



DEPARTAMENTO DE CIENCIAS FORENSES
ORGANISMO DE INVESTIGACIÓN JUDICIAL (OIJ)
PODER JUDICIAL, COSTA RICA
**Comparación de voces mediante software
Nuance Forensics**

**PROCEDIMIENTO DE
OPERACIÓN NORMADO
ESPECIFICO**

P-DCF-ECT-ISF-36

VERSION: 02

Rige desde: 19/02/2024

PAGINA: 1 de 8

Elaborado o modificado por:	Revisado por Líder Técnico:
Lic. José Alfonso Rodríguez Arias Perito en Acústica 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 Imagen y Sonido Forense

CONTROL DE CAMBIOS A LA DOCUMENTACIÓN

Versión	Fecha de Aprobación	Fecha de Revisión	Descripción del Cambio	SCD	Solicitado por
01	10/08/2023	19/02/2024	Versión Inicial del Procedimiento	008-SCD-ISF-2023	MHC
02	19/02/2024	-	Agregar indicación de validación	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 02	PAGINA: 2 de 8
Comparación de voces mediante software Nuance Forensics	P-DCF-ECT-ISF-36	

1 Objetivo:

Estandarizar la metodología para realizar análisis de comparación de voces con el uso del software Nuance Forensics.

2 Alcance:

Este procedimiento se refiere a la comparación de dos o más grabaciones de audio en las que intervengan voces, con la finalidad de clasificar una señal de voz de origen desconocido, como perteneciente a una señal entre un conjunto de posibles locutores, por medio de análisis biométrico de la voz, utilizando el software forense Nuance Forensics.

El uso de este procedimiento es de acatamiento obligatorio para todos los miembros de la Unidad de Acústica Forense, que realicen comparaciones de locutores que se requieran someter a comparaciones.

[Esta metodología se encuentra validada en el Informe de validación 002-ISF-VAL-2023 e Informe de validación 001-ISF-VAL-2024.](#)

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.
- Barros, José Luis y Moser, Michael. (2009). Ingeniería Acústica. España: Springer.
- European Network of Forensic Science Institutes (ENFSI). (2015). Guideline for Evaluative Reporting in Forensic Science.
- Evett IW, Jackson G, Lambert JA and McCrossan. "The impact of the principles of evidence interpretation on the structure and content of statements". Science & Justice, 40 (4) 233-239, 2000.
- Evett IW. Towards a uniform framework for reporting opinions in forensic science casework. Science & Justice; 38(3) 198-202. 1998.

4 Equipos y Materiales:

4.1 Equipos

- Audífonos profesionales de orejera completa.
- Computadora de trabajo.
- Software Adobe Audition.
- Software Nuance Forensics.

4.2 Materiales

- Formulario "Mantenimiento Preventivo y/o Verificación de Software".

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 02	PAGINA: 3 de 8
Comparación de voces mediante software Nuance Forensics	P-DCF-ECT-ISF-36	

- 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:

N/A.

6 Condiciones Ambientales:

N/A.

7 Procedimiento:

Nota 1: La verificación del funcionamiento del software Nuance Forensics la lleva a cabo una empresa externa al Poder Judicial, deje registro de