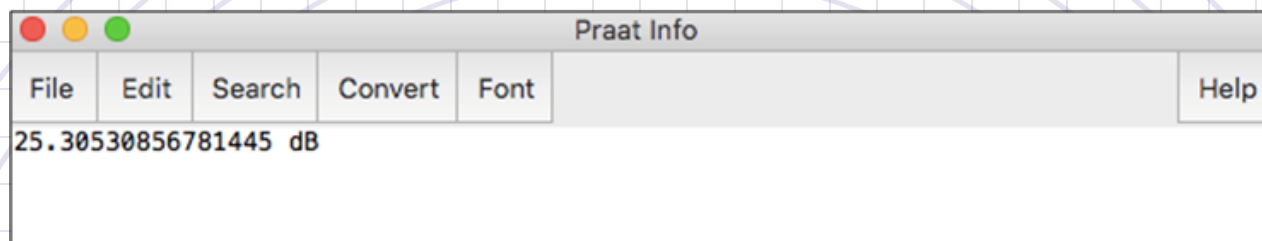


Imagen 77. Configuración para obtener el valor del segundo peak

Por último, se resta el valor del primer peak con el segundo peak y obtenemos el valor del SPR.

1-5 Khz / 5-8 Khz

- Para calcular este valor, debemos alargar el rango. Por tanto, debemos hacer todo de nuevo. Con el archivo de audio en la ventana **Praat objects** ir a **Analyse spectrum** y luego seleccionar **Ltas (Pitch corrected)**.

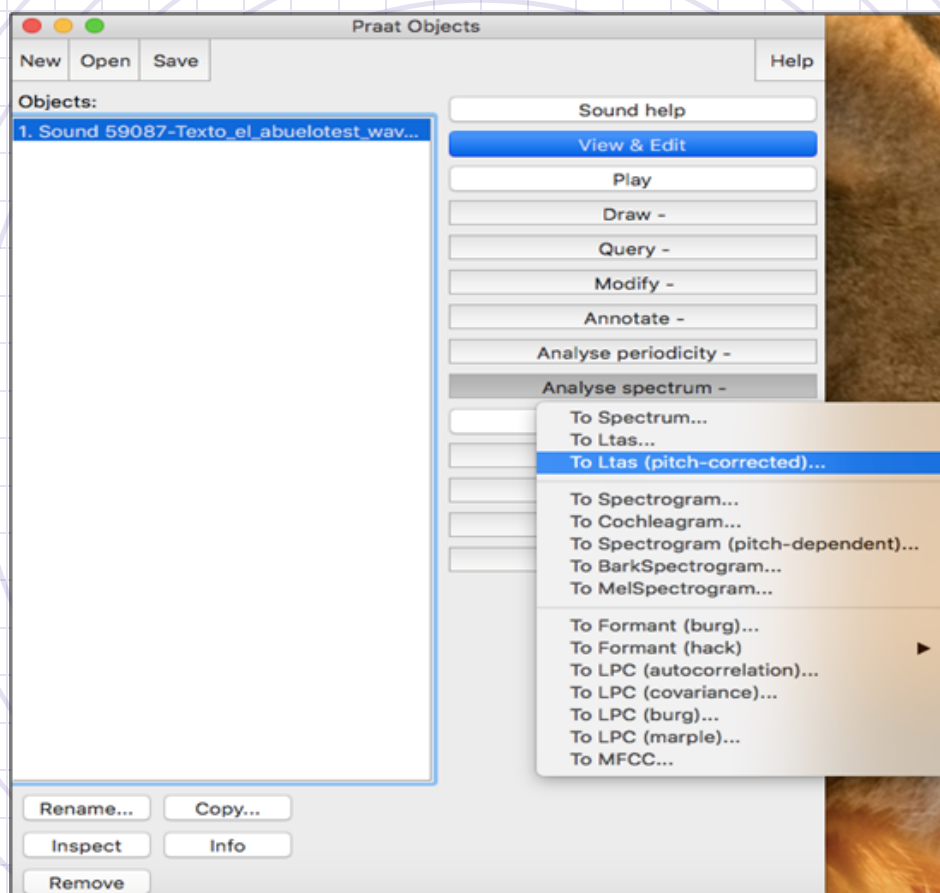


Imagen 79. Selección Ltas (Pitch-corrected)

