



DEPARTAMENTO DE CIENCIAS FORENSES
ORGANISMO DE INVESTIGACION JUDICIAL (OIJ)
PODER JUDICIAL, COSTA RICA

Análisis preliminar de muestras de voz

**PROCEDIMIENTO DE
OPERACIÓN NORMADO
ESPECIFICO**

P-DCF-ECT-ISF-21

VERSION: 04

Rige desde: 19/02/2024

PAGINA: 1 de 8

Elaborado o modificado por:	Revisado por Líder Técnico:
Lic. José Alfonso Rodríguez Arias Perito de Acústica Forense y Biometría de Voz	Lic. José Alfonso Rodríguez Arias Líder Técnico de Sección/Unidad de Acústica Forense
Visto Bueno Encargado de Calidad:	Aprobado por:
Lic. José Alfonso Rodríguez Arias Encargado de Calidad de la Sección de Imagen y Sonido Forense	Lic. Marco Herrera Charraun Jefatura, Sección de Imagen y Sonido Forense

CONTROL DE CAMBIOS A LA DOCUMENTACIÓN

Versión	Fecha de Aprobación	Fecha de Revisión	Descripción del Cambio	SCD	Solicitado por
01	23/11/2017	06/05/2020	Versión Inicial del Procedimiento	003-SCD-ISF-2017	RVF
02	06/05/2020	28/09/2021	Cambios por actualización de software Batvox y cambios en redacción	028-SCD-ISF-2020	MHC
03	28/09/2021	19/02/2024	Incluir SNR como criterio de aceptación y/o rechazo y corrección en numeración.	002-SCD-ISF-2021	MHC
04	19/02/2024	-	Cambios en tiempo mínimo de habla neta para audios de test e inclusión de leyenda de validación sobre cambio en tiempo de habla neta y cambios en redacción y nombre	001-SCD-ISF-2024	MHC

**ESTE PROCEDIMIENTO ES UN DOCUMENTO CONFIDENCIAL
PARA USO INTERNO DEL DEPARTAMENTO DE CIENCIAS FORENSES
SE PROHÍBE CUALQUIER REPRODUCCIÓN QUE NO SEA PARA ESTE FIN**

La versión oficial digital es la que se mantiene en la ubicación que la Unidad de Gestión de Calidad defina. La versión oficial impresa es la que se encuentra en la Unidad de Gestión de Calidad. Cualquier otro documento impreso o digital será considerado como copia no controlada.

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 04	PAGINA: 2 de 8
Análisis preliminar de muestras de voz	P-DCF-ECT-ISF-21	

1 Objetivo:

Unificar los pasos a seguir para realizar análisis preliminar de muestras de voz, con la finalidad de definir si son aptas o no para la pericia de comparación de voces.

2 Alcance:

El alcance de este procedimiento se concentra en la determinación de las razones técnicas del por qué una muestra de voz es apta o no para ser sometida a un análisis de comparación de voces. Aplica para todo el personal que realiza una sesión de análisis preliminar de muestras en la sección de Imagen y Sonido Forense del Departamento de Ciencias Forenses.

Ver Informe de Validación 001-ISF-VAL-2024, acerca del requisito de habla neta para comparación de voces.

3 Referencias:

- Andrzej Drygajlo, Michael Jessen, Stefan Gfroerer, Isolde Wagner, Jos Vermeulen and Tuija Niemi. Methodological Guidelines for Best Practice in Forensic Semiautomatic and Automatic Speaker Recognition, including Guidance on the Conduct of Proficiency Testing and Collaborative Exercises. European Network of Forensic Science Institutes (ENFSI). 2015.
- Issues in Data Processing and Relevant Population Selection. OSAC Technical Guidance Document 0006. OSAC Speaker Recognition Subcommittee. The Organization of Scientific Area Committees (OSAC) for Forensic Science. September 23, 2022.

