

Manual de usuario

LogoSpeech Studio es una herramienta diseñada especialmente para la investigación fonética o fonológica de señales bioacústicas. Desarrollada para ser una herramienta multiplataforma, nace con la idea de dar apoyo en la docencia logopeda en aspectos como el reconocimiento del habla o la instrucción fonadora. Por ello se diseñó, íntegramente, buscando un funcionamiento simple y ameno para su correcto uso por usuarios menos experimentados. Esta sección pretende servir de guía para la iniciación en las tareas básicas a desarrollar con la misma.

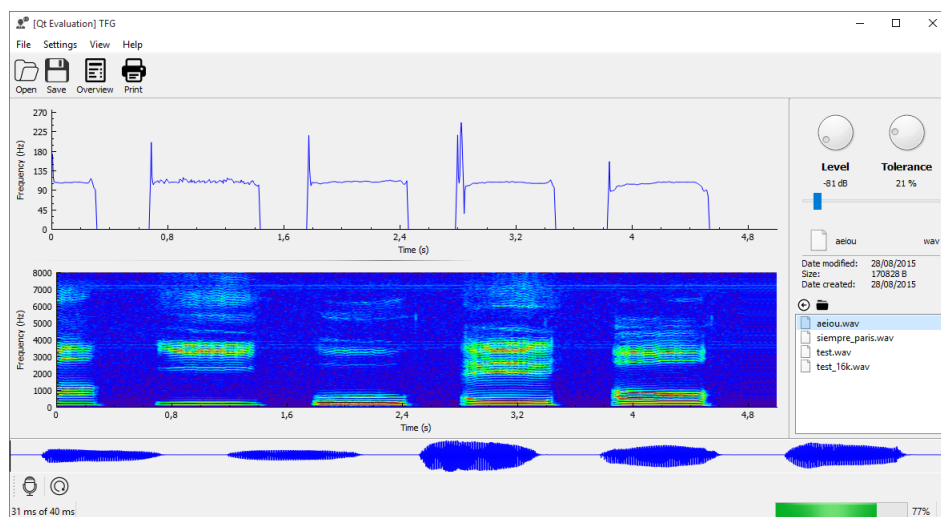


Figure 1: LogoSpeech Studio, entorno de usuario

En la Figura 1 podemos observar la interfaz principal de la aplicación. Se ha elegido una interfaz clásica y sencilla, con elementos de gran tamaño y fácilmente reconocibles para su rápida integración con el usuario. De esta forma se pueden distinguir principalmente 4 secciones de control, figura 2:

- Barra de control y conjuración: conforma el menú principal de la aplicación. A través del mismo tendremos un control total de la aplicación. Permite la realización de distintas operaciones y así como la configuración del proceso de extracción de parámetros tanto de la señal capturada como de la señal cargada desde un archivo.
- Representación de resultados: conforma el conjunto de gráficas que visualizan los resultados obtenidos una vez procesados los datos. Es totalmente redimensionable y se puede elegir que gráficas se desean mostrar a través de los paneles superiores. Así mismo permite la exportación de resultados a través de menús contextuales.
- Panel lateral: permite un rápido acceso al directorio actual para facilitar la carga de archivos. Así mismo incorpora un panel de control superior a través del cual se define el factor de comprensión/expansión del espectrograma así como los niveles de limitación y la tolerancia del detector de actividad sonora utilizado en el proceso de estimación de la frecuencia fundamental.

- Barra de captura de datos o reproducción de archivos: incluye unos controles para el inicio/pausa tanto del proceso de captura como de reproducción. Así mismo incluye una barra de progreso con el porcentaje de tiempo utilizado para el cálculo y visualización de parámetros en relación al tamaño del propio buffer de captura del dispositivo. Representa, además, de forma global el archivo de audio cargado, si lo hubiera.

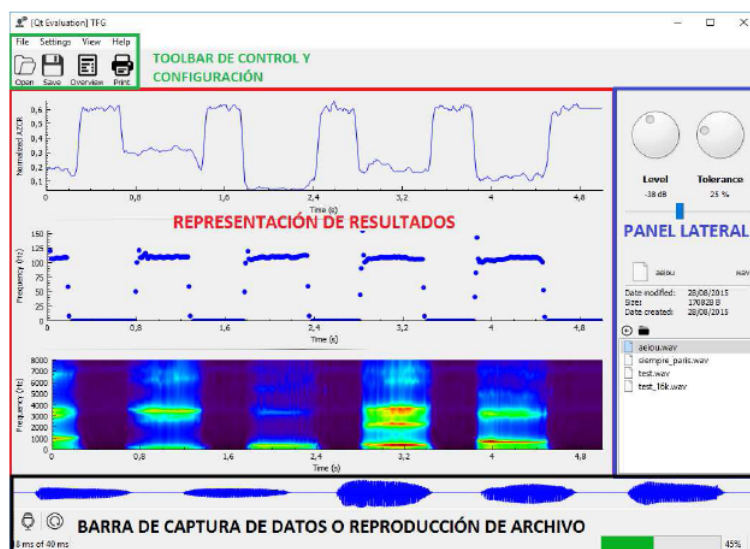


Figure 2: LogoSpeech Studio, secciones principales

0.1 Captura de datos

La captura de datos en esta aplicación se ha implementado para que sea fácilmente configurable. El usuario puede iniciar la captura al lanzar la aplicación a través del propio botón habilitado para esta labor, o a través de la barra principal. El proceso de captura es totalmente personalizable, el usuario puede decidir el dispositivo de captura deseado, el formato de salida del archivo capturado así como la frecuencia de muestreo o el codec utilizado. Estos parámetros son introducidos a través del dialogo de configuración del proceso de captura de datos, figura 3.

El usuario debe de guardar la configuración para reiniciar la captura y ser utilizada en futuras ocasiones. Así mismo siempre podrá volver a la configuración por defecto de la propia aplicación.

0.2 Conguración de procesado

Esta es sin duda la sección más importante de la aplicación. Es la que controla todo el proceso de tratamiento previo y extracción de parámetros de la señal de entrada. Viene estructurada en tres secciones bien diferenciadas:

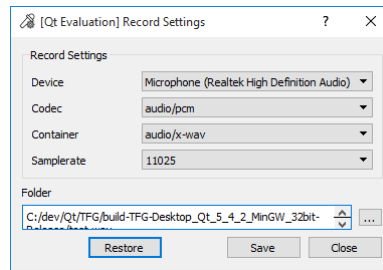


Figure 3: Dialogo para la configuración de captura de sonido

- Preprocesado: incluye la configuración de las técnicas de preprocesado desarrolladas. El usuario puede determinar si desea utilizar un filtro de compensación de tensión DC así como de realce de componentes de alta frecuencia. Por último y no por ello menos importante puede elegir la ventana que desea utilizar en el proceso de fragmentación de señal.

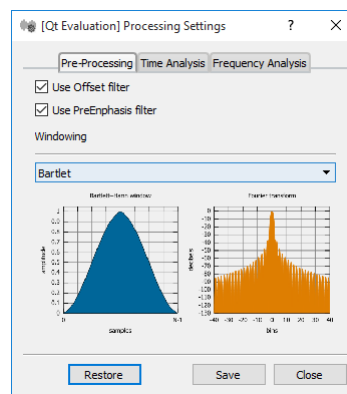


Figure 4: Dialogo para la configuración del preprocesado

- Análisis temporal: • incluye el dimensionado de las ventanas a utilizar así como el tamaño buffer global temporal que será visualizado para el usuario. Desde esta sección podremos configurar el tamaño de la ventana así como el tiempo de salto. Por otro lado permite la selección del parámetro a calcular en la gráfica de representación temporal (Energía o tasa de cruces por cero) así como la conjuración del método a utilizar para la estimación de la frecuencia fundamental así como el margen dinámico de la misma.
- Análisis de frecuencia: es el campo más importante. Permite la configuración de los distintos espectrogramas que pueden visualizarse. Permite definir el tamaño de las transformadas así como la propia técnica a utilizar para su visualización. Actualmente permite la estimación de distintos campos: :
 - Espectro real FFT
 - Espectro LPC, indicando el orden del mismo.

- Espectro Cepstral, indicando el número de muestra a utilizar.
- Espectro por banco de filtros, indicando el número de filtros que serán distribuidos en el rango de frecuencias a escala Mel.

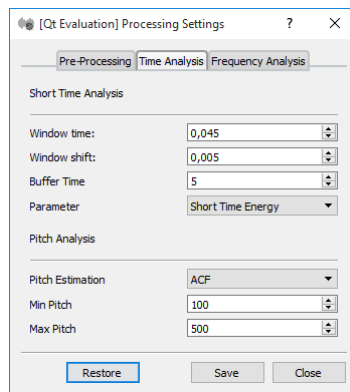


Figure 5: Dialogo para la configuración del análisis temporal

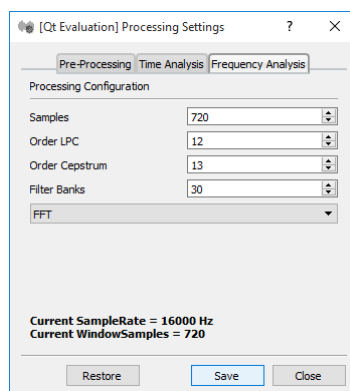


Figure 6: Dialogo para la configuración del análisis frecuencial

0.3 Configuración de gráficas

LogoSpeech Studio permite la configuración de los datos así como la forma en que estos se visualizarán. Para ello incorpora una sección dedicada a la personalización de las propias gráficas. A través de la misma se pueden elegir:

- El uso de ejes verticales/horizontales para la representación de la escala en que se visualizan los datos.
- El uso de un grids..
- La selección del color así como del grosor de la línea que visualiza los datos.

- La elección de la escala de colores utilizada para la representación del espectrograma.
- La elección del factor de expansión/compresión de los datos visualizados en el espectrograma.

Todo ello se integra en subsecciones fácilmente configurables. Cada una de ellas dedicada a cada una de las gráficas representadas. La capacidad de personalización es tal que nos permite seleccionar que gráficas deseamos visualizar así como un redimensionado arbitrario a gusto del propio usuario, figura 8.

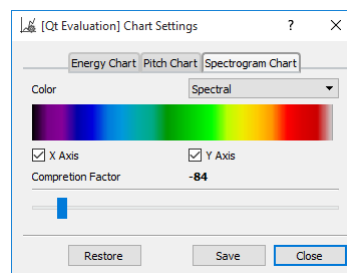


Figure 7: Dialogo para la configuración de la visualización del espectrograma

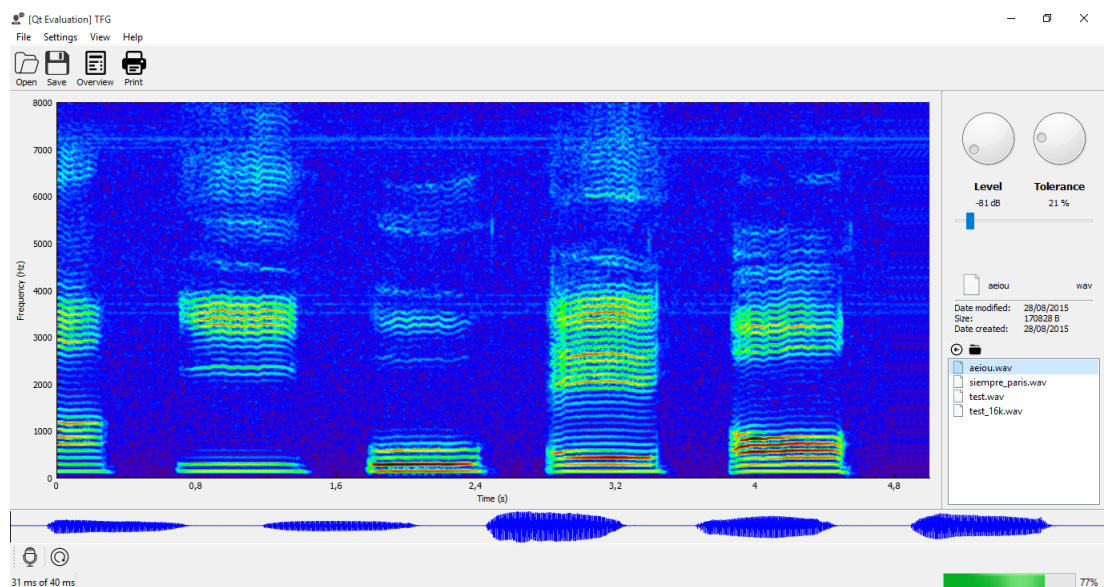


Figure 8: LogoSpeech Studio, visualización personalizada