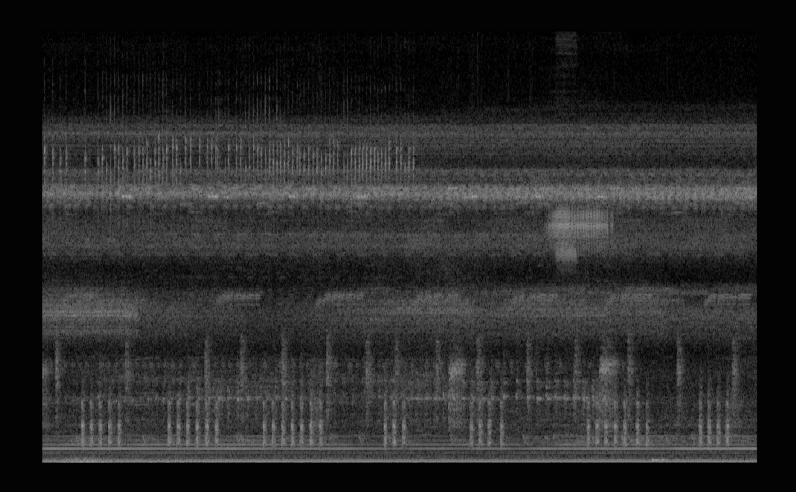


#### Instituto de Investigación de Recursos Biológicos Alexander von Humboldt

Investigación en biodiversidad y servicios ecosistémicos

# PROTOCOLO DE ANOTACIÓN MANUAL DE SEÑALES ACÚSTICAS PARA EL APRENDIZAJE AUTOMÁTICO EN GRABACIONES DE MONITOREO ACÚSTICO PASIVO





### Instituto de Investigación de Recursos Biológicos Alexander von Humboldt

Investigación en biodiversidad y servicios ecosistémicos para la toma de decisiones

## PROTOCOLO DE ANOTACIÓN MANUAL DE SEÑALES ACÚSTICAS PARA EL APRENDIZAJE AUTOMÁTICO EN GRABACIONES DE MONITOREO ACÚSTICO PASIVO

Elaborado por Maria Paula Toro Gómez Bióloga

Agosto

2022

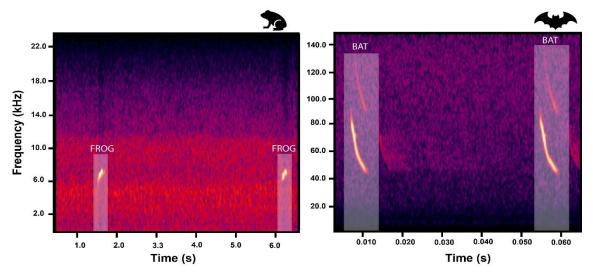
#### Tabla de contenido

INTRODUCCIÓN	3
INTRODUCCIÓNSOFTWARE Y MANEJOAudacity	4
Audacity	4
Abrir un archivo de audio	5
Visualización del espectrograma	6
Configuración del espectrograma	
Parámetros de configuración del espectrograma	8
Dividir canales de grabación	9
Como moverse en la grabación	
Antes de comenzar las anotaciones	
ANOTACIÓN MANUAL DE SEÑALES ACÚSTICAS	
Códigos de anotación	13
Indicaciones de calidad de señal	
Selección de la señal acústica	15
Creación de etiquetas	16
Exportar e importar etiquetas	17
REFERENCIAS	

#### INTRODUCCIÓN

El monitoreo acústico pasivo (MAP) ha sido considerado un método eficiente para evaluar el estado de la biodiversidad y los ecosistemas (Desjonquères et al. 2020). Este método consiste en la grabación autónoma de los sonidos presentes en un ambiente, a través de sensores acústicos autónomos resistentes a las condiciones exteriores, permitiendo recopilar datos acústicos en diferentes frecuencias (e.g., sonidos ultrasónicos) en una escala temporal y espacial mayor (Browning et al. 2017). Dado que el MAP permite detectar y extraer información sobre la presencia de especies a través de sus sonidos, este método se ha utilizado para realizar inventarios de especies, entender patrones de diversidad, monitorear especies de interés para la conservación e incluso detectar especies invasoras (e.g., Ulloa et al. 2016; Sugai et al. 2019). Si bien el monitoreo acústico pasivo brinda enormes beneficios (e.g. alta tasa de muestreo, mayor rentabilidad financiera), la gran cantidad de datos que genera supone un desafío en términos de procesamiento y análisis. El aprendizaje automático, aplicado en los sistemas de procesamiento de los datos ha facilitado en gran medida la extracción de información relevante para el análisis de grabaciones acústicas pasivas (Digby et al. 2013, Astaras et al. 2017, Ulloa et al. 2018).