4 Equipos y Materiales:

4.1 Equipos

- Audífonos de orejera completa con rango de frecuencia entre 8 y 25 kHz o similar.
- Estación de trabajo para análisis de audios.
- Programa(s) informático(s) tales como Adobe Audition, Acu-Expert, Media Info o similares, que permitan observar:
 - Espectrograma de banda ancha y estrecha.
 - Formantes de la voz.
 - Frecuencia espectral.

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 04	PAGINA: 3 de 8
Análisis preliminar de muestras de voz	P-DCF-ECT-ISF-21	

- Tono fundamental.

- Oscilograma .

- Propiedades de los archivos de audio: frecuencia de muestreo, profundidad de bits, kilobits por segundo (kbps).

- Software Batvox y/o Nuance Forensics.
- Software ffmpeg, similar o superior.
- Software FTK Imager, HashMyFiles o similar.

4.2 Materiales

- Formulario de Trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica".
- Formulario de Trabajo "Espacio Adicional (Anexo de todos los formularios periciales)".

5 Reactivos y Materiales de Referencia:

No aplica.

6 Condiciones Ambientales:

No aplica.

7 Procedimiento:

7.1 Realice un duplicado del material aportado en la estación de trabajo de la Unidad Acústica Forense en alguna unidad de disco con la creación de una carpeta con el número de orden de trabajo. Dicho duplicado se debe hacer mediante un software que asegure la réplica completa del contenido de cada dispositivo analizado, similar o superior al FTK Imager.

7.2 Realice para cada uno de los dispositivos de almacenamiento aportados el cálculo de los códigos MD5 y/o SHA1 del material mediante el software FTK Imager, HashMyFiles o similar. Cada archivo resultante será convertido a PDF mediante el software de elección del analista.

Incorpore los PDF con los cálculos MD5 y/o SHA1 del contenido de cada uno de los dispositivos de almacenamiento al legajo durante el proceso de registro de la apertura de indicios.

Si el material aportado contiene un archivo describiendo algún cálculo Hash ejecutado anteriormente coteje con los resultados obtenidos durante la apertura. Reporte en el informe si se encontraron diferencias.

7.3 Revisión de propiedades de la grabación.

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 04	PAGINA: 4 de 8
Análisis preliminar de muestras de voz	P-DCF-ECT-ISF-21	

7.3.1 Revise en software MediaInfo o similar las propiedades digitales de la grabación: peso del archivo, formato, códec o contenedor de grabación, frecuencia de muestreo registrada, profundidad de bits registrada, tasa de bits por segundo (kbps), indicar si es constante o variable, compresión y tiempo total de grabación. Registre dichas propiedades en el formulario de trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica".

Nota 1: Idealmente la frecuencia de muestreo debe ser 8000 Hz o superior, la profundidad de bits debe ser 16 bits o superior y la tasa de bits deber ser 128 kbps o superior, en caso de que alguna, varias o todas las anteriores sean inferiores a los valores mencionados verifique la información de los formantes de la voz como se indica en el punto 7.4 antes de descartar la muestra, ya que estos no son criterios de aceptación o rechazo.

7.3.2 En ocasiones el formato en el que viene la muestra no permite importarlo al software Adobe Audition o similar, por lo que se debe convertir a un formato sin pérdida wav, proceda con algún software de conversión, preferiblemente ffmpeg.

7.3.3 Revise la presión sonora de la señal de audio en software Adobe Audition, si la misma supera los 0dB (cero decibeles) normalícelo a -1dB. Deje registro en el formulario de trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica".

7.3.4 Revise en el software de comparación (Batvox o Nuance Forensics) el tiempo de habla neta del locutor de interés y SNR (relación señal-ruido).

7.3.5 Verifique que el tiempo de habla neta mínima sea de 30 segundos para las muestras que serán utilizadas como modelo (audios de entrenamiento) y 15 segundos para los audios de test. En caso contrario, los audios aportados no son aptos para comparación de voces.

7.3.6 Revise la SNR, si la misma es inferior a 10 dB la muestra de audio no es apta para comparación.

7.4 Realice análisis acústico con un software de análisis de audio similar o superior al AcuExpert de la siguiente manera:

7.4.1 Visualice en un espectrograma LPC la totalidad del audio, con el objetivo de ver si se mantiene el registro de los formantes a lo largo de toda la grabación. Para mejorar el aspecto de la representación gráfica ajuste según sea necesario el tamaño, la resolución y la intensidad del espectrograma.

