

# **ANAGRAF**

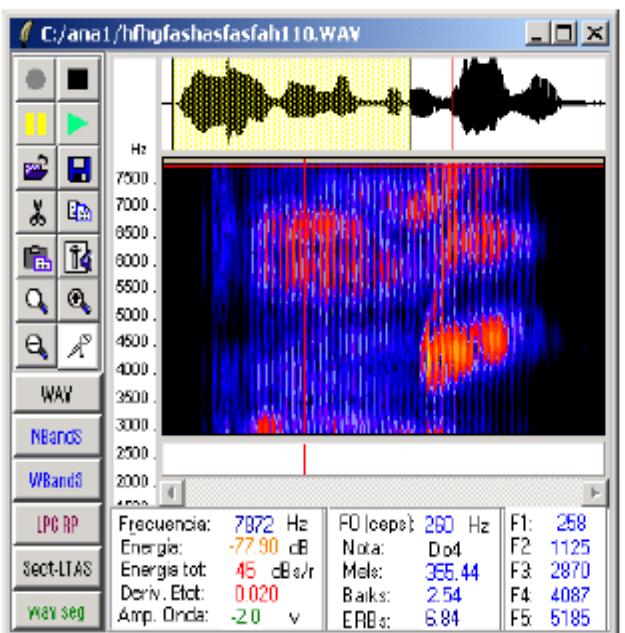
Análisis Acústico y Graficación  
de Señales de Habla

BlackVOX - 2016

## Pantalla Principal

La pantalla se divide en dos partes, una con ventanas a la derecha y otra con varios botones a la izquierda:

BOTONES  
de  
FUNCION



VENTANA  
DE  
GRAFICOS

VENTANA DE  
DATOS

### Botones de FUNCION

Botones de Visualización (Muestran u ocultan ventanas de gráficos)



WAVeform. Forma de Onda Completa, en color negro

Espectrograma de Banda Angosta. "Narrow Band." Se corresponde con ventanas temporales largas. Se visualizan los armónicos, con buena resolución de frecuencias

Espectrograma de Banda Ancha. "Wide Band". Se corresponde con ventanas temporales pequeñas. Se visualizan los detalles en el tiempo.

LPC Row Peaks. Picos máximos de los formantes. Para estudios de voz es recomendable usar armónicos con los formantes superpuestos.

Section Spectra - Long Term Average Spectra. La sección muestra la distribución de formantes con su energía relativa para el instante seleccionado. El LTAS promedia los espectros para el segmento seleccionado.

waveform segment. Segmento de onda en color verde, que ha sido pintado en la onda negra

Zero crossings: Cruces por cero. Energía total. "Overall energy, Root Mean Square. RMS"

F0- Frecuencia Fundamental calculada por la función get\_f0 de Entropic Signal Processing System (ESPS)  
Cp: calculo de F0 por el método de cepstrum, cc: calculo de F0 ciclo a ciclo.



ToBI y otros etiquetados. Espacios para transcripción de acuerdo al método ToBI y ToBI ampliado.

Etiquetas de **Misceláneas**

Alfabetos sampa e ipa. Se requiere instalar el conjunto de caracteres usando el archivo que se provee, llamado Silipa.exe.

Ventana de Evaluación de los Índices de perturbación y Formantes F1 F2 normalizados. Este botón puede figurar también como Riesgo Vocal.

Ventana de **Configuración**. Permite cambiar el color de fondo a blanco para imprimir. Puede ajustar la frecuencia de muestreo a 16000Hz para ver correctamente los formantes. Las escalas pueden ser ajustadas aquí para expandir los contornos de energía y F0.

Ventana de **Datos numéricos** y botón de salida

### *ANAGRAF Análisis Acústico de Señales de Habla*

**Puerta de salida.** Es siempre preferible utilizar este botón para salir del programa y no la X de la ventana. De este modo se garantiza limpiar cualquier archivo residente en memoria.

#### Botones de Grabación y Edición

ABRIR un Archivo



SALVAR el archivo

GRABACION

PARAR la Grabación

PAUSA

REPRODUCCION

CORTAR SEGMENTO  
Onda Negra

VOLUMEN de Reproducción

COPIAR SEGMENTO  
Onda Negra

PEGAR SEGMENTO Onda Negra

ZOOM Menos Onda Verde

ZOOM Más Onda Verde

La grabación se realiza con los parámetros indicados en el menú de Configuración. Si va a realizar una grabación de uso clínico elija 16000 Hz y 16 bits.

## Botones de transformación de la onda

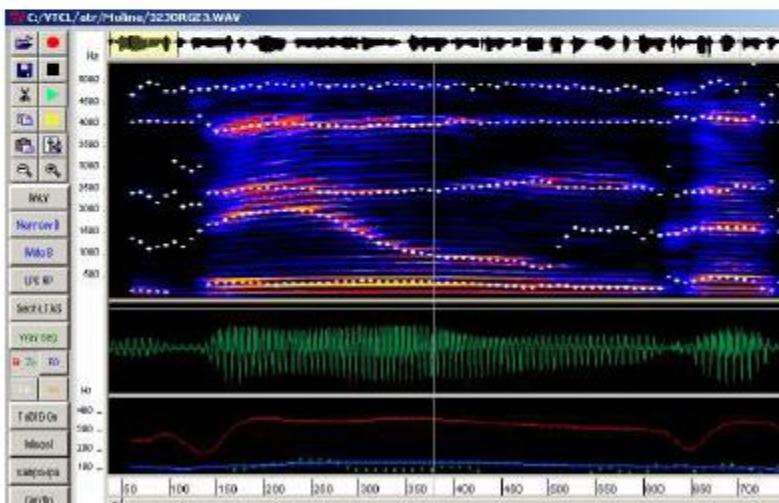
**INVERSION de fase** Invierte los picos positivos y negativos.