En el aprendizaje automático los sonidos se clasifican utilizando diferentes algoritmos de aprendizaje. El aprendizaje supervisado consiste en algoritmos que se entrenan para clasificar muestras utilizando observaciones (i.e., archivo de audio) con sus respectivas etiquetas (i.e., categoría a la cual pertenece la observación). Las etiquetas o anotaciones son brindadas por expertos que revisan y anotan manualmente las observaciones, indicando cuales son los atributos que el modelo de aprendizaje automático debe aprender a reconocer. Una anotación o etiqueta, por ejemplo, puede indicar que una grabación incluye el canto de una rana o el pulso de un murciélago (Figura 1). Por el contrario, el aprendizaje no supervisado utiliza algoritmos que se entrenan para clasificar muestras, de acuerdo con la estructura presente en los mismos datos, es decir, sin etiquetas (Theobald 2017).



**Figura 1.** Representación gráfica de una anotación o etiqueta que indica a que animal pertenece la señal acústica que se visualiza en el espectrograma.

El éxito de muchos modelos de aprendizaje automático supervisado depende en gran medida de la calidad de los datos de entrenamiento que se utilizan. Para indicarle al ordenador cuales son los atributos debe aprender a reconocer, se utiliza la anotación o etiquetado de elementos. El conjunto de datos construido a partir de estas etiquetas es el principal componente de entrenamiento de los algoritmos de aprendizaje automático. Así, el sistema de aprendizaje podrá analizar el resto de las grabaciones y encontrará los sonidos que ya habían sido previamente identificados. Una anotación o etiquetado deficiente puede dar lugar a un modelo de aprendizaje propenso a errores. Por lo tanto, para que los algoritmos aprendan de forma eficiente y que, además, puedan reconocer patrones similares en el futuro, la anotación realizada en los datos debe ser precisa, consistente y relevante para la tarea que se le pide a la máquina.

Uno de los aspectos más importantes para generar datos de entrenamiento de calidad es la consistencia en la anotación o etiquetado de datos, esto es, asegurarse de que todas las etiquetas de los elementos estén etiquetadas de la misma manera y bajo los mismos criterios. Así se reducen ambigüedades que pueden alterar el comportamiento del modelo de aprendizaje y hacer que su evaluación sea inexacta. Para garantizar etiquetas consistentes es necesario establecer aquellos criterios, umbrales y pautas de anotación que sirvan de referencia para los expertos que realizan el proceso de anotación en proyectos de aprendizaje automático. Este protocolo describe el procedimiento paso a paso para realizar anotaciones detalladas de señales acústicas en grabaciones de monitoreo acústico pasivo. Además, presenta las bases que se utilizan para realizar un etiquetado, que se pueda emplear como referencia en futuros proyectos que incluyan aprendizaje automático para el análisis de grabaciones de monitoreo acústico pasivo.

#### SOFTWARE Y MANEJO Audacity

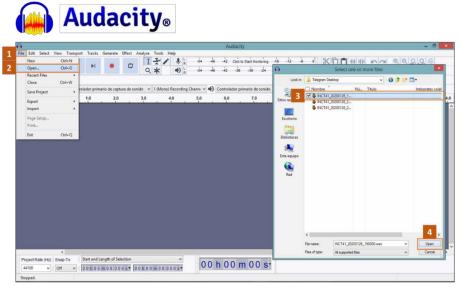
Para anotar señales acústicas en grabaciones de audio, se requiere un software que permita la visualización, edición y análisis de sonido; y que, además, permita crear y exportar etiquetas. **Audacity** es un software gratuito de código abierto para la edición y grabación de audio disponible en múltiples plataformas (Windows, macOS, GNU/Linux y otros sistemas operativos). En este protocolo se describe el proceso de anotación manual utilizando este software, ya que permite desarrollar un método estandarizado gratuito y su interfaz intuitiva facilita su uso para realizar anotaciones acústicas. Para instalar el software Audacity, diríjase a "audacityteam.org" y descargue el instalador para su sistema operativo (Figura 2). Puede encontrar información completa sobre Audacity en el manual del usuario del programa (<a href="https://manual.audacityteam.org/index.html">https://manual.audacityteam.org/index.html</a>)



Figura 2. Página web del programa Audacity donde se puede descargar el instalador del programa

#### Abrir un archivo de audio

Para abrir el archivo de audio que contiene la grabación, diríjase al menú *File* + *Open* y seleccione desde la ventana emergente la grabación en la cual va a realizar las anotaciones. Asegúrese que el archivo de audio se encuentre en un formato compatible (e.g WAV). A continuación, seleccione *Open* en la ventana emergente (Figura 3). Una manera alterna de realizar este proceso es arrastrar el archivo de audio y soltarlo dentro de la ventana de Audacity.