- Luego se abre una ventana para calibrar Sound: **To Ltas (Pitch-corrected)**. En
- **Minimum pitch** se deja 75 Hz (nunca 0, para eliminar ruidos de baja frecuencia). En **Maximum pitch** se deja 600 Hz lo cual es adecuado para voz hablada, si es

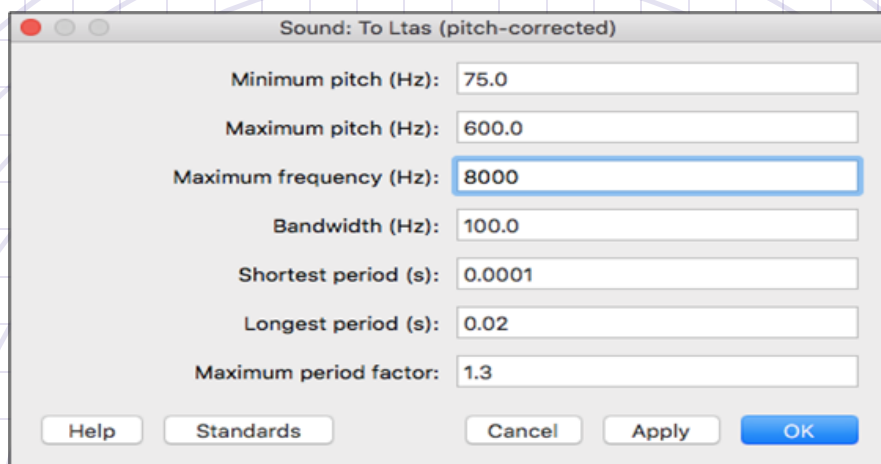


Imagen 80. Configuración Sound To Ltas (pitch-corrected)

- Luego de apretar **OK** el software eliminó todos los sonidos afonos y segmentos sin voz, quedando esta nueva señal la ventana Praat Objects.

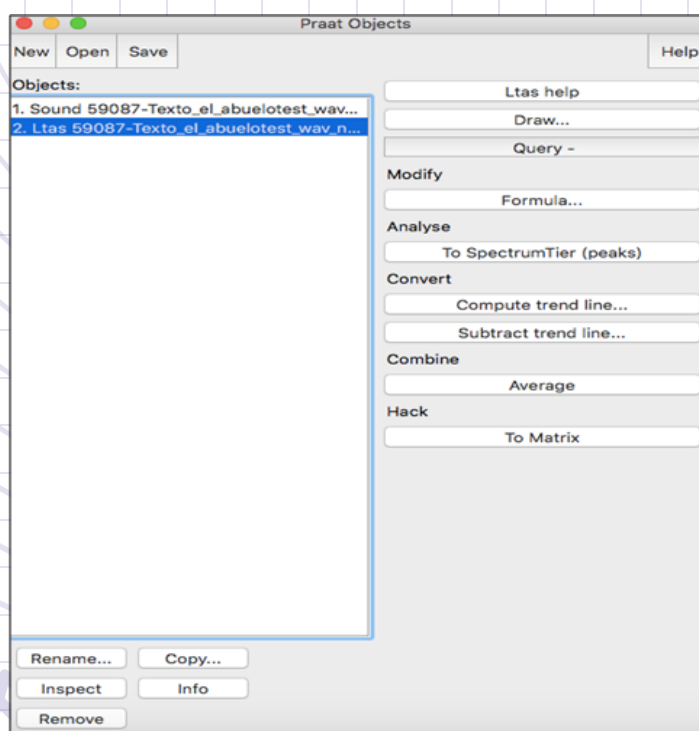


Imagen81. Archivo creado mediante Ltas (pitch-corrected)

- Con el archivo **Ltas** creado en nuestra ventana **Praat Objects**, vamos **Query** y
- luego la opción **Get slope** y se abrirá una ventana para configurar llamada **Ltas: Get slope**.

