



# SELECCIÓN, ANOTACIÓN Y ETIQUETADO DE LOS CANTOS DE LAS RANAS DEL ECUADOR

Andrés Estrella<sup>1</sup>, Daniel Nicolalde<sup>2</sup>, Damián Nicolalde<sup>3</sup>, Daniela Pareja<sup>4</sup>

## RESUMEN

El presente trabajo explica el estándar y prácticas de selección, anotación y etiquetado del material bioacústico para grabaciones que contengan cantos de ranas. El objetivo es disponer de una base de datos de señales acústicas de cantos de ranas presentes en el Parque Nacional Yasuní en Ecuador considerando que los cantos anotados son una valiosa fuente de información para investigación de reconocimiento automático de especies. Este documento explica el procedimiento de selección de audio en buen estado, así como el procedimiento de anotación. En la selección de audio se evitarán sonidos con baja relación-señal-a-ruido, con presencia de voz humana, con ruidos mecánicos y señales satura-

das. En el procedimiento de anotación, se usan etiquetas estándar para marcar el intervalo de anotación, los cantos de rana por especie y cada nota. Esta base de datos servirá como referencia para desarrollo de aplicaciones bioacústicas, bioinformáticas y de procesamiento digital de señales en tareas que requieran reconocimiento automático de cantos de rana.

**Palabras clave:** cantos de ranas; bioacústica; anotar cantos; etiquetar cantos; procesamiento digital de señal

## ABSTRACT

The present work explains the standards and practices of selection, recording and labeling of bioacoustic

---

<sup>1</sup> (andrew.srocks@yahoo.com).

<sup>2</sup> (danielnicolalde@gmail.com).

<sup>3</sup> Pontificia Universidad Católica del Ecuador, Facultad de Ingeniería, Quito, Ecuador (danicolalde@puce.edu.ec).

<sup>4</sup> Pontificia Universidad Católica del Ecuador, Facultad de Ciencias Exactas y Naturales, Quito, Ecuador (dparejameja@gmail.com).

material found in recordings of frog calls in the wild. The goal is to have a consistent audio database of annotated frog calls found in Yasuní National Park in the eastern region of Ecuador for development and testing of automatic recognition algorithms. This document describes the procedure for audio selection as well as the recording procedure. Signals with low signal-to-noise-ratio, presence of human voice, mechanical noises, and clipped are discarded during

the selection step. In the annotation task, standard labels are chosen to mark the annotation interval, a frog call, and each note. This database is expected to serve as ground-truth for the development of automatic species recognition applications in bioacoustics, bioinformatics, and digital signal processing.

**Key words:** songs of frogs; bioacoustics; record songs; tag songs; digital signal processing

## INTRODUCCIÓN

Ecuador es hábitat de 566 especies de ranas conocidas (Ron, S. R., Guayasamín, J. M., Yáñez-Muñoz, M. H., Merino-Viteri, A. Ortiz, D. A. y Nicolalde, 2016). Sin embargo, nuevas especies se van descubriendo constantemente. Los machos cantan para atraer a las hembras con fines reproductivos y también para defender su territorio. Los cantos son compuestos de una o varias notas como se observa en Figura 1. Por sus características particulares, los investigadores consideran los cantos de rana como una valiosa fuente de información para identificación y clasificación de especies. Específicamente, se ha reportado el uso de características del canto discriminatorias entre diferentes especies. Estas características son: frecuencia fundamental y dominante, número de pulsos, número de armónicas, duración del canto, tiempo de subida, etc (Duellman, W. E., 1986).

En general, el procedimiento de obtención de características de canto de rana se realiza de forma manual y con auxiliares de software que permiten extraer las características de cada canto, que serán usadas por un clasificador estadístico. En muchos casos, la complejidad del hábitat de las ranas permite registrar únicamente el canto de las especies.

Este trabajo brindará un protocolo que les permita a los investigadores extraer las características de los cantos de las ranas para futuros análisis, por ejemplo: investigar el desempeño de modelos de combinación de gaussianas (GMM) (Reynolds & Rose, 1995), un método clásico de reconocimiento de locutor, para reconocimiento de especies de rana.