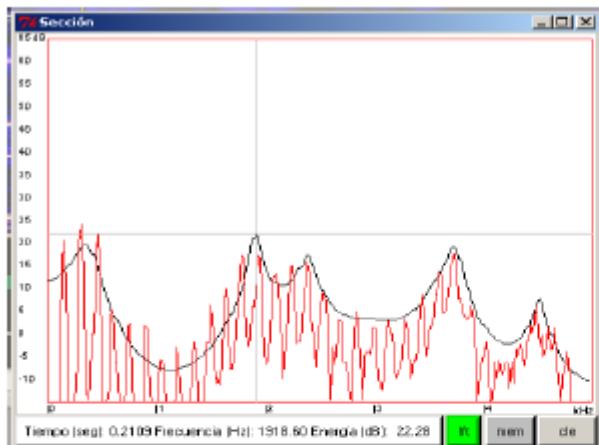
**ELIMINACION de la componente continua.** Se usa para nivelar la onda respecto al nivel de referencia cero.

## Ventana Principal

Muestra la forma de Onda Completa (negro), el Espectrograma, el segmento de forma de Onda (verde), y los Contornos de energía total (rojo), de F0 (azul) y Cruces por cero (verde), todos sincronizados en el tiempo.

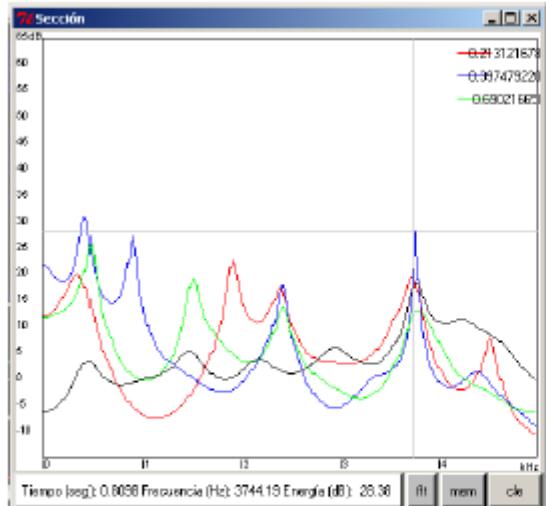


## Ventana de Sección y LTAS



Haga clic en el instante deseado (Sección o Espectro Instantáneo) o pinte un segmento en la pantalla principal.(LTAS). Entonces oprima el botón Sect-LTAS. Puede conservar el contorno obtenido apretando "MEM". Puede visualizar o no los armónicos apretando "FFT". En ambos caso se iluminará la tecla en verde.

Espectro instantáneo de formantes y de armónicos superpuestos para visualizar el grado de aprovechamiento de energía medida con B. Un aumento del B1 es indicador de perdida de energía por nasalización. La nasalización puede visualizarse si existen componentes nasales, como picos espectrales entre los 250-400Hz (por debajo del F1 oral de la vocal /a/), o alrededor de los 900-1100 Hz (entre el F1 y F2 de la vocal /i/).



El espectro LTASS (Long Term Average Speech Spectrum) se obtiene con frases fonéticamente completas. Para que sea balanceada se requieren dos o más frases (45 segundos) que representen al histograma de frecuencias típico. El espectro promedio LTASS permite observar características de largo tiempo asociadas a la perdida de componentes frecuenciales en la producción de habla.

Espectros de formantes superpuestos con la tecla MEM y FFT inactivo. En azul la vocal /a/, en rojo la /i/ y en verde la /e/.

## Ventana de Datos

1. Todas las mediciones serán visualizadas en esta ventana.
2. Los valores instantáneos de: frecuencia, energía, y tiempo se observan de acuerdo a la posición de la flecha del Ratón en el Espectrograma. Los valores instantáneos de energía total y F0 en la ventana correspondiente.
3. Si está activado el botón de LPC Raw Peaks (puntos blancos), se medirán los valores de los formantes en donde este posicionado el cursor vertical. Se leerán las Frecuencias, F1 a F5 en Hz, y los Anchos de Banda B1 a B5 en Hz.
4. Los valores calculados de Jitter, Shimmer, Relación H/N, y F0 promedio se obtienen pintando un segmento de onda y apretando el botón amarillo CAL.
5. La duración de un segmento pintado puede utilizarse para medir un periodo de la onda y obtener su inversa la F0.

Frecuencia:	256	Hz	F0:	0	Hz	F1:	381	B1:	127	Tiempo:	0.1869	s
Energía:	24.91	dB	Nota:	exced		F2:	1998	B2:	61	Duración:	0.5231	s
Energía tot:	-18	dBs/r	Mels:	0.00		F3:	3373	B3:	436	1/Dur:	.2	c/s
Deriv. Etot:	0.000		Barks:	0.00		F4:	4428	B4:	834	F0 inst.:	256	Hz
Ruido Amb:	18	dBr/1v	ERB:	0.00		F5:	4865	B5:	175	Amp. Onda:	.1.2	v

## Botón de Cálculo

Calculo de los parámetros de Voz. El valor de Jitter (variación de F0 ciclo a ciclo se confirma como válido al igualarse los valores (promedio) y (ciclo a ciclo) de F0. En algunas voces patológicas no se igualan los valores de F0 y puede aceptarse una variación entre las mediciones. Esta variación será penalizada al calcular el índice de perturbación en forma automática.

Los valores de la relación armónico ruido pierden confiabilidad si la medida de Jitter es muy alta.

La medición de Jitter puede dar errores, para ello acepte las advertencias y vuelva a pintar la onda en segmentos más cortos. Si el valor de Jitter no se puede obtener confiablemente, cambie la frecuencia de muestreo a 10.000 en el menú de configuración, convierta y vuelva a calcular.

## Armónico ruido vs. Cruces por cero

Para la medida de ruido, la mejor medida es la relación armónico ruido (debe ser mayor a 4). Los cruces por cero dan información cualitativa dentro del segmento de habla analizado visualmente donde hay más ruido.

## Ataque soplado vs. aspereza

Cuando el ruido aparece al inicio de la emisión vocalica puede aparecer débil y el f0 estar ausente. En otras ocasiones el F0 está entrecortado no hay ruido y se percibe una aspereza.

Esto es debido a la sensación que producen los cortes de F0. En una tercera alternativa el F0 puede aparecer cortado y en cada corte aparecer el ruido débil. Aquí se combinan ambos efectos mencionados en primer término.