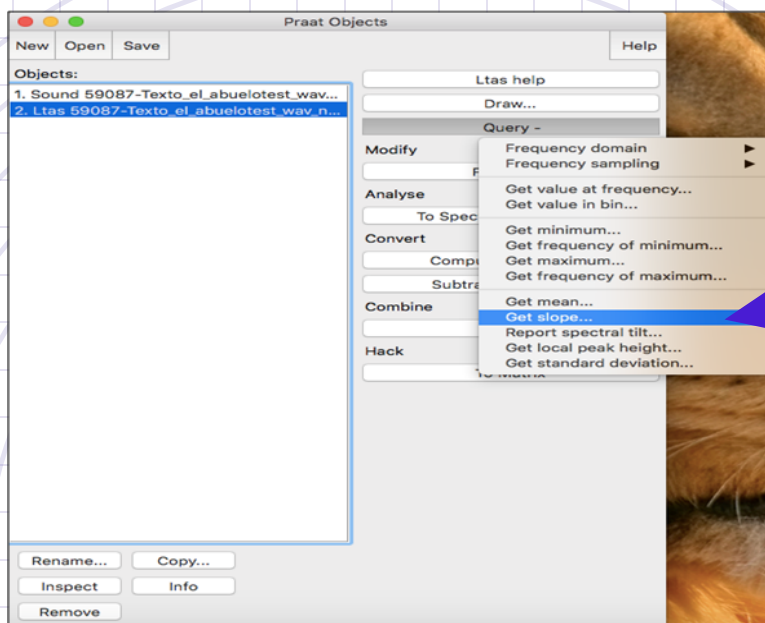


Imagen 82. Selección Get slope

- Con la ventana de configuración **Ltas: Get slope** generada, debemos colocar en
- **Low band (Hz)** 1000-5000. Luego en **High band (Hz)** 5000-8000. En la opción **Averaging method** se selecciona **energy**.

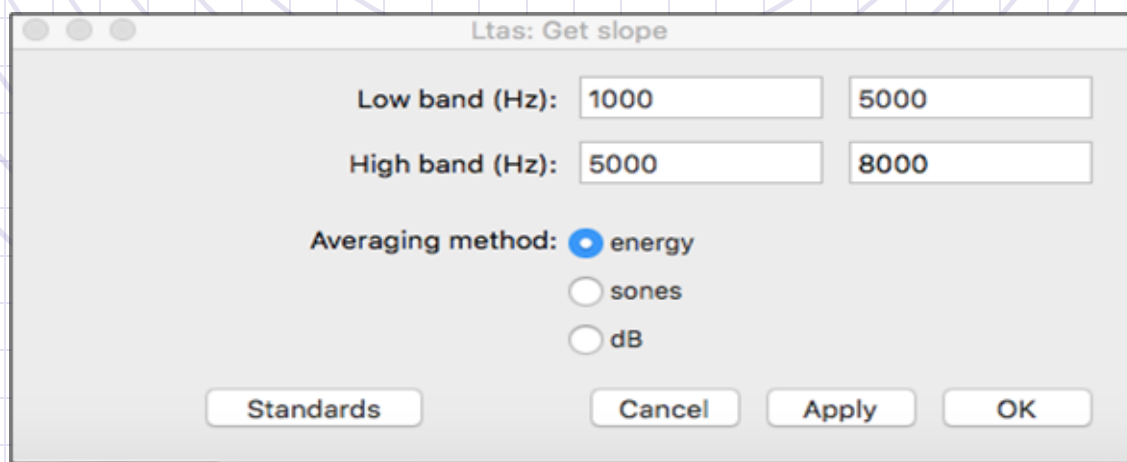


Imagen 83. Configuración 1-5/5-8 Khz

- Presionamos OK y obtenemos el valor.