Para el proceso de reconocimiento de las especies de ranas, se consideran



dos fases: entrenamiento y pruebas. En la fase de entrenamiento se estiman los parámetros de los modelos de cada rana en base a su registro acústico anotado, y en la fase de pruebas se evalúa el desempeño del algoritmo a través de métricas (Jaafar & Ramli, 2013). En ambos casos es necesario contar con un corpus de evaluación anotado y etiquetado cuidadosamente.

El objetivo de este documento es proveer instrucciones detalladas para anotar un archivo de audio que contiene el canto de rana y producir un corpus de evaluación (*ground truth*) consistente y acertado para investigación. Este documento también va a servir

como referencia a investigadores usando los datos anotados. Los procedimientos mostrados de aquí en adelante fueron desarrollados usando como referencia los métodos propuestos en (Kasuri et al., 2009) (Jaafar & Ramli, 2013) y recogiendo la experiencia de los investigadores en el proyecto.

El documento está organizada de la siguiente forma. En la sección II se analiza la selección de audio que se usará para la anotación. Posteriormente, en la sección III (A, B, C, D, E, F, G, H, I, J, K, L) se explica el procedimiento de anotación y etiquetado de audio. Finalmente, en la sección IV se emitirán las conclusiones.

## MATERIALES Y MÉTODOS

### Selección de audio

El hábitat de la rana es un ambiente acústico muy complejo y diverso donde varias especies compiten por ser escuchadas con fines reproductivos y territoriales (Duellman, W. E., 1986). Además, las ranas cantan en coros compuestos de varios individuos de la misma especie y durante la caída de lluvia (Jaafar & Ramli, 2013)(Taylor, Watson, Grigg, & McCallum, 1996). También existen diferencias en las prácticas de grabación seguidas por cada investigador. Algunas veces se pueden introducir ruidos mecánicos en el procedimiento de graba-

ción. En otras palabras, existen fuentes de ruido que degradan la calidad del audio grabado en campo.

Debido a esto, es necesario realizar una selección del material acústico que se va a usar para entrenar el sistema automático antes de realizar la anotación y etiquetado del mismo. La selección se hace en base al análisis visual del espectrograma y escuchando cada segmento para asegurar buena calidad del corpus de evaluación.

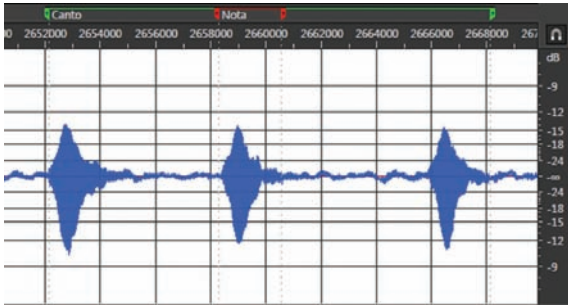


Figura 1: Un canto de *H. lanciformis* compuesto de 3 notas (sílabas)

#### a. Registro acústico de cantos de ranas

Es importante seguir un método sistemático de grabación de audio en el campo para generar muestras homogéneas y de calidad suficiente para análisis e inventario de material bioacústico (Brandes, 2008). Un manual de grabación bioacústica propuesto por *Obritz et al.* está disponible en («Bioacoustic approaches in biodiversity inventories», 2010).

#### b. Herramientas de software

Para selección, anotación y etiquetado de los cantos de ranas se ha escogido usar el software Adobe Audition 6.0 en combinación con la herramienta de edición CueListTool 1.7 («CueListTool 1.7», s. f.) para generar el archivo de intercambio.

#### c. Formato de audio

Para inventario y análisis de material bioacústico es necesario usar audio digital de alta calidad sin pérdidas («Bioacoustic approaches in biodiversity inventories», 2010). En particular, se van a usar archivos en formato *Waveform Audio File Format WAVE* («Multimedia Programming Interface and Data Specifications 1.0», 1991) con extensión .WAV. La frecuencia de muestreo a partir de 44100 Hz y con resolución de 16-bits en adelante. No usar archivos en MP3.