<b>F0 (pt):</b>	<b>193</b>	<b>Jitter (pt)</b>	<b>0.30</b>
<b>F0 (cc):</b>	<b>193</b>	<b>Jitter (cc):</b>	<b>1.02</b>
<b>H/N (dB):</b>	<b>4.36</b>	<b>Shimmer (%):</b>	<b>0.94</b>
<b>Fase</b>	Lyapunov Max. Exp.: 0.33		

Amplitud del Cepstrum y Cruces por Cero

Jitter Promedio y Jitter Ciclo a Ciclo

Relación Armónico Ruido y Shimmer

Coeficiente de Lyapunov

CAL : calcula los valores. Fase: Dibuja el gráfico de Lyapunov

## Cálculos de F0

El método de cálculo estándar (puntos azules grandes) permite ver el contorno de F0 solo en los segmentos de la señal que son altamente confiables (amplitud suficiente y periodicidad comprobable). Su utilización es más de uso fonético que foniático.

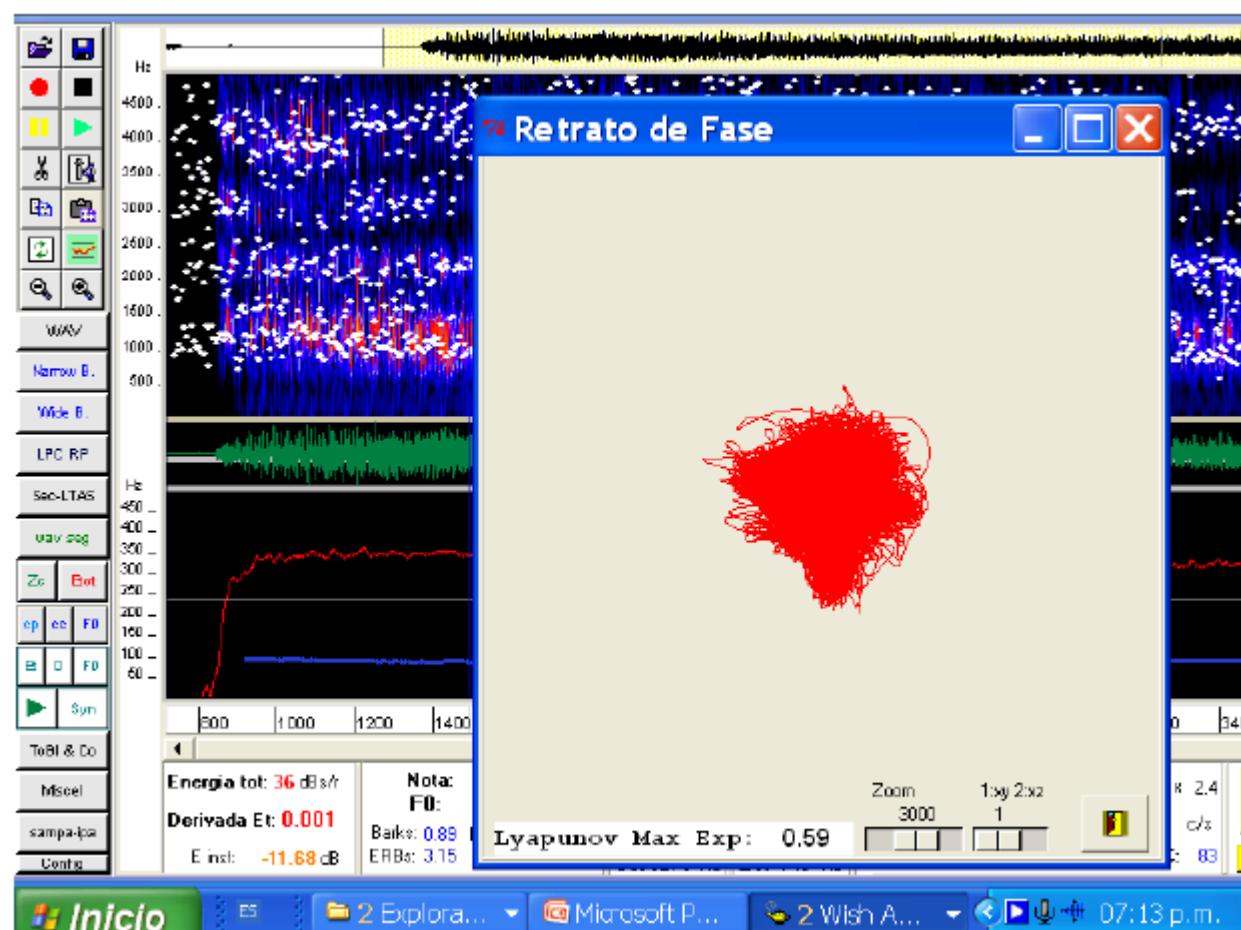
El calculo de F0 por cepstrum da más información para fines foniáticos. En las zonas de ausencia de F0 o de ruido, grafica una nube de puntos sin continuidad.

El calculo de F0 ciclo a ciclo es altamente impreciso, por tener menos restricciones para la medición. Puede ser útil como método más arriesgado, cuando los otros dos métodos se inhiben o no muestran valores que serían esperables. Para aceptar algún valor particular, el criterio a seguir es mirar la continuidad de la curva de F0.

## Retrato de Fase

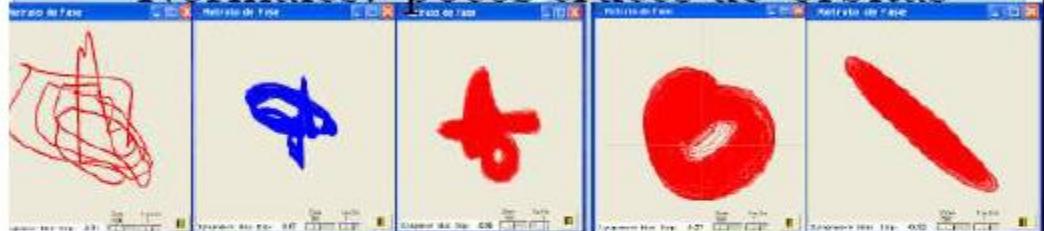
El gráfico que se presenta al oprimir el botón de Fase se conoce con el nombre de Retrato de Fase. Este gráfico de tres dimensiones puede aparecer como una madeja bien delineada para una vocal /a/ perfecta obtenida mediante la síntesis o como un ovillo muy denso para alteraciones muy severas. La medición objetiva está aún en desarrollo y consiste en obtener un número que ejemplifique los tipos de gráficos comentados. Se incluyen trabajos publicados preliminares en la carpeta de artículos. Con la cuantificación se pretende obtener una medición ligada al número de huecos de la figura tridimensional. Cuantos más huecos hay más correcta es la emisión (valores pequeños y negativos) y cuanto más relleno el ovillo, mas alteración (valores mayores).