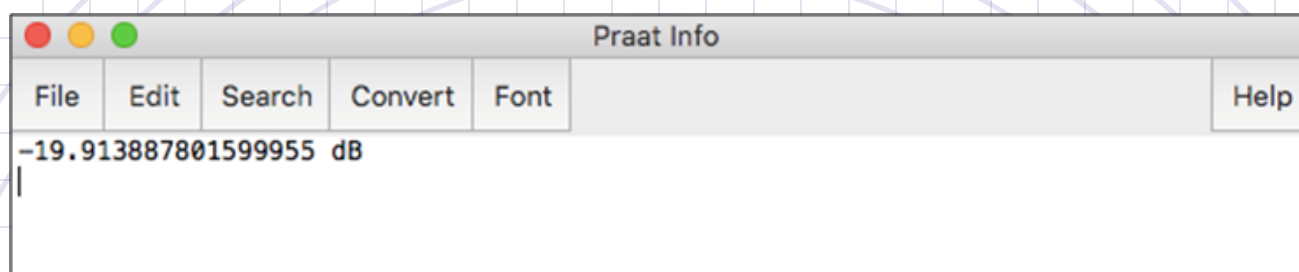


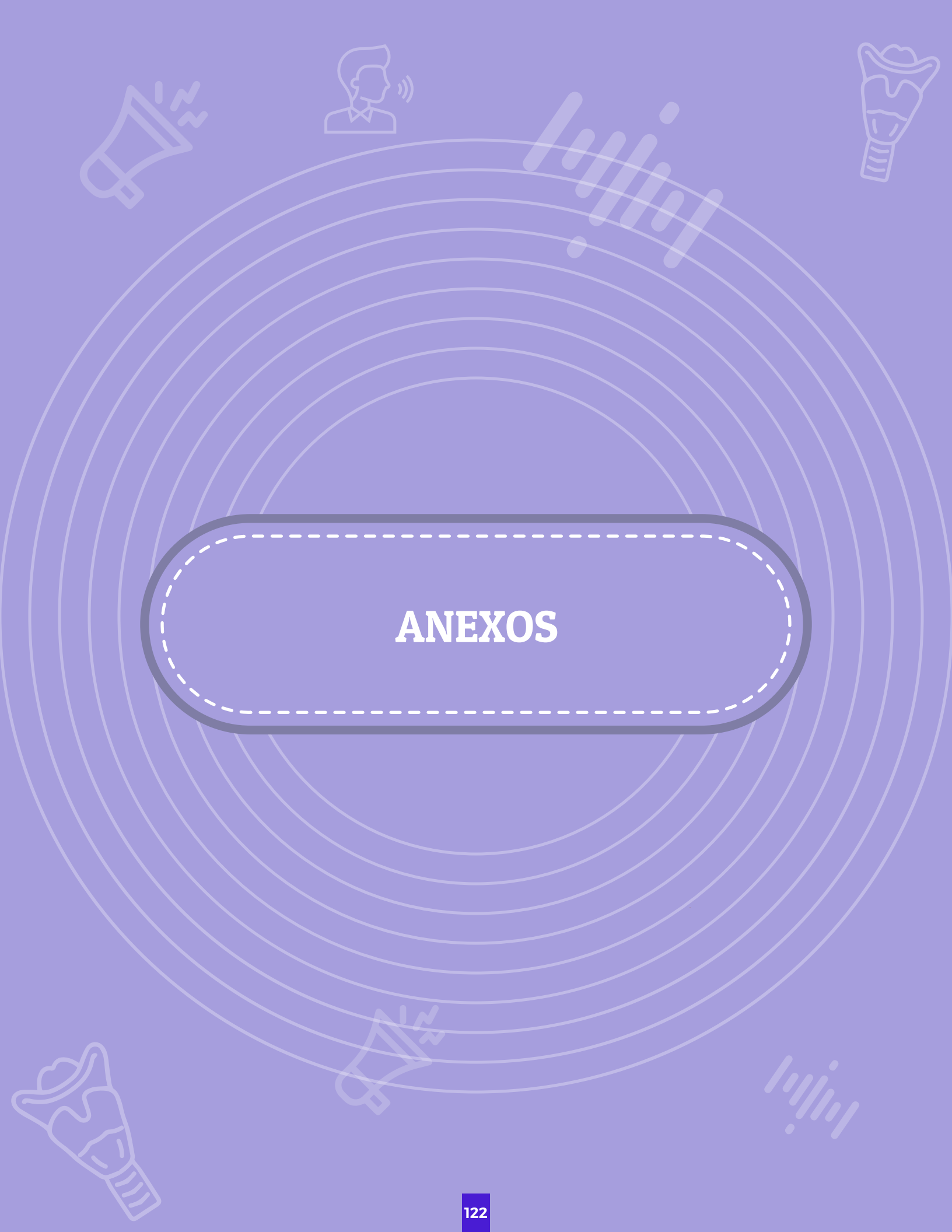
Imagen 84. Valor obtenido 1-5/5-8 Khz

Referencias

1. Aronson, A.S. (1990). *Clinical voice disorders*. New York: Theme medical publishers.
2. Farías, P. (2016). *Guía Clínica para el Especialista en Laringe y Voz*, Buenos Aires: Akadia.
3. Martínez, L., Cabezas, C., Labrar, M y Hernández, R. (2006). La logopedia en Iberoamerica. In XXV Congreso de Logopedia, Foniatría y Audiología.
4. Dejonckere, P.H. (2000). "Valoración perceptual y de laboratorio de la disfonía. En Rosen, C.A. y Murry, T. (eds): *Trastornos de la voz y fonocirugía (parte II)*. Clínicas Otorrinolaringológicas de Norteamérica. México: McGraw-Hill Interamericana.
5. Jackson, M. (2002). *La voz patológica*. Buenos Aires: Editorial Médica Panamericana.
6. Cecconelo, L. (2012). *Aplicación del analisis acústico en la clínica Vocal-Trabajando con Anagraf*. Buenos Aires: Akadia.
7. Jerez, M., Portillo, M., Sastoque, M y Fretes, M. (2016). *Evaluación clínica de la voz*. Buenos Aires: Akadia.
8. Sacket, DL., Rosemberg Wm., Gray, JA, Haynes RB y Richardson, WS. (1996). *Evidence based medicine: what it is and what it isn't*. BMJ; 312:71-72.
9. Farías, P. (2015). *Ejercicios para restaurar la función vocal: Observaciones clínicas*. Buenos Aires: Akadia.