**Figura 3.** Interfaz principal del programa Audacity. Los números representan el paso a paso para abrir un archivo de audio.

Una vez abierto el archivo de audio, podrá visualizar el oscilograma en modo estéreo, es decir, los canales de grabación izquierdo (arriba) y derecho (abajo) (Figura 4)

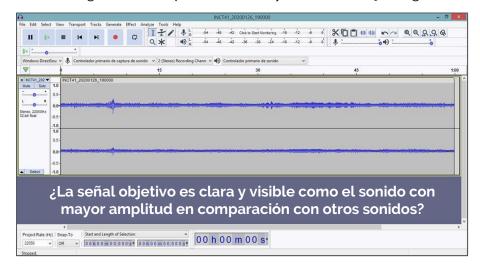


Figura 4. Oscilograma de una grabación acústica pasiva

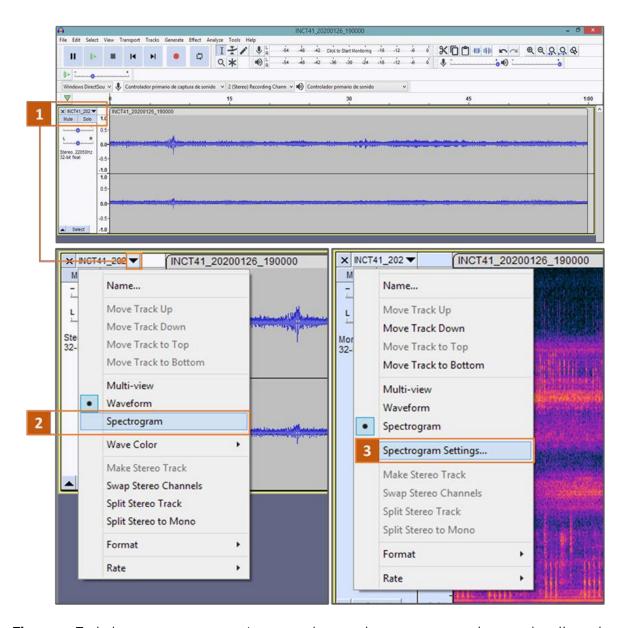
En grabaciones acústicas pasivas, el oscilograma puede no ser el adecuado para realizar anotaciones de las señales acústicas, debido a que la señal acústica objetivo puede presentar superposición con otros sonidos, como el ruido de fondo, los sonidos de insectos o llamadas de heteroespecíficos. Por esta razón, para realizar anotaciones de señales acústicas en este tipo de grabaciones se utilizará el espectrograma. Además, el uso del espectrograma brinda información visual que facilita la identificación de señales, ayudando a obtener anotaciones más estandarizadas.

#### Visualización del espectrograma

Para visualizar el espectrograma de la grabación, despliegue el menú de pistas de audio, haciendo clic en el triángulo invertido [▼] que se encuentra junto al nombre del archivo (Figura 5). En el menú desplegable, seleccione *Spectrogram*. Una vez se encuentra en la vista de espectrograma, se activa la opción *Spectrogram settings* (Configuración de espectrograma) en la entrada del menú despegable, esta opción le permite le permite ajustar algunos parámetros de visualización del espectrograma.

#### Configuración del espectrograma

El espectrograma proporciona información sobre la distribución de frecuencias de los sonidos a lo largo del tiempo, y la intensidad con la que se presenta un color representa la energía con la que se produce el sonido. En Audacity puede realizar ajustes a la configuración del espectrograma de la siguiente manera: seleccione la opción *Spectrogram settings* (Configuración de espectrograma) en la entrada del menú despegable de la grabación (Figura 5).



**Figura 5**. En la imagen se representan con números los pasos a seguir para visualizar el espectrograma de una grabación.

En la ventana emergente (Figura 6) podrá encontrar una serie de parámetros que podrá ajustar para configurar la visualización del espectrograma y facilitar la anotación o etiquetado. Los parámetros de visualización de la configuración de espectrograma (*Spectrogram settings*) incluyen *Escala* (Scale), *Colores* (Colors) y *Algoritmos* (Algorithms). Cada uno de ellos modifica la visualización de un aspecto del espectrograma. Aquí se presentan (resaltados Figura 6) la configuración establecida para una adecuada visualización de grabaciones de monitoreo acústico pasivo para realizar anotaciones en el espectro audible tenga en cuenta que estos parámetros puedan variar de acuerdo con el objetivo de cada proyecto. Estos parámetros se deben ajustar al principio y dejar fijos para mantener consistencia al momento de realizar las anotaciones.