El grafico del retrato de fase incluye la cuantificación. Este valor es el mismo que aparece en la pantalla principal cuando se oprime el botón CAL de cálculo.



## Retratos de Fase . normales y alterados

Normales: pocos cruces de órbitas

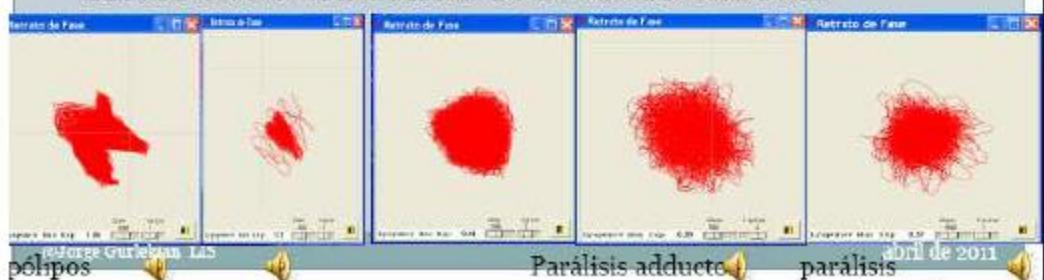


Vocal artificial y natural  
masc

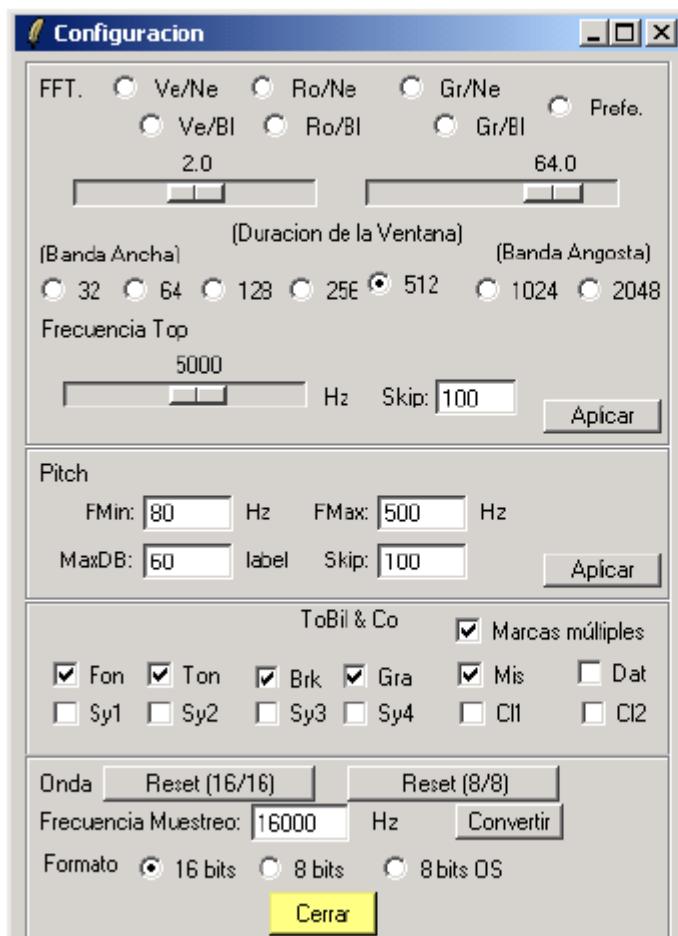
Nat fem

Soprano con apoyo

Alterados: Muchos cruces de órbitas



### Ventana de Configuración



Define los colores de frente y fondo.

Contraste y Brillo

Duración de la Ventana de análisis

Frecuencia Tope del Espectrograma

Separación de las ventanas de análisis

Escalas de F0 y Energía Total

Fon: Fonético

Ton: Tonal

Brk: Pausas

Gra: Grafémico

Mis: Misceláneas

Dat: Val. Acústicos

Sy: Sintáctico

Cl: Clases de palabras

Convertidor de la frecuencia y nivel de muestreo.

Transcripciones ToBI &Co.

Una vez seleccionados los niveles de transcripción en el menú de configuración, podrá activar espacios de transcripción en distintos niveles en sincronía con los gráficos temporales. Al concluir la transcripción guarde el archivo de onda con el mismo nombre y automáticamente se guardarán todos los archivos en formato de texto (formato WAVES) asociados a los distintos niveles de transcripción que haya utilizado. Luego podrá importar esos archivos desde Excel o utilizarlos con otros programas.

**Las opciones son:**

1. Hacer una marca: con un doble click
  2. Escribir la transcripción: con el teclado, inmediatamente después de hacer la marca
  3. Editar la transcripción: con el teclado, posicionando el cursor del mouse en el texto
  4. Borrar la marca: colocando el puntero del mouse sobre la marca hasta que aparezca una doble flecha y apretando en el teclado Supr o Del
  5. Mover la marca: colocando el puntero del mouse sobre la marca hasta que aparezca una doble flecha oprimir y sostener el botón izquierdo y moviendo el mouse

Cuando aparezca un aviso de guardado, salve el archivo de onda o ignore la advertencia.

Las etiquetas fonéticas y tonales podrá encontrarlas en la lista de trabajos publicados.

Los cambios del tipo de letra (botón SAMPA IPA) en el etiquetado se concretan al mover levemente una etiqueta de su lugar.

Ventana SAMPA e IPA

Elija transcribir con los alfabetos SAMPA o IPA apretando el botón correspondiente. (En el caso de transcribir con IPA deberá tener los simbolos IPA cargados en su PC. Para ello active el programa *silipa.exe* que se encuentra en el directorio donde se ha instalado el programa Anagraf). Si elige transcribir con SAMPA lea el artículo incluido en la documentación para usar los caracteres del teclado común.