#### d. Tipos de ruido por evitar en archivo de audio

En el audio seleccionado para entrenar el sistema automático, es importante evitar perturbaciones que introduzcan errores en los modelos estocásticos. Primero, impedir grabaciones de rana con conversación de fondo. En Figura 2

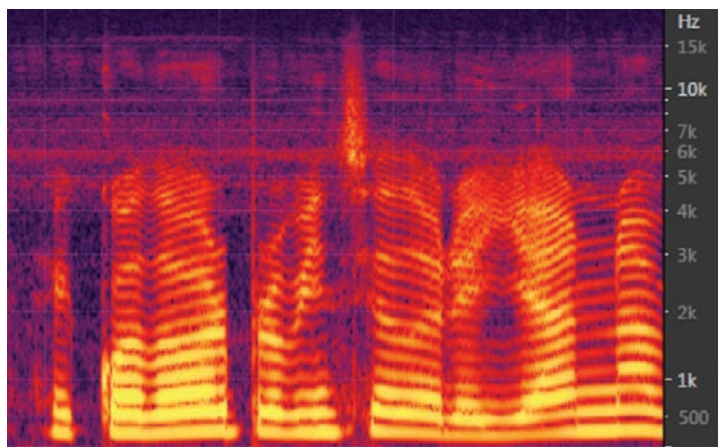


se puede ver un segmento de espectrograma típico de voz humana. La voz concentra energía en una zona del espectro típicamente usada por las ranas por lo que es necesario seleccionar archivos sin esta clase de interferencia.

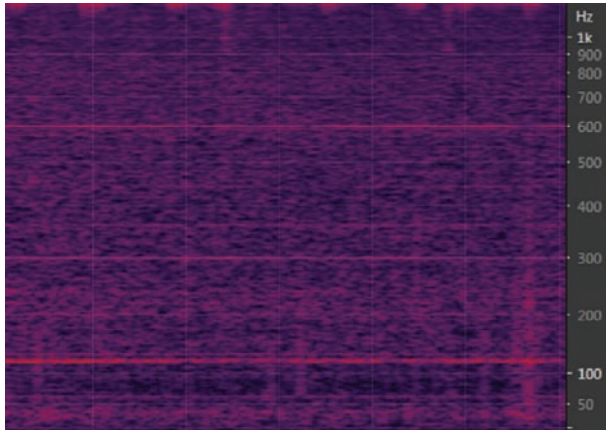
Segundo, omitir archivos con zumbido en 60 Hz o con sonido de motores. Las armónicas de estas señales pueden también interferir en los cantos. En la figura 3 se puede ver un espectrograma típico de un generador AC de fondo con sus armónicas durante la grabación. Ruido de este tipo puede ser tratado con filtros de audio. Sin embargo, el foco

de este trabajo es en análisis sin profundizar en restauración. Evitar archivos con este ruido porque introduce tendencias erróneas en los modelos de rana.

Tercero, no escoger grabaciones con baja relación de señal a ruido (SNR). Esta situación se genera cuando el canto de rana ha sido grabado a distancia y es apenas perceptible por sobre el ruido de fondo. Para entrenar los modelos de rana se busca segmentos donde el canto de rana sea claramente más fuerte que el ruido de fondo. En lo posible usar  $SNR > 20$ .



**Figura 2: Espectrograma típico de voz humana**



**Figura 3: Espectrograma de un generador AC en 60 Hz con sus armónicas**

Cuarto, evitar grabaciones que presenten artefactos mecánicos como golpes en el micrófono, fricción o mala conexión del cable. En la figura 4 se presenta el espectrograma de golpe y fricción en el micrófono. Su espectro interfiere con los cantos de rana por lo que su presencia degrada significativamente la calidad del material.