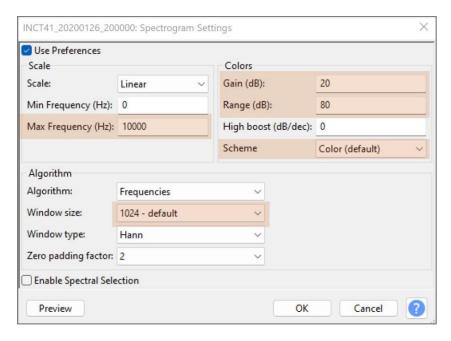


Figura 6. Ventana de configuración de parámetros de visualización del espectrograma

#### Parámetros de configuración del espectrograma

#### Escala (scale)

Esta opción permite modificar la escala vertical del espectrograma, es decir la escala en la que se presenta la frecuencia de los sonidos de la grabación. Se recomienda usar una escala lineal.

Frecuencia máxima y Frecuencia mínima (Min Frequency and Max Frequency) El ajuste de este parámetro permite enfocar la atención de búsqueda visual a una franja del espectrograma. Este valor corresponde a los límites de la escala vertical, es decir la frecuencia. Para grabaciones pasivas y análisis en espectro audible se recomienda ajustar según los límites de frecuencia de los sonidos de interés. Por ejemplo, si se requiere anotar la presencia del llamado de un anfibio con canto entre 2 y 5 kHz, se puede ajustar el valor de Escala con: una Min Frequency (Hz) de 1 kHz y Max Frequency de 7 kHz, Las frecuencias fuera de este rango no serán visibles.

#### Colores (Colors)

El uso de colores para representar la energía del sonido en espectrogramas es una cuestión de gustos, sin embargo, asegúrese de que los ajustes de visualización le permitan visualizar de forma clara e informativa las señales acústicas presentes en la grabación. Los parámetros que debe ajustar aquí son:

Ganancia (Gain in dB): Esta opción le permite aumentar/disminuir el brillo de la pantalla. Para señales pequeñas donde la pantalla es principalmente (oscura) puede aumentar este valor para ver colores más brillantes y ver más detalles. Para grabaciones acústicas pasivas y

análisis en espectro audible el valor establecido para ganancia es de 20 dB.

Rango (Range in dB): Esta opción establece el rango de tamaños de señal que se mostrarán como colores, y a su vez el valor mínimo de intensidad que se mostrará en el espectrograma. El valor por defecto de este parámetro es de 80 dB, las señales a partir de este valor se escalonarán según la paleta y espaciado de colores.

Esquema (Scheme) Aquí puede escoger entre dos formatos de color o dos configuraciones en escala de grises. Tenga en cuenta que la paleta de color "clásica" puede ser poco sensible para personas daltónicas.

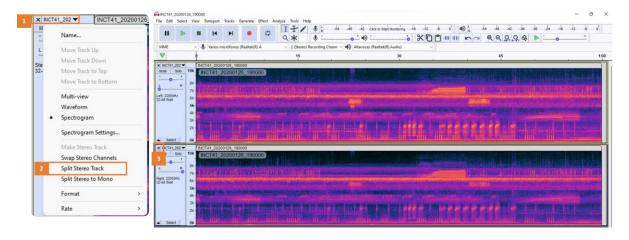
#### Algoritmo (Algorithm)

Tamaño de la ventana (Window size)-Puede elegir el tamaño de ventana (Transformación rápida de Fourier- FFT), con el que desea visualizar la grabación, este parámetro afecta la cantidad de detalles verticales (en frecuencia) que puede visualizar. Un tamaño de ventana FFT con valores más grandes aporta mayor resolución de baja frecuencia, pero menor resolución temporal. En la práctica el ancho de ventana FFT se debe ajustar según la resolución temporal y de frecuencias óptima para el tipo de grabación que se está analizando. Por ejemplo, un FFT de 1024 funciona bien para grabaciones con frecuencia de muestreo de 44100 o 48000 Hz, permitiendo visualizar de manera clara las señales acústicas de varios grupos taxonómicos.

Para aplicar los cambios en los ajustes de configuración del espectrograma, debe presionar **OK**. Tenga en cuenta que estos cambios en visualización solo se aplican en la grabación abierta y permanecerán solo mientras la ventana del proyecto esté abierta. Para realizar cambios permanentes en la configuración predeterminada del espectrograma, utilice *Preferencias de espectrograma*, de esta manera cada vez que abra un nuevo archivo la configuración establecida previamente estará predeterminada. Para establecer las preferencias diríjase a **Edit + Preferences + tracks + spectrograms**, en esta ventana ajuste los parámetros y seleccione **OK**. Puede reestablecer la configuración del espectrograma en cualquier momento.