## Ventana de Misceláneas

Emisión de salida		Fonación	Modo	Punto	
<input type="radio"/> sus	Sustituida	<input type="radio"/> agu	Aguizada	<input type="radio"/> lab	Labializado
<input type="radio"/> agr	Agregada	<input type="radio"/> gra	Agraviada	<input type="radio"/> den	Dentalizado
<input type="radio"/> cmi	Omitida	<input type="radio"/> mon	Monótona	<input type="radio"/> pal	Palatalizado
<input type="radio"/> ini	Ininteligible	<input type="radio"/> fue	Fuerte	<input type="radio"/> vel	Velalizado
<input type="radio"/> ina	Inaudible	<input type="radio"/> glo	Golpe Glótico	<input type="radio"/> far	Faringealizado
<input type="radio"/> sl	Silencio	<input type="radio"/> lar	Laringealizada	<input type="radio"/> ant	Anterior
<input type="radio"/> epe	Epéntesis	<input type="radio"/> ten	Tensiva	<input type="radio"/> pos	Posterior
<input type="radio"/> da	Alargada	<input type="radio"/> nor	Normal	<input type="radio"/> ele	Elevado
<input type="radio"/> eco	Acoitada	<input type="radio"/> deb	Débil	<input type="radio"/> des	Descendido
<input type="radio"/> dup	Duplicada	<input type="radio"/> sop	Soplada		
<input type="radio"/> dip	Díptongada	<input type="radio"/> sor	Ensordecida		
<input type="radio"/> que	con Quiebres	<input type="radio"/> son	Sonorizada		
		<input type="radio"/> dif	Diplófonica		
		<input type="radio"/> lim	extensión Limitada		

Insertar

Categorías de Misceláneas Elija una categoría por título y haga clic en el nivel de misceláneas para insertar el acrónimo.

## Botones de Síntesis

Correlatos acústicos de la prosodia: Duración, Energía y F0



Al oprimir los botones de Et: Energía total, D: Duración ó F0: Frecuencia fundamental podrá visualizar círculos sobre los contornos reales respectivos. Si oprime cada botón otra vez los círculos desaparecen. Los círculos también desaparecen si se oprime el botón F0 que controla la visualización de los contornos originales.

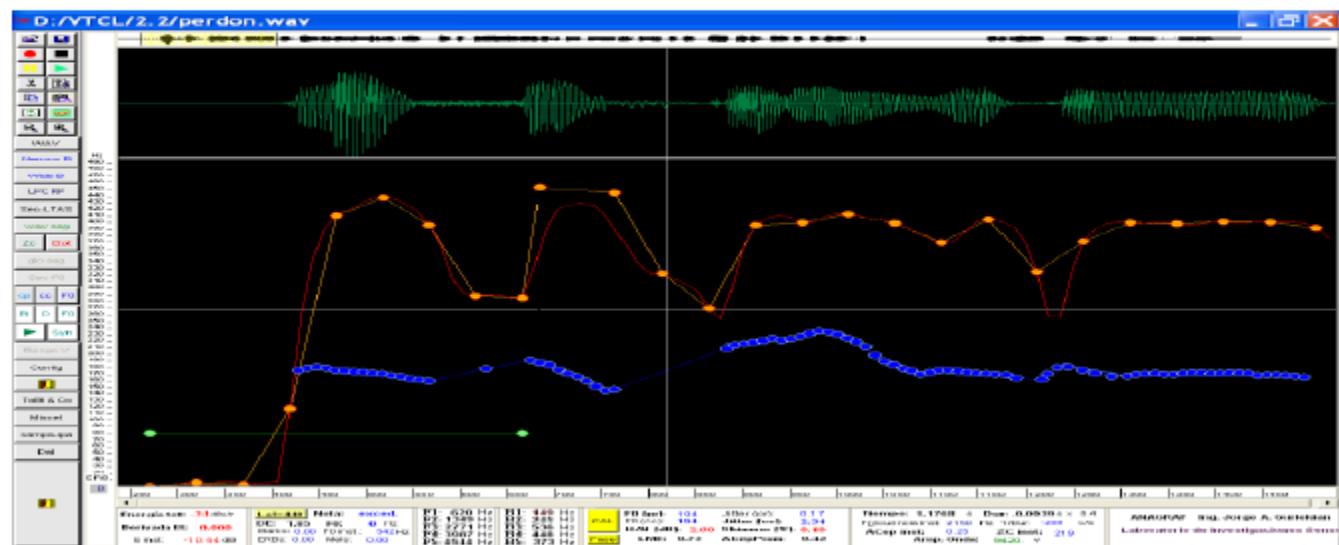
**Atención:** Para activar el desplazamiento de círculos deberá hacer clic en la ventana de contornos y desplazar el ratón hacia algún extremo de la ventana de contornos, por ejemplo en parte superior izquierda para que los cursores horizontal y vertical no interfieran con la operación de desplazamiento de círculos.

Coloque el puntero del ratón en la parte superior interna del círculo y oprima el botón derecho sosteniéndolo mientras desplace el círculo. Al soltarlo, el círculo cambiará de color para indicarle que ha sido capturado. Puede desplazar los círculos en todas las direcciones. Además puede superponer varios círculos si desea visualizar sólo los que considera importantes.

Luego oprima el botón Syn: Sintetizar, le permitirá oír la señal con el/los contorno/s modificados. Oprima play verde para volver a oír lo sintetizado.

El control de la duración esta diseñado de la siguiente forma. Al oprimir D: Duración, aparecerá en pantalla una línea horizontal posicionada en una Línea 1 (velocidad original) de referencia, con círculos. Para *aumentar* la duración (lentificación) de los segmentos deberá colocar los círculos por *arriba* de la Línea 1 y para *disminuir* la duración (aceleración) deberá mover los círculos hacia *abajo*. El factor de multiplicación de la duración podrá leerse en la ventana de datos numéricos.

El control de energía actúa igual que el de F0.