7.4.2 Si durante esa visualización se observan los formantes de manera parcial (ubicado en ciertas áreas de la grabación) o total, proceda a realizar el análisis en al menos cuatro vocales por cada 60 segundos de habla neta. Si durante la visualización no se observa al menos el registro de los tres primeros formantes de la voz, verifique los formantes realizando análisis en al menos cuatro vocales por cada 60 segundos de habla neta, si de igual forma no se observa el registro de los tres primeros formantes la muestra de audio no es apta para comparación.

7.4.3 Para estos dos puntos realice capturas de imagen de la representación del espectrograma e incluya las mismas en el Formulario de Espacio Adicional.

7.4.4 Calcule el tono fundamental (pitch) y posteriormente defina el mismo en el cepstrograma.

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 04	PAGINA: 5 de 8
Análisis preliminar de muestras de voz	P-DCF-ECT-ISF-21	

7.4.5 Calcule los formantes y busque las vocales necesarias según lo indicado en el punto 7.4.2.

7.5 Registre las propiedades indicadas en los puntos 7.3 y 7.4 en el formulario de trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica". Asimismo utilice el "Formulario de Espacio Adicional" para adjuntar imágenes y otros documentos que sean necesarios.

7.6 Agregue al legajo del caso el formulario de trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica" y el formulario de "Espacio Adicional".

Nota 1: Además de los casos citados anteriormente no son admisibles para análisis de comparación de voces:

- Los archivos que se determine que se han grabado a partir de un dispositivo donde se está reproduciendo el audio de interés, por ejemplo reproducir audios a partir de un teléfono celular y grabarlo con otro dispositivo, grabar una llamada telefónica en altavoz con otro dispositivo, entre otros.
- Audios que presenten incrustado dentro del mismo canal y/o frecuencia, música en un nivel similar al habla.
- Audios que se identifique que la voz está siendo interrumpida o modificada debido a que la persona de interés se está tapando la boca o usando algún filtro para variar la voz.

8 Criterios de Aceptación o Rechazo de Resultados:

No.	Criterio de rechazo	Valor Límite	Corrección Aplicable
1	Audios que no cuenten con registro en al menos los primeros tres formantes.	Mínimo registro en el primero, segundo y tercer formante	Se rechaza la muestra y se indica que no es apta para comparación
2	Audios modelos o de entrenamiento con una duración inferior a 30 segundos de habla neta.	30 segundos de habla neta para audios modelo o de entrenamiento	Se indica que la muestra no es apta
3	Audios de test con una duración inferior a 20 segundos de habla neta.	20 segundos de habla neta para audios de test	Se indica que la muestra no es apta
4	Audios con una SNR inferior a 10 dB.	10 dB	Se indica que la muestra no es apta si el valor es inferior a 10 dB
5	Archivos que se identifique que se han grabado a partir de un	No aplica	Se indica que la muestra no es apta

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 04	PAGINA: 6 de 8
Análisis preliminar de muestras de voz	P-DCF-ECT-ISF-21	

	dispositivo donde se está reproduciendo el audio de interés		
6	Audios que presenten incrustado dentro del mismo canal y/o frecuencia música en un nivel similar al habla.	No aplica	Se indica que la muestra no es apta
7	Audios que se identifique que la voz está siendo interrumpida o modificada debido a que la persona de interés se está tapando la boca o usando algún filtro para variar la voz.	No aplica	Se indica que la muestra no es apta

9 Cálculos y evaluación de la incertidumbre:

No aplica.

10 Reporte de Análisis y Resultados:

10.1 Indique al solicitante mediante informe o dictamen pericial, si la muestra o muestras aportadas son aptas o no para un análisis de comparación de voces. Utilice como referencia los siguientes textos:

"El archivo "nombre del archivo" cumple con los parámetros digitales y acústicos necesarios para ser utilizado como audio modelo y/o de test en un análisis de comparación de voces".