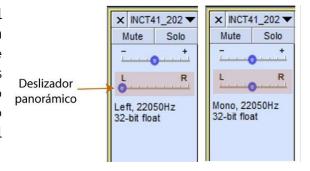
#### Dividir canales de grabación

 con un borde amarillo, para excluir uno de los dos canales seleccione  $\mathbf{X}$  en el mismo menú de la grabación (Figura 7).



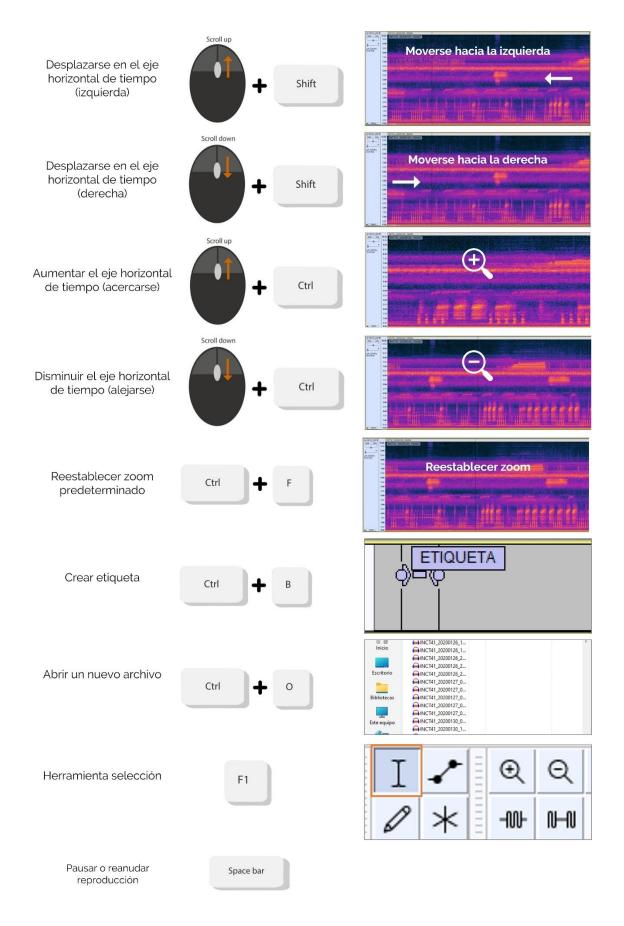
**Figura 7.** Pasos para dividir los canales de una grabación estéreo. El paso tres (3) representa como excluir o cerrar alguno de los canales.

Cuando ha dividido una grabación, el deslizador panorámico del canal de grabación que ha conservado, mantiene su dirección de canal inicial, lo que hace que la señal sea más fuerte en el altavoz/ auricular izquierdo o derecho según sea el caso, para evitar esto asegúrese de posicionar el deslizador en el centro.

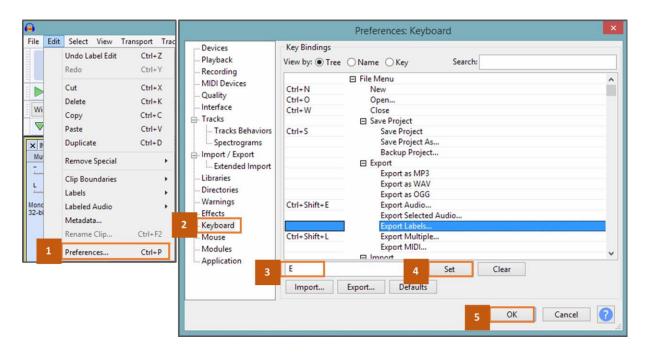


#### Como moverse en la grabación

Ahora es momento de moverse dentro de la grabación, Audacity tiene una serie de atajos y accesos directos que facilitan acciones como moverse, acercarse, alejar, pausar la grabación, entre otras. Aquí se presentan algunas de las funciones más útiles al momento de realizar anotaciones.



Igualmente, Audacity permite agregar atajos a su teclado para funciones específicas, para hacerlo realice lo siguiente: *Edit+Preferences+Keyboard*. En la ventana que se muestra podrá buscar la función a la cual desea agregar el atajo, en este caso vamos a agregar un atajo a la función "export labels" con la asignación del teclado "*E*", a continuación, presione el botón *set* y luego presione *OK*. De esta manera cuando ya tenga las etiquetas podrá acceder al menú de exportación de manera rapida tan solo presionando la tecla "*E*" (Figura 8).



**Figura 8.** Esquema ilustrando como agregar atajos a funciones específicas, los números representan el paso a paso.

#### Antes de comenzar las anotaciones

En este momento la grabación se encuentra lista para realizar el etiquetado de las señales acústicas objetivo. Antes de empezar el proceso de anotación tenga en cuenta las siguientes recomendaciones.