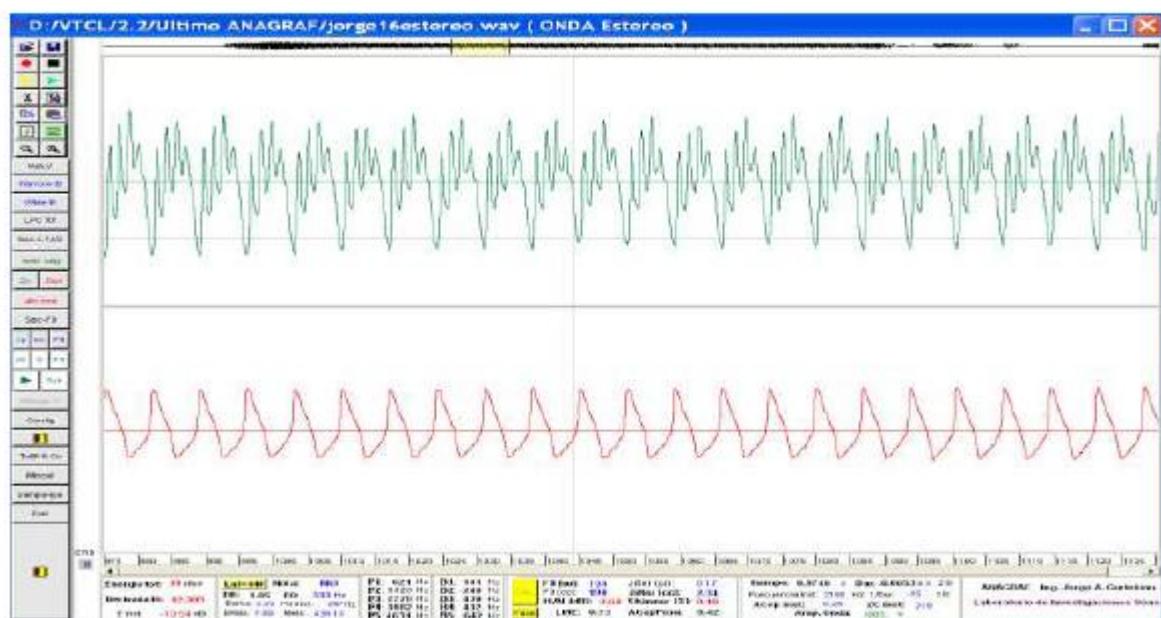


La modificación se logra ubicándose en el circulito y arrastrándolo con el botón derecho del mouse.

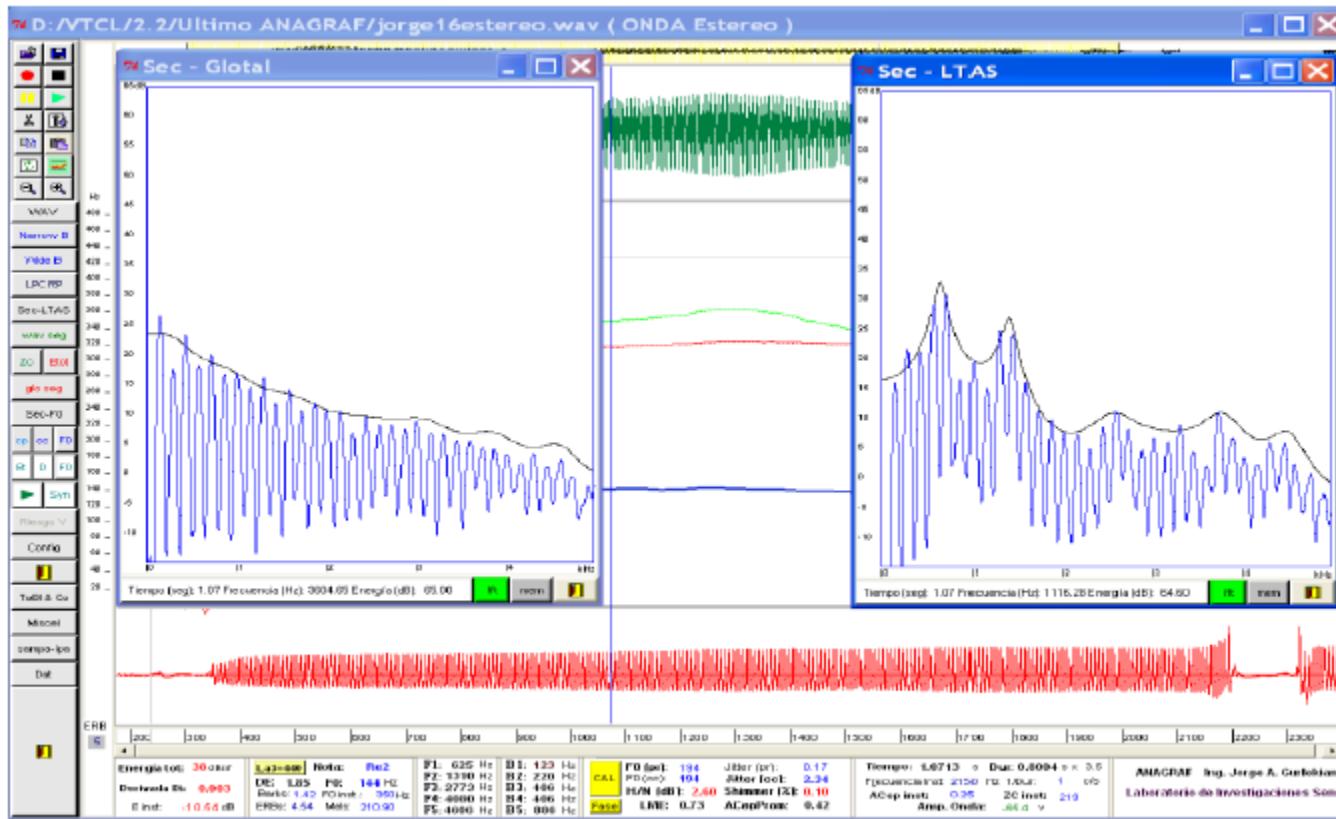
#### Botones de visualización de la onda Glotal (Requiere un equipo EGG).



Permite la visualización de una señal obtenida por un electroglotógrafo. Se muestra la onda glótica (glo seg) y el espectro instantáneo y de largo plazo (Sec-F0) de la señal glótica como se ve en la figura. (Equipo en desarrollo).