Finalmente, no escoger archivos que presenten *clipping*. Descartar grabaciones que presenten saturación que

produce distorsión del canto. La figura 5 muestra un ejemplo de canto saturado. Se recomienda escoger cantos que no sobrepasen el nivel -6dB que es la pérdida esperada de presión acústica a 1 metro de distancia de la rana con un micrófono correctamente calibrado («Bioacoustic approaches in biodiversity inventories», 2010).

El material seleccionado se anota y etiqueta siguiendo el procedimiento detallado en la siguiente sección.

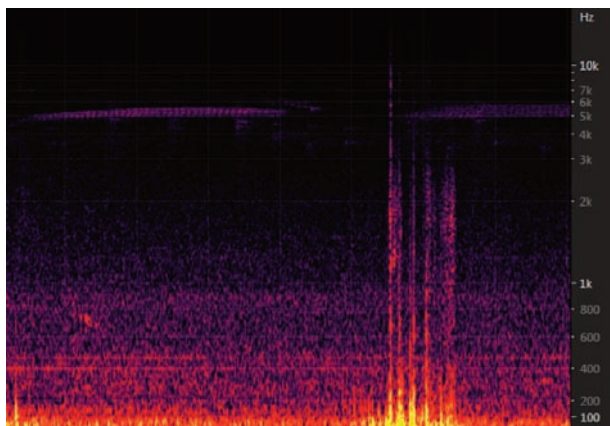


Figura 4: Allobates insperatus con ruido de golpes en micrófono y generador AC de fondo

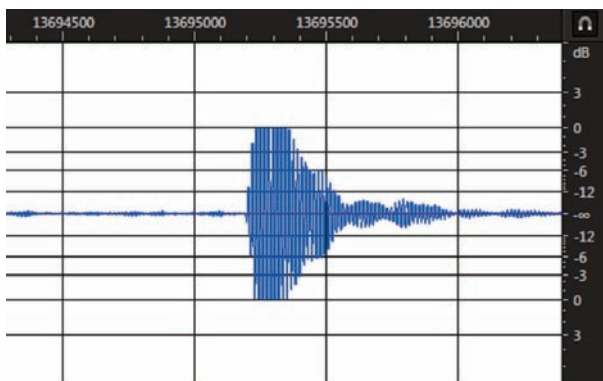


Figura 5: Canto de rana con clipping

### Anotación y Etiquetado de Audio

Una vez que el corpus de evaluación ha sido seleccionado es necesario

anotar y etiquetar los cantos de acuerdo con las características de cada especie. El corpus se anota entre 5 especialistas en identificación manual con un *cross-*

check del 10% del material anotado por cada uno. A continuación se define el dominio y la identificación única por cada especie de rana. Después se introducen la nomenclatura y el procedimiento de anotación e intercambio.

### e. Matriz de Anotación

Se define un dominio "Yasuní" con 128 especies de ranas. Para el estudio de reconocimiento automático de especie de ranas se escogen inicialmente 10 especies que se muestran en tabla 1.

**Tabla 1: Lista de ranas escogidas**

FAMILIA	GÉNERO	CÓDIGO
Bufo	<i>Rhinella margaritifera</i>	R0009
Craugastoridae	<i>Pristimantis conspicillatus</i>	R0023
Hylidae	<i>Dendropsophus bifurcus</i>	R0056
Hylidae	<i>Dendropsophus triangulum</i>	R0068
Hylidae	<i>Hypsiboas alfaroi</i>	R0070
Hylidae	<i>Hypsiboas calcaratus</i>	R0072
Hylidae	<i>Hypsiboas cinerascens</i>	R0073
Hylidae	<i>Osteocephalus fuscifacies</i>	R0084
Leptodactylidae	<i>Engystomops petersi</i>	R0105
Leptodactylidae	<i>Leptodactylus discodactylus</i>	R0106

### f. Elementos de anotación

Se definen tres elementos para anotar en una grabación: intervalo de anotación, canto de rana y nota. Primero, el intervalo de anotación se escoge siguiendo los criterios de selección y debe contener la mayor cantidad de cantos posible. Marcar inicio y final de cada intervalo en el archivo. Segundo, un canto de rana está formado por una o más notas. La figura 1 muestra un canto compuesto de tres notas. Marcar como segmentos la extensión del canto y cada nota individualmente.