"El archivo "nombre del archivo" no es apto para comparación debido a que no cumple con los parámetros digitales y/o acústicos necesarios, específicamente "motivo por el cual no es apto".

10.2 Para los casos en los que en la misma solicitud piden análisis preliminar y análisis de comparación de voces (en caso de ser aptas las muestras), se debe generar una nueva orden de trabajo, a partir de la cual se realizará un consumo de la orden de trabajo de análisis preliminar, por lo que el análisis preliminar de las muestras se describirá en el mismo dictamen que se emitirá para el análisis de comparación de voces.

11 Medidas de Seguridad y Salud Ocupacional:

No aplica.

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 04	PAGINA: 7 de 8
Análisis preliminar de muestras de voz	P-DCF-ECT-ISF-21	

12 Simbología:

dB: decibeles.

Kbps: Kilobit por segundo.

KHz: KiloHertz.

Hz: Hertz.

LPC (Linear predictive coding): Codificación predictiva lineal.

PDF: Portable Document Format.

SNR: Relación señal-ruido (por sus siglas en inglés signal to noise ratio).

XLR Siglas de external live return, lo cual refiere a un tipo de conector balanceado más utilizado para conexiones de audio profesional.

13 Terminología:

- Audio de entrenamiento: Fichero de audio que contiene la voz del sospechoso y que se usa para crear un modelo biométrico de su voz.

- Audio de test: Audio que contiene voz cuya identidad es desconocida y que se utiliza para compararse con modelos de sospechoso con el objetivo de encontrar alguna coincidencia que permita identificar.

- Batvox: Programa biométrico forense de comparación de voces.

- Código hash: Algoritmo que consigue crear a partir de información registrada, un código alfanumérico de longitud normalmente fija que representa el contenido y asegura que no se ha modificado una transmisión o duplicación.

- Compresión: Proceso por el cual se reduce la tasa de bits de una señal digital de audio buscando como fin la reducción de su peso en disco, a medida que se comprime el audio se reducen estos valores, con lo cual se pierde fidelidad con respecto al sonido físico.

- Decibeles (dB): unidad de medida de la presión sonora.

- Espectrograma: representación gráfica de la energía del contenido frecuencial de la señal de una de una grabación de sonido y sus variaciones en el transcurso del tiempo que ésta dura.

- Formante de la voz: pico de intensidad en el espectro de un sonido; se trata de concentración de energía (amplitud de onda) que se da en una determinada frecuencia.

- Formato de grabación: propiedades digitales de la forma en la que se almacenan los datos de una grabación de audio y cómo puede utilizarse en un sistema de administración de archivos.

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 04	PAGINA: 8 de 8
Análisis preliminar de muestras de voz	P-DCF-ECT-ISF-21	

- Formato wav: formato de audio digital el cual normalmente se utiliza sin compresión de datos.

-Frecuencia espectral: muestra todos los elementos sónicos individuales de un sonido.

- Frecuencia de muestreo: Es la cantidad de muestras de audio que se toman en un segundo. Esta se mide en Hertz (Hz). La cantidad de muestras tomadas definen la calidad de la muestra y su correspondencia respecto al original.

- Habla neta: tiempo real de duración de un relato leído, espontáneo o dialogado sin silencios ni interrupciones de un mismo locutor.

- Hertz (Hz): unidad de medida de frecuencia del Sistema Internacional de Unidades.

- Oscilograma: representación de la presión sonora en función del tiempo.

- Remuestreo: Reasignación en post-producción, de la cantidad de muestras de audio que se toman por segundo.

- Relación señal-ruido: se define como la proporción existente entre la potencia de la señal que se transmite y la potencia del ruido que la corrompe. Este margen es medido en decibelios.

- Tono fundamental: es el número de veces que vibran los pliegues vocales por segundo.

- Vúmetro: indicador del nivel de volumen que entra a un equipo de audio expresado en decibeles (dB).

14 Anexos

No aplica.