Ondas tomadas a la salida del tracto vocal (arriba) y a la entrada (abajo).



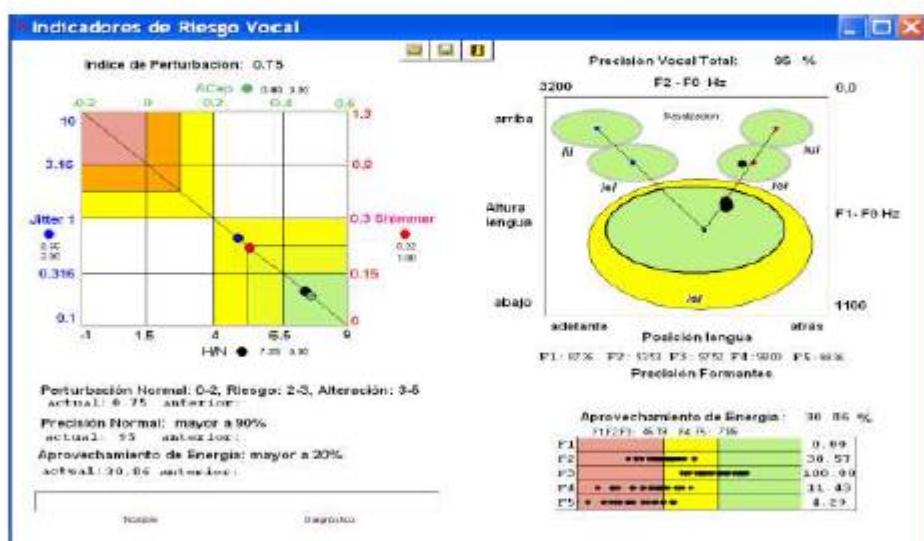
I

Riesgo V

**Botón de Riesgo Vocal.** Para que se active debe estar presente el contorno de Frecuencia Fundamental y haber realizado con el botón CAL el cálculo de los índices de perturbación. Para observar el grado de precisión visualice en pantalla los contornos de formantes (puntos blancos).

## Ventana de Evaluación del Riesgo vocal

Al oprimir el botón aparecerá esta ventana.



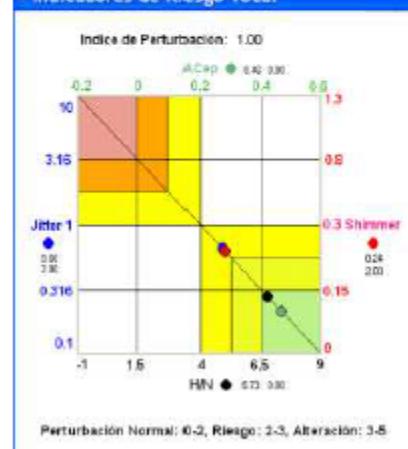
Ventana de evaluación. Se presenta un índice de perturbación global que agrupa cuatro índices clásicos de perturbación vocal, el grado de precisión vocal y el grado de aprovechamiento de energía de los tres primeros formantes con relación a los armónicos más próximos.

## Índice de perturbación integrado (IPI)

Presenta una diagonal que va desde los valores normales (abajo a la derecha y rectángulo verde) hasta valores patológicos (arriba a la izquierda y rectángulo rojo). La diagonal cruza en su centro los valores estadísticos de los umbrales de normalidad.

Para obtener el gráfico correspondiente a un hablante ó paciente, se deben calcular previamente los índices con el botón CAL. Luego, al oprimir el botón Índices de perturbación los valores se dibujaran sobre la diagonal. Los botones abrir y guardar permiten almacenar los índices para su comparación futura. La ubicación de los índices en la zona verde clara provee un pronóstico de disfonía que depende tanto de la proximidad de las índices al centro como de la cantidad de índices.

Indicadores de Riesgo Vocal



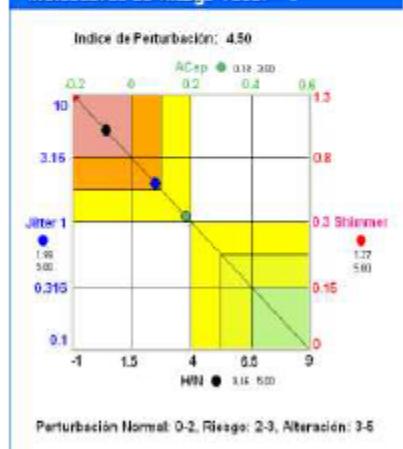
Voz normal

Indicadores de Riesgo Vocal



Voz con riesgo

Indicadores de Riesgo Vocal



Voz con alteración

## Índice de precisión vocal. (IPV)

Presenta las áreas F1 F2 normalizadas respecto de la frecuencia fundamental para las cinco vocales.

Para obtener el gráfico correspondiente a un hablante ó paciente, se deben calcular previamente los formantes LPC RP y el contorno de Frecuencia Fundamental (F0).

Opcionalmente haciendo click en cada instante en el espectrograma se agregarán puntos F1F2 sobre las áreas normalizadas. Los botones abrir y guardar permiten almacenar los datos F1F2 para su comparación futura.

La ubicación de los pares F1F2 normalizados fuera de los límites de las elipses de referencia, pueden corresponder a valores aceptables de emisión si a juicio del profesional actuante no están en conflicto con otras vocales. El índice de precisión vocal se calcula a partir de la desviación estándar de acuerdo a esta fórmula: Precisión= 100 - (2 DE / VM ) \* 100

Los circulitos negros sobre las áreas de las vocales son los valores de F1yF2 en un instante dado.

Si el paciente no mueve sus articuladores durante la emisión sostenida de la vocal /a/, los circulitos deberían caer siempre en el mismo lugar. En ese caso la precisión sería del 100%.

El programa promedia la estabilidad de los cinco formantes dándole más peso a los primeros tres.

Como no siempre se mantienen los articuladores, se ha experimentado (Gurlekian y Molina) que la precisión se considera normal hasta un valor del 90%. Por debajo de ese valor se habla de falta de precisión.