10. Farías, P (2010). Laboratorio de voz en capítulo de Laringe de "Otorrinolaringología y afecciones conexas". Buenos Aires: Ateneo.
11. Saldías, M. (2017). Apuntes de Aprendizaje para el curso: "Evaluación objetiva de la voz" Modulo 1 parte 1. Comunikrte.
12. Stemple, J., Glaze, L y Klaben, B. (2010). *Clinical voice pathology: Theory and management*. San Diego: Plural Publishing.
13. Sataloff, R. (2005). *Voice science*. San Diego: Plural Publishing.
14. McCoy, S. (2015). *Your voice: The basics*. Delaware: Inside View Press.
15. Titze, I y Verdolini, K. (2012). *Vocology: The science and pratique of voice habilitation*. Salt Lake City: National Center for Voice and Speech.
16. Cobeta, I., Núñez, F y Fernández, S. (2013). *Patología de la voz*. Barcelona: Marge Médica Books.
17. Casado, C y Adrián, J., Cobeta, I., Núñez, F., González, R., Rodríguez, M., Coll, R y Muñoz, F. (2002). *La evaluación clínica de la voz. Fundamentos médicos y logopedicos*. Málaga: Aljibe.
18. Titze, I. (1995). *Workshop on acoustic voice analysis: Summary statement*. Iowa Denver, National Center for Voice and Speech.
19. Saldías, M. (2017). Apuntes de aprendizaje N° 5 "Medidas de perturbación". Comunikrte.
20. Kay PENTAX (2008). Software instruction manual: Multi-Dimentional Voice Program (Model 5105). PENTAX Medical Company.