- Escuche la grabación una vez por completo, mientras visualiza de cerca al espectrograma, antes de realizar cualquier etiqueta.
- Tenga a la mano un canto de referencia de la señal objetivo que se le ha encargado etiquetar.
- Investigue los tipos de canto y frecuencias en los que puede emitir señales acústicas su especie objetivo.
- Utilice audífonos al momento de etiquetar los datos, esto le permitirá tener una mayor resolución al momento de identificar la señal objetivo.
- Tenga en cuenta la información asociada a esta grabación, por ejemplo, especies registradas, fecha, localidad, hábitat, otras especies en la grabación, etc. Lo anterior

permitirá detectar artefactos o problemas de grabación en esta etapa.

#### ANOTACIÓN MANUAL DE SEÑALES ACÚSTICAS

#### Códigos de anotación

Al momento de realizar anotaciones es importante definir la terminología y códigos que serán utilizados para etiquetar los elementos. En este caso la señal acústica será anotada de acuerdo con el nombre científico de cada especie, utilizando la siguiente codificación:

Se asignarán las tres primeras letras del género, seguido de las tres primeras letras del epíteto, utilizando siempre mayúscula sostenida.

Por ejemplo:

Physalaemus cuvieri se codificará como PHYCUV.

Scinax ruber = **SCIRUB** 

<u>Den</u>dropsophus <u>cru</u>zi = **DENCRU** 

Andinobates bombetes = **ANDBOM** 

Para aquellas especies cuyas letras coincidan entre sí se asignará a una de las especies la cuarta o quinta letra consecutiva según corresponda, para evitar errores de asignación en las etiquetas.

Por ejemplo:

Boana albopunctata se codificará como BOAALB

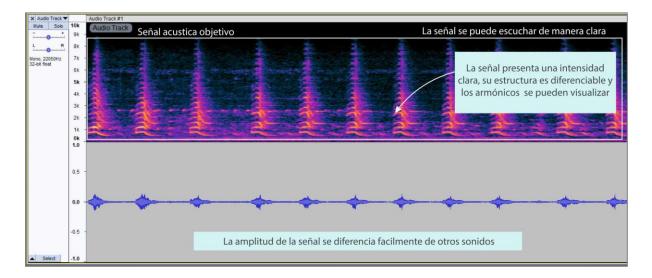
Boana albomarginata = **BOAALM** 

#### Indicaciones de calidad de señal

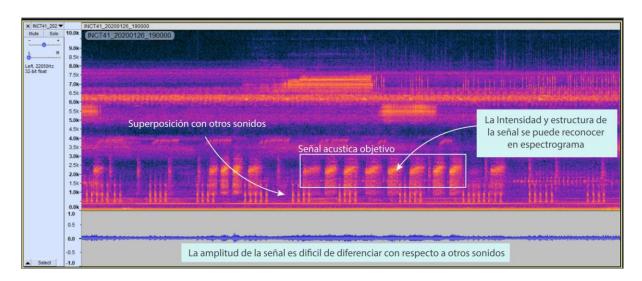
Muchos factores pueden alterar la calidad de una grabación y a su vez de las señales acústicas presentes en ellas. Por ejemplo, el ruido de fondo, la lluvia, un río que fluye rápidamente, el viento, emisiones de sonido de otros animales, distancia del emisor a la grabadora. Por lo tanto, para mejorar la calidad de las anotaciones, incrementar la usabilidad de los datos y mejorar el análisis de errores del modelo de aprendizaje, se incluirá en el etiquetado la calidad con la que se presenta la señal objetivo de acuerdo con los siguientes

criterios de calidad:

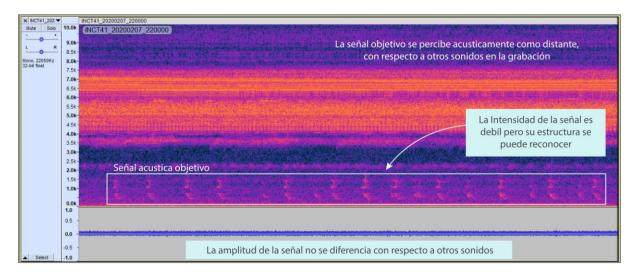
**CLEAR** ("C") Una señal acústica *clear* o clara es aquella en la que un solo individuo está emitiendo la señal, se puede escuchar con claridad y bien separada de otros sonidos. En el espectrograma se puede identificar la estructura de la señal con una mayor intensidad acompañada de sus armónicos. En el oscilograma la amplitud de una señal acústica puede visualizarse fácilmente con respecto a los demás elementos en la grabación.