Por ejemplo un paciente con temblor en el tracto vocal dará valores del 50 o 60% (Gurlekian y Sigal, Tesis en preparación). En una voz profesional se busca que la precisión sea alta para que pueda aprovechar bien la energía de los armónicos glotales. En un paciente no profesional, da una indicación que el paciente debe mejorar el control de su articulación.

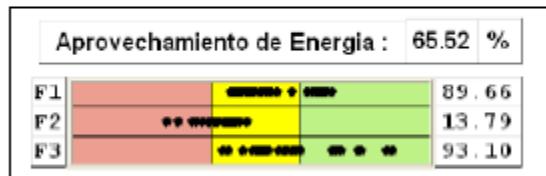
## Índice de aprovechamiento de energía. (IAE)

Se calcula a partir de la medición del ancho de banda de los formantes F1 F2 y F3.

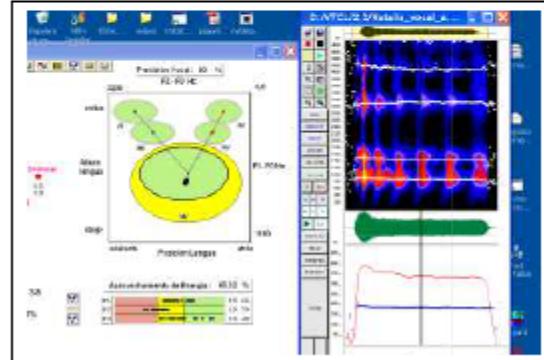
Cuando los anchos de banda son iguales o mayores a los esperados el índice de aprovechamiento disminuye. Se contabiliza esta situación y se presenta el valor en relación a un número total de mediciones.

Para obtener el gráfico correspondiente a un hablante ó paciente, se deben calcular previamente los formantes LPC RP y el contorno de Frecuencia Fundamental (F0).

Opcionalmente puede mover el cursor a lo largo de todo el cuerpo de la vocal oprimiendo el botón izquierdo para agregar medidas adicionales de ancho de banda. El color rosado indica un pobre aprovechamiento y el verde agua buen aprovechamiento.



En este ejemplo se visualiza un buen aprovechamiento en el F1 y F3, formantes que en la vocal /a/ normalmente provienen de la cavidad faríngea y un pobre aprovechamiento de la cavidad oral anterior.



## Informe Fonoaudiológico e Impresión

Comience por abrir una plantilla vacía de informe fonoaudiológico en un programa de edición de textos. Una plantilla de ejemplo se encuentra en el CD donde están los programas originales

En el Anagraf si lo desea cambie el fondo a color blanco, quite la barra del menú inferior de Windows, coloque los formantes superpuestos a los armónicos, F0cp como contorno de F0 y calcule los parámetros de perturbación para los informes de voz.

Para los informes de habla use banda ancha con los formantes superpuestos y el contorno de F0 convencional.

Cuando termine esta preparación oprima en el teclado la tecla *Print Screen* o *Imprimir Pantalla*, esto es equivalente al comando "copiar". Luego vaya a la plantilla seleccione el cuadrado de texto y pegue la imagen.

Para los primeros informes utilice esta la plantilla que figura en el CD. Así tendrá un formato orientativo del informe fonoaudiológico. Coloque los datos del paciente y su diagnóstico ORL en el primer cuadrado de texto. Luego, coloque la figura elegida en el programa ANAGRAF que capturó con ImprPant, mediante el comando o ícono "pegar" en el segundo cuadrado de texto de la plantilla. Utilice el tercer recuadro de texto para hacer su informe, indicando cual ha sido su percepción auditiva, la visualización de su audición en el grafico presentado, y la convalidación cuantitativa, si corresponde, indicada en forma de dato numérico en la tabla del cuadrado de texto final (vea la plantilla), todo en ese orden. Guarde e imprima en color.

Cuando lo deseé, diseñe un informe más personal con logo incluido. Elija las imágenes que serán informadas y distribúyalas en varias hojas. Cada gráfico deberá tener una explicación del contenido.

En una hoja separada incluya la ventana de evaluación con los índices de perturbación y las áreas vocálicas normalizadas informando sobre los datos obtenidos y su pronóstico.

## Referencias

- Rabiner, L.R. and Shafer, R.W.,(1978). "Digital Signal Processing of Speech Signals", Prentice Hall, Inc., Englewoods Cliffs, New Jersey.
- Von Bekesy,G. (1960). "Experiments in Hearing".(edited by E.G. Wever.) McGraw-Hill,New York
- Seneff, S.(1985). "Pitch and Spectral Analysis of Speech based on an Auditory Synchrony Model", PhD Thesis, M.I.T., Department of Electrical Engineering and Computer Science
- Seneff,S.(1986). "A computational model for the peripheral auditory system: application to SR research", Proc. IEEE ICASSP-86, Tokyo, Abril 7-11, 1986, pp. 37.8.1-37.8.4.
- Zwicker, E. (1961). "Subdivision of the Audible Frequency range into Critical Bands (Frequenzgruppen)", JASA 33,248-249.

## Bibliografia

- Índice de Perturbación, de precisión vocal y de grado de aprovechamiento de energía para la evaluación del riesgo vocal. Gurlekian, J.A. y Molina N. Logopedia, Foniatria y Audiología. Elsevier Doyma. 2012, 32, 156-163.
- El alfabeto fonético SAMPA y el diseño de cónsonas fonéticamente balanceado. Gurlekian, J.A., Colantoni, L y Torres, H. Fonoaudiológica. Asalfa, 2001. 47,3, 58-69
- Evaluación Acústica de parálisis unilaterales mediante coeficientes de Lyapunov. Elisei,N.,Evin;D., Gómez Fernández F. Gurlekian, J.A. Poster. Jornadas de Investigación de la UBA
- Preparación de un Informe Clínico para tres patologías de la voz. Gurlekian, J.A. Facal, M.L y Salvatori J.M. *Topicos em Fonoaudiología*. CEFAC, Editorial LOVISE (Marchesan, Zorzi, Gomes, eds.) Pp. 539-554

## Para referenciar el sistema de análisis Anagraf:

Gurlekian, J. A. "El laboratorio de Audición y Habla del LIS". En Procesos Sensoriales y Cognitivos. Editorial Dunken. (editor: M. Guirao).1997. pp. 55-81.