1. Casado, J y Pérez, A. (2009). TRASTORNOS DE LA VOZ: DEL DIAGNÓSTICO AL TRATAMIENTO. Málaga: Aljibe.
2. McCoy, S. (2012). Your voice: An inside view. Delaware: Inside View Press.
3. Yanagihara, N. (1967). "Significance of harmonic changes and noide components in hoarseness". En: *Journal of Speech and Hearing Research*. 10: 531-541.
4. Rabiner, L y Juang, B. (1993). *Fundamentals of speech recognition*. New Jersey: PIR Prentice-Hall
5. Oppenheim, A y Schafer, R. (2004). *From frequency to quefrequency: A history the cepstrum*. IEEE Signal Procesing Magazine. 21: 95-106.
6. Saldías, M. (2017). Apuntes de aprendizaje N° 10: "Análisis Espectral: LTAS". Comunikrte.
7. Nordenberg, M y Sundberg, J. (2003). *Effect on LTAS of vocal loudness variation*. TMH-QPSR. KTH. 45: 93-100.
8. Tanner, K., Roy, N., Ash, A y Buder, E. (2005). Spectral moments of the long-term average spectrum: "Sensitive indices of voice change after therapy?". En: *Journal of voice*. 19(2): 211-222.
9. Sunberg, J. (1994). Vocal fold vibration patterns and phonatory modes. TMH, QPSR, KTH. 35:069-080.
10. Verdolini, K., Druker, D., Palmer, P y Samawi, H. (1999). "Laryngeal adduction in resonant voice". En: *Journal of voice*. 12(3): 315-327.
11. Guzmán, M., Higuera, D., Fincheira, C., Muñoz, D., Guajardo, C y Dowdall, J. (2013). "Immediate acoustic effects of Straw phonation exercises in subjects with dysphonic voices". En: *Logoedics Phoniatics Vocology*. 38:35-45.
12. Leino, T. (2009). "Long-term average spectrum in screening of voice quality in speech: Untrained male university students". En: *Journal of Voice* 23(6): 671-676.
13. Guzman, M., Castro, C., Madrid, S., Olvarria, C., Leiva, M., Muñoz, D., Jaramillo E y Laukkanen, A. (2015). "Air Pressure and contact Quotient Measures During Different Semioccluded Postures in Subjetes with Different Voice Conditions". En: *Journal of Voice*. Vol. 30, No. 6, pp. 759.e1-759.e10.
14. Guzman, M., Angulo, M., Muñoz, D y Mayerhoff, R. (2013). " Effect on long-term average spectrum of pop singers vocal warm-up with vocal function exersices". En: *International Journal of Speech-Language Pathology.*, 15(2): 127-135.
15. Droguett, Y. (2017). "Aplicaciones clínicas del análisis acústico de la voz". En: *Rev. Otorrinolaringol. Cir. Cabeza Cuello*; 77:474-483.
16. Hertrich, I., y Ackermann, H. (1995) "Gender-specific vocal dysfunctions in Enfermedad de Parkinson's disease: electroglottographic and acoustic analyses", En: *Ann Otol Rhinol Laryngol*. Mar;104(3):197-202.
17. Jiang, J., Lin, E., Wang, J., y Hanson, D. (1999). "Glottographic Measures Before and After Levodopa Treatment in Enfermedad de Parkinson's Disease", en: *The Laryngoscope. The American Laryngological, Fthinological and Otological Society, Inc*. Aug;109(8):1287-94.