**MEDIUM** ("M") Una señal acústica *medium* es aquella que se puede escuchar e identificar visualmente en el espectrograma, pero que presenta superposición con otros sonidos bióticos o abióticos. La intensidad de la señal puede ser similar a la de otros elementos en la grabación, pero la estructura de la señal se puede identificar. En el oscilograma es difícil identificar su amplitud, y puede ser emitida por uno o varios individuos.



**FAR** ("F") Una señal acústica *far* es aquella en la que se puede identificar un solo individuo emitiendo una señal que se puede escuchar distante y estar parcialmente enmascarada con otros sonidos. En el espectrograma la señal se representa con poca intensidad, pero su estructura se puede identificar. Una señal acústica *far* es difícil de visualizar en el oscilograma.



Para incluir la calidad de la señal en la anotación, después del código de la especie, adicione un guion al piso ("\_") seguido de la letra correspondiente a la calidad de la señal (*i.e.*, "C", "M", "F"). De esta forma una etiqueta completa quedaría de la siguiente forma:

PHYCUV\_F

PHYCUV\_M

PHYCUV\_C

#### Selección de la señal acústica

Una vez identificada una señal acústica para una especie objetivo deberá seleccionarla para agregar la etiqueta. Para esto, escoja la herramienta **seleccionar** de la barra de herramientas (Figura 9) o acceda mediante su atajo en este caso la tecla "**F1**". Amplíe la sección para conocer la estructura de la señal y seleccione incluyendo parte del ruido de fondo, como se muestra en la figura 9.

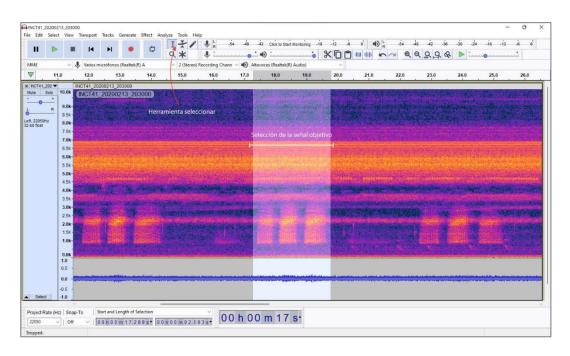


Figura 9. Selección de la señal acústica de una especie objetivo.

#### Creación de etiquetas

Una vez seleccionada la señal podrá agregar la etiqueta de la siguiente manera, diríjase a *Edit + labels + add labels at selection* o bien acceda desde el acceso directo con la combinación de teclas *Ctrl + B.* Inmediatamente le aparecerá bajo el espectrograma una nueva sección llamada *label track* en esta sección usted podrá escribir y etiquetar su selección según corresponda (Figura 10).

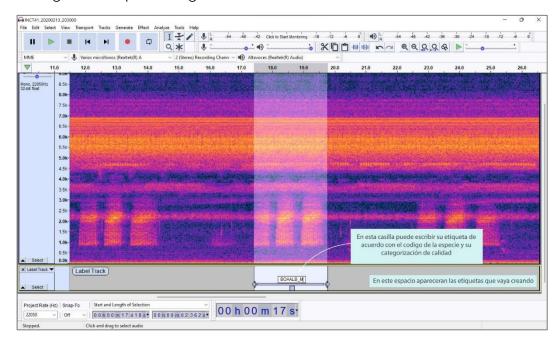
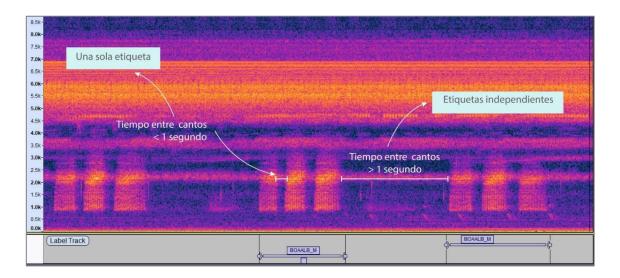


Figura 10. Sección de creación de etiquetas (Label track)

**NOTA**: Tenga en cuenta que su etiquetado solo se encuentren los caracteres asociados al código de la especie y calidad de la señal, asegúrese de que no tenga espacios, puntos u otro carácter que pueda generar inconsistencias.

Para facilitar el procedimiento de anotación, las señales acústicas de una misma especie se anotan en conjunto si la duración entre vocalizaciones no es mayor a un segundo (1s), tal y como se muestra en la figura 11. Si las señales acústicas de una misma especie están separadas por más de un segundo deben dividirse y etiquetarse de manera independiente.



**Figura 11**. Criterio de selección de señales y creación de etiquetas de acuerdo con el tiempo entre señales.

Al momento de buscar las estructuras y seleccionarlas tenga presente lo siguiente:

- Las señales distorsionadas (recortadas o distorsionadas por el micrófono) no se anotan.
- Solo se anotan las señales que pueden escucharse y observarse en el espectrograma.