### g. Nomenclatura

Para anotar los segmentos seleccionados en cada archivo se usan marcadores de *cue list*. Usar en la etiqueta las palabras en mayúsculas INICIO para el punto inicial y FINAL para el punto final del intervalo de anotación tal como se muestra en la figura 7. Solo anotar los cantos entre un INICIO y FINAL. Para marcar todo el canto de rana se usa la etiqueta RXXXXA en canto de apareamiento y RXXXXT en canto territorial. Para marcar cada sílaba componente de un canto se etiqueta con RXXXXAYY en



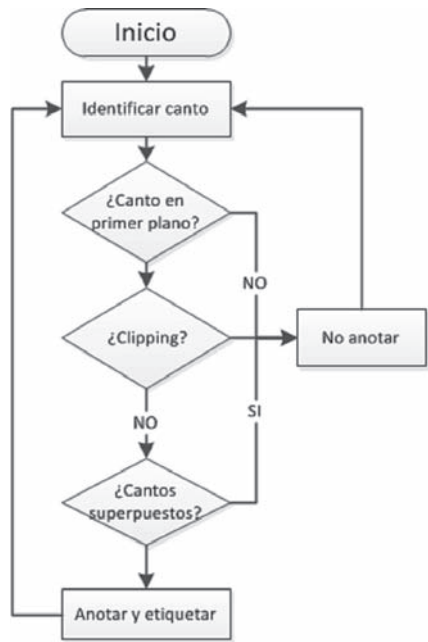


canto de apareamiento y RXXXXTTY en canto territorial. Donde XXXX representa la Identificación numérica única de cada especie de rana en la TABLA 1, y YY es un número de dos dígitos que indica el orden de ocurrencia de cada sílaba en un mismo canto. Cada canto YY inicia en 01.

### ***h. Procedimiento de anotación***

En cada especie de rana se dispone de una lista de archivos \*.WAV que contienen grabaciones de cantos de rana. Identificar en cada archivo intervalos de anotación que contengan la mayor cantidad de cantos de rana posible. Puede existir uno o más intervalos de anotación por archivo. Estos intervalos deben limitarse a condiciones libre de ruido descritos en la sección II. En cada intervalo de anotación se seleccionan los cantos a anotar siguiendo el flujo de la figura 6 con ayuda de la forma de onda y el espectrograma. Primero identificar un canto que pertenezca a la especie buscada. Si el canto tiene suficiente volumen para estar en primer plano, comprobar que no se encuentre saturado. Si el canto, además, no está superpuesto con cantos de la misma u

otra especie, anotar y etiquetar. Si alguna de las condiciones anteriores no se cumple, no anotar, y continuar al siguiente canto hasta llegar al final del intervalo de anotación.



**Figura 6. Flujo de anotación de un canto de rana**

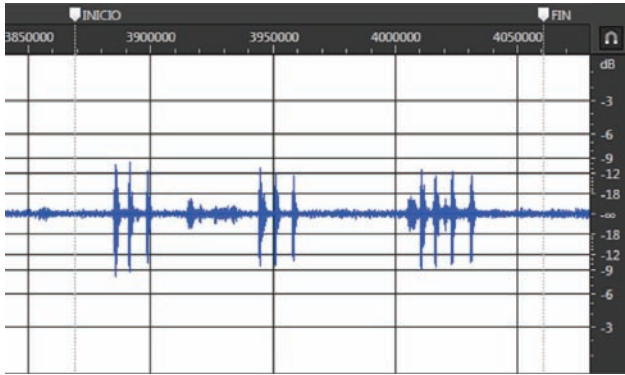


Figura 7. Inicio y final de un intervalo de anotación

Con Adobe Audition 6.0 el procedimiento de anotación y etiquetado se realiza de la siguiente forma:

1. Seleccionar con el "mouse" el elemento de anotación a marcar.
2. Presionar 'M' o en el menú principal dar *click* en "Edición-> Marcador-> Añadir Marcador". Escoger marcador de señal (*cue*).
3. Editar el marcador con la nomenclatura correspondiente.
4. Continuar al siguiente elemento de anotación y repetir los pasos 1-3 de ser el caso.
5. Guardar al culminar la anotación.

#### ***i. Marcar el Intervalo de Anotación***

Los intervalos de anotación deben contener a todos los cantos anotados en una grabación. Marcar más de un intervalo de anotación por archivo de ser necesario. Posicionar siempre el INICIO en una zona con ruido de fondo estable y localizado al menos unos 200 milisegundos antes del primer canto en el intervalo ya que este marcador se usa también para estimar la potencia del ruido de fondo. Usar el mismo criterio para posicionar el FINAL. La figura 7 presenta un ejemplo de selección de un intervalo de anotación con las etiquetas INICIO y FIN colocadas adecuadamente.

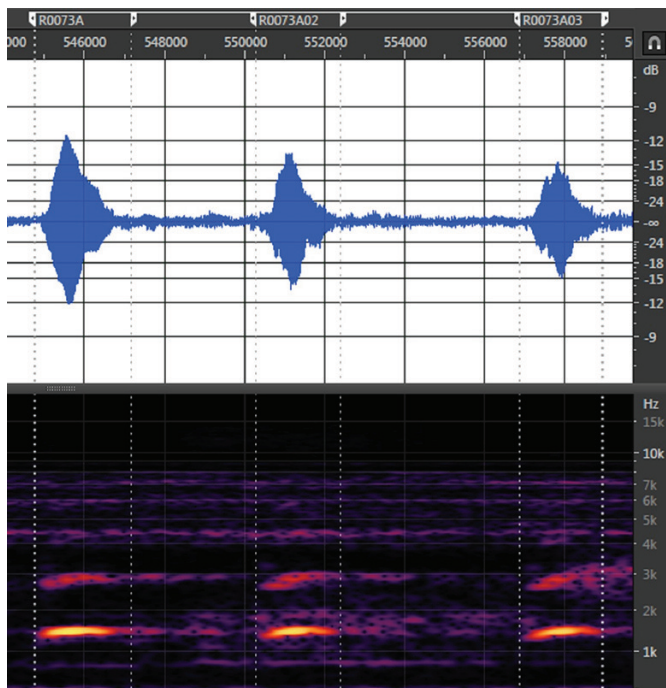


Figura 8. Marcado de un canto con tres notas de *Hypsiboas cinerascens*

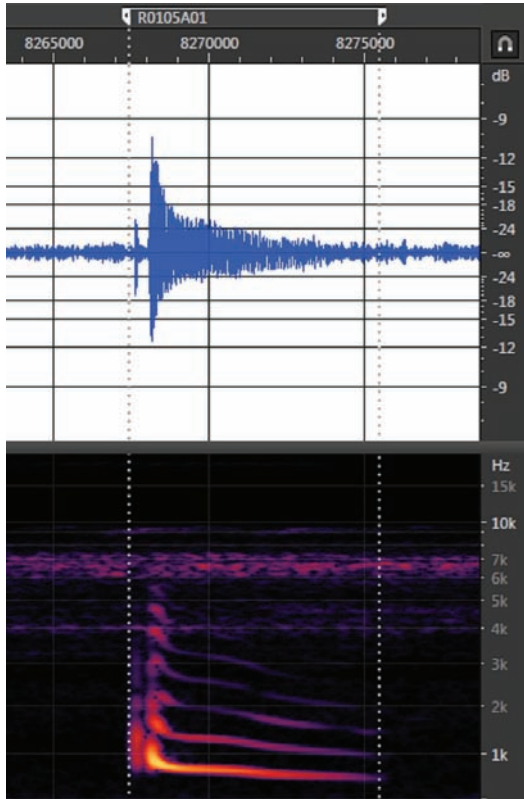


Figura 9. Marcado de un canto de *Engystomops Petersi*

#### ***j. Marcar el canto de rana***

Seleccionar los cantos de acuerdo con el flujo de la figura 6. Usar el espectrograma como guía para reconocer superposición de cantos. Al anotar, marcar los límites del segmento en la forma de onda. Las figuras 8 y 9 muestran un ejemplo de anotado de cantos de cor-

tejo compuestos de una nota y varias notas. Las etiquetas se asignan R0073A y R0105A respectivamente.

#### ***k. Marcar las notas del canto de rana***

Las notas del canto de rana se marcarán de manera secuencial en el canto de acuerdo con su su orden de



ocurrencia. Por ejemplo, el canto de *Hypsiboas Cinerascens* en la figura 8 tiene 3 notas. Se asigna el nombre para el marcador de la siguiente forma:

1. Nota 1: R0073A01
2. Nota 2: R0073A02
3. Nota 3: R0073A03

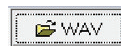
Las notas de un canto se encuentran separadas por silencios. Como regla general se registran las notas de un canto si los silencios intermedios son de duración igual o mayor a la notas. En caso que no exista una separación obvia, marcar todo el canto como una sola nota como se muestra en la figura 9.

### ***1. Generar el archivo de intercambio con CueListTool***

Debido a que los marcadores creados con Adobe Audition 6.0 se guardan directamente en el archivo .WAV, se usa CueList Tool 7.0 para generar una lista de marcadores en texto simple para intercambio. El archivo generado debe tener el mismo nombre del archivo .WAV, pero con extensión .RTF.

El procedimiento para generar los archivos es el siguiente:

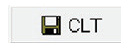
1. Abrir el archivo .WAV con CueListTool:  
File > Open Wav File o click en



2. Aplicar el botón "Read Cue list file from Wav File" dar click en



3. Guardar por defecto con el mismo nombre del archivo y en la misma carpeta. Aplicar el botón "Save Cue list file to disk"



4. Guardar también el archivo de texto .RTF aplicando el botón "Save text window to disk" dando click en



## **RESULTADOS Y DISCUSIÓN**

### ***Lista de especies de ranas seleccionadas, anotadas y etiquetadas***

El número de cantos seleccionados, anotados y etiquetados bajo el pro-

colo desarrollado en el presente trabajo fue de 67; en la TABLA 2 se puede observar el número de cantos por especie.

**TABLA 2. Número de cantos por especie**

FAMILIA	GÉNERO	CÓDIGO	NO. CANTOS
Bufonidae	Rhinella margaritifera	R0009	6
Craugastoridae	Pristimantis conspicillatus	R0023	8
Hylidae	Dendropsophus bifurcus	R0056	8
Hylidae	Dendropsophus triangulum	R0068	4
Hylidae	Hypsiboas alfaroi	R0070	5
Hylidae	Hypsiboas calcaratus	R0072	8
Hylidae	Hypsiboas cinerascens	R0073	9
Hylidae	Osteocephalus fuscifacies	R0084	5
Leptodactylidae	Engystomops petersi	R0105	11
Leptodactylidae	Leptodactylus discodactylus	R0106	3

### **Cross-check**

Para realizar el *cross-check* de los cantos anotados el procedimiento a seguir es:

Aleatoriamente se redistribuirán los cantos anotados entre personas diferentes a las que realizaron la anotación del canto. Estas personas deben revisar todos los cantos anotados y de ser el caso, corregir si alguno presenta alguna falla en la anotación.

Por cada especie, quién esté realizando el *cross-check* anota nuevamente un canto –aleatoriamente–, escogiendo el mismo punto de inicio y fin.

Luego de realizar un estudio en la literatura, no se encontró un protocolo para la selección, anotación y etiquetado de material bioacústico, específicamente

para el canto de las ranas y que permita intercambio. El presente trabajo sugiere un protocolo y se propone usar un archivo de texto simple generado con el software *CueListTool* para intercambio.

El uso de este protocolo con un archivo de intercambio permite la automatización de los procesos de entrenamiento y testado a través de software especializado sin necesidad de analizar el archivo digital de audio que podría ser demasiado grande. Además, permite el fácil intercambio de información de contenido de bioacústica en archivos que son generalmente de larga duración entre los interesados sin enviar el audio digital cada vez.

Un archivo de *CueList* ofrece un detalle del contenido de la grabación como referencia para una rápida extrac-



ción de información sin necesidad de escuchar toda la grabación, lo que podría llevar mucho tiempo.

## CONCLUSIONES

El presente documento propone un estándar de selección de material acústico, explica como marcar los cantos de rana con sus respectivas notas usando etiquetas predefinidas. Es importante tener una base de datos consistente con cantos de ranas de diferentes especies. La presencia de ranas es un indicador muy importante del buen estado del ambiente en regiones como la selva amazónica. Esta base de datos será base para investigadores en áreas de bioacústica, bioinformática, procesamiento digital de señales, entre otras.

Se sugiere instrucciones detalladas de selección, anotación y etiquetado de los cantos para generar un corpus de evaluación y archivo de material bioacústico que servirá como referencia para investigaciones futuras.

Una aplicación específica de la base de datos de cantos de rana, con sus respectivas anotaciones, es para el reconocimiento automático de especies de rana a base de sus cantos. Aquí es muy importante un modelo estocástico por especie, donde se requieren cantos de rana previamente identificados.



## BIBLIOGRAFÍA

- Bioacoustic approaches in biodiversity inventories. (2010). En *Manual on Field Recording Techniques and Protocols for All Taxa Biodiversity Inventories* (Vol. 8, pp. 68-99). Abc Taxa.
- Brandes, T. S. (2008). Feature vector selection and use with hidden Markov models to identify frequency-modulated bioacoustic signals amidst noise. *Audio, Speech, and Language Processing, IEEE Transactions on*, 16(6), 1173–1180.
- CueListTool 1.7. (s. f.). [Software Documentation].
- Duellman, W. E., and L. T. (1986). *Biology of Amphibians*. (McGraw-Hill Book, Ed.). New York.
- Jaafar, H., & Ramli, D. A. (2013). Automatic syllables segmentation for frog identification system. En *Signal Processing and its Applications (CSPA), 2013 IEEE 9th International Colloquium on* (pp. 224-228). <http://doi.org/10.1109/CSPA.2013.6530046>
- Kasturi, R., Goldgof, D., Soundararajan, P., Manohar, V., Garofolo, J., Bowers, R., ... Zhang, J. (2009). Framework for Performance Evaluation of Face, Text, and Vehicle Detection and Tracking in Video: Data, Metrics, and Protocol. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(2), 319-336. <http://doi.org/10.1109/TPAMI.2008.57>
- Multimedia Programming Interface and Data Specifications 1.0. (1991, agosto).
- Reynolds, D. A., & Rose, R. C. (1995). Robust text-independent speaker identification using Gaussian mixture speaker models. *IEEE Transactions on Speech and Audio Processing*, 3(1), 72-83. <http://doi.org/10.1109/89.365379>
- Ron, S. R., Guayasamín, J. M., Yáñez-Muñoz, M. H., Merino-Viteri, A. Ortiz, D. A. y Nicolalde, D. A. (2016). AmphibiaWebEcuador. Recuperado 20 de febrero de 2016, a partir de <http://zoologia.puce.edu.ec/vertebrados/anfibios/EspeciesEstadisticas.aspx>
- Taylor, A., Watson, G., Grigg, G., & McCallum, H. (1996). Monitoring Frog Communities: An Application of Machine Learning. En W. J. Clancey & D. S. Weld (Eds.), *AAAI/IAAI, Vol. 2* (pp. 1564-1569). AAAI Press / The MIT Press.