1. Lin, E., Jiang, J., Hone, S., Hanson, D. (1999). *Photoglottographic Measures in Enfermedad de Parkinson's Disease*. Journal of Voice, Vol. 13, No. 1, pp. 25-35.
2. Santos, L., Reis, L., Bassi, L., Guzella, F., Reis, C. y Cortes, A. (2010). Acoustic and hearing-perceptual voice analysis in individuals with idiopathic Enfermedad de Parkinson's disease in "on" and "off" stages. *Arq Neuropsiquiatri*. [Versión electrónica].
3. Ikui, Y., Nakamura, H., Sano, D., Hyakusoku, H., Kishida, H., Kudo, Y., Joki, H., Koyano, S., Yamauchi, A., Takano, S., Tayama, N., Hirose, H., Oridate, N., y Tanaka, F. (2015). "An aerodynamic Study of Phonations in Patients With Parkinson Disease (PD)", en: *Journal of Voice*. May, 29(3): 273-80.
4. Holmes, R.J., Oates, J.M., Phyland, D.J., y Hughes, A.J. (2000). "Voice characteristics in the progression of Parkinson's disease", en: *Int J Lang Commun Disord*. Jul-Sep; 35(3): 407-18.
5. Jannets, S y Lowit, A. (2014). Cepstral analysis of hypokinetic and ataxic voices: correlations with perceptual and other acoustic measures". En: *Journal of voice*. Nov; 28(6): 673-80.
6. Smith, L y Goberman, A. (2014). "Long-time average spectrum in individuals with Parkinson disease". En: *NeuroRehabilitation*. 35(1): 77-88.
7. Silbergleit, A.K., Lewitt, P.A., Peterson, E.L y Gardner, G.M. (2015). "Quantitative Analysis of Voice in Parkinson Disease Compared to Motor Performance: A Pilot Study". En: *J Parkinsons Dis*. 5(3): 517-24
8. Baken, R y Orlikoff, R. (2000). *Clinical Measurement of speech and voice*. San Diego: Singular.
9. Titze, I. (1990). Interpretation of the Electrolottographic Signal. En: *Journal of Voice*. 4 (1): 1-9.
10. Guzman, M. (2017). Apuntes de Aprendizaje: "Medidas aerodinámicas de la fonación". Comunikrte.
11. Stemple, J. (2014). *Clinical voice pathology*. San Diego: Plural Publishing.
12. Titze, I. (2010). *Principles of Voice Production*. Salt Lake City: National Center for Voice and Speech.
13. Titze, I y Sundberg, J. (1992). Vocal intensity in speakers and singers. En: *Journal of the Acoustical Society of America*. 91, 2936-2946.
14. Verdolini, K., Titze, I y Druker, D. (1990). Changes in phonation threshold pressure with induced conditions of hydration. En: *Journal of Voice*. 4: 142-151.
15. Colton, R., Casper, J., & Leonard, R. (2011). *understaniding voice problems: A Physiological Perspective for Diagnosis and Treatment*. Philadelphia: Lippincott Williamns & Wilkins.

The background of the page is a solid light purple color. In the center, there are several concentric white circles. Overlaid on these circles is a large, horizontal, rounded rectangle with a dashed white border. The word "ANEXOS" is written in white, bold, uppercase letters inside this rectangle. Surrounding the central graphic are several white line-art icons: a megaphone in the top-left, a person speaking in the top-center, a hand holding a megaphone in the top-right, a hand holding a megaphone in the bottom-left, a megaphone in the bottom-center, and a hand holding a megaphone in the bottom-right. There are also some abstract line patterns in the top-right and bottom-right areas.

ANEXOS

Pauta de Evaluación Vocal

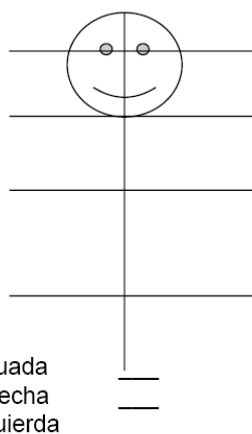
Nombre:

Fecha: __/__/__

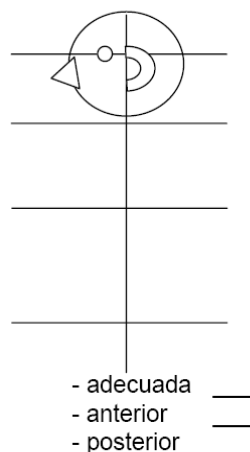
I.- Postura

- Postura Estática en reposo

VISTA ANTERIOR



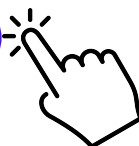
VISTA LATERAL



Descripción:

Pincha este link

https://youtu.be/_fnL5Qg38ZI



II.- Tonicidad

- Dinámica
(cabeza-cuello)

	Adecuada	Insuficiente	Nula	Crujido	Dolor
Flexión	[]	[]	[]	[]	[]
Extensión	[]	[]	[]	[]	[]
Lateralización	[]	[]	[]	[]	[]
Rotación	[]	[]	[]	[]	[]

- Palpación (tonicidad)

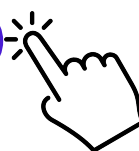
Músculos	0	1	2	3
Músculos Suprahioideos				
Músculos tirohioideos				
Músculos Cricotiroideos				

(Criterio para clasificación de tensión muscular extrínseca. Murray Morrison)

- Altura Laríngea en Reposo [adecuada] [elevada] [descendida]
- Craqueo [presente] [ausente]
- Musculatura ECM [Normal] [Hipertónico] [Hipotónico]

Pincha este link

<https://youtu.be/P9qOGXf833E>



III.- Respiración

Tipo Respiratorio	Costal Superior []	Costodiafragmático []	Abdominal []
Modo Respiratorio	Oral []	Nasal []	Mixto []
Permeabilidad Nasal. Prueba de glatzel	Simétrica []	Asimétrica []	
CFR	Adecuada []	Alterada []	Muy Alterada []
TME /s/:	[]	[]	[]
Apoyo respiratorio en fonación	Presente []	Escaso []	Ausente []

Pincha este link

<https://youtu.be/K0Hv3ahho1Q>



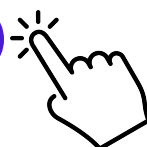
IV.- Parámetros Vocales

Volumen conversacional:	Normal []	Bajo []	Alto []	
Volumen Intenso	Logrado []	No logrado []	Logrado con tensión []	Logrado con cambio de tono []
Rango dinámico	Bajo [] dB	Alto [] dB		
Ataque vocal:	Normal []	Duro []	Soplado []	
Quiebres Tonales:	Ausente []	Al Agudo []	Al grave []	Habla espontánea []
Extensión Tonal	De _____ Hz Hasta _____ Hz De _____ Hasta _____			
Glissando ascendente (desde modal a falsete)	Logrado []	No logrado []		
Glissando descendente (desde falsete a modal)	Logrado []	No logrado []		
Prosodia :	Adecuada []	Exagerada []	Monótona []	
Resonancia :	Adecuada []	Hipernasal []	Hiponasal []	Posterior []
Apertura Bucal:	Adecuada []	Aumentada []	Escasa []	

Temblor de voz :	Presente []	Ausente []		
TMF /a/ :	[]	[]	[]	
Índice s/a:	[]			

Pincha este link

<https://youtu.be/SXRIHFf6gql>



V.- Escala RASATI

	0	1	1-2	2	2-3	3
Ronquera						
Aspereza						
Soplosidad						
Astenia						
Tensión						
Inestabilidad						

VI.- Órganos Fonoarticulatorios

VII.- Análisis Acústico

Índices de perturbación y ruido a corto plazo

Parámetros	Valor obtenido	Referencial (Programa praat)
Frecuencia Fundamental (f0)		
Jitter (rap)		<0.680%
Jitter (ppq5)		<0.840%
Shimmer (local)		<3.810%
Shimmer (apq5)		<3.07%
HNR		>20dB

NHR		<0.19
------------	--	-----------------

Espectro promedio a largo plazo (LTAS)

Medida a largo plazo	Valor obtenido
Alpha Ratio	
L1-L0	
1-5/5-8	
SPR	

Formantes

Formante (vocal a)	Valor frecuencial	Ancho de banda
F1		
F2		
F3		
F4		
F5		