De acuerdo con la especie objetivo, la señal acústica podría ser una sola llamada, o una serie corta o larga de llamadas, por lo que la selección depende de la estructura de la señal.

#### Exportar e importar etiquetas

Cuando categorice y etiquete según corresponda las señales acústicas a lo largo de la grabación, puede exportar las etiquetas en formato de texto (.txt)

Acceda al menú *File + Export + Export labels*, y guarde el archivo con el mismo nombre de la grabación (Figura 12). El nombre de la grabación lo puede copiar accediendo al menú de pistas de la grabación y seleccionado *Name*, y luego pegarlo como nombre de archivo al momento de exportar las etiquetas.

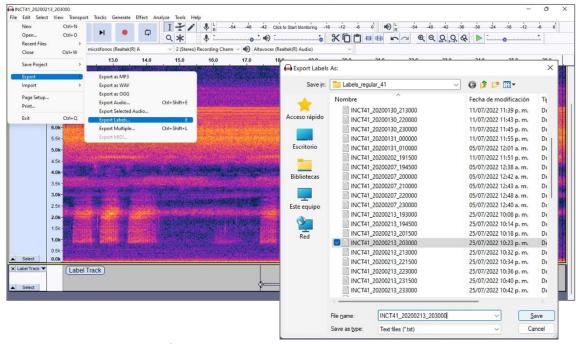


Figura 12. Ruta para exportar etiquetas

Una vez ha exportado su archivo de etiquetas, quedará guardado en la ubicación asignada en su ordenador. Este archivo contendrá el tiempo inicial y final de la señal acústica anotada y a la cual ha asignado especie y calidad de la señal.

*INCT41_20200126_200000: Bloc de notas		
Archivo	Editar Ver	
1.062793 2.928585 6.671977	1.558763 4.829803 7.250609	BOAALB_M BOAALB_M BOAALB_F

Para visualizar un archivo de etiquetas, debe abrir el archivo de audio y dirigirse a **Archivo+ importar+ etiquetas**. Una vez seleccionado el archivo .txt que contiene las etiquetas, podrá visualizarlas en la grabación (Figura 13). Tenga en cuenta que solo se visualizará correctamente las etiquetas correspondientes al audio seleccionado.

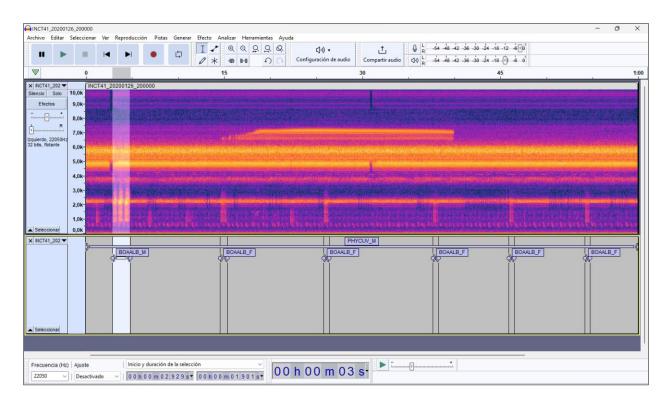


Figura 13. Visualización de etiquetas de una grabación acústica pasiva.

#### **REFERENCIAS**

Astaras C, Linder JM, Wrege P, Orume RD, Macdonald DW. 2017. Passive acoustic monitoring as a law enforcement tool for Afrotropical rainforests. *Frontiers in Ecology and the Environment* 15: 233–234.

Browning, E., Gibb, R., Glover-Kapfer, P., & Jones, K. E. (2017). Passive acoustic monitoring in ecology and conservation. *WWF Conservation Technology Series*, 1(2), 1–75.

Desjonquères, C., Gifford, T., & Linke, S. 2020. Passive acoustic monitoring as a potential tool to survey animal and ecosystem processes in freshwater environments. *Freshwater Biology*, 65(1), 7-19.

Digby A, Towsey M, Bell BD, Teal PD. 2013. A practical comparison of manual and autonomous methods for acoustic monitoring. *Methods in Ecology and Evolution* 4: 675–683.

Sugai, L, Silva T, Ribeiro J, & Llusia, D. 2019. Terrestrial passive acoustic monitoring: review and perspectives. *BioScience*, 69(1), 15-25.

Theobald, O. 2017. *Machine learning for absolute beginners: a plain English introduction* (Vol. 157). Scatterplot press.

Ulloa JS, Aubin T, Llusia D, Bouveyron C, Sueur J. 2018. Estimating animal acoustic diversity in tropical environments using unsupervised multiresolution analysis. *Ecological Indicators* 90: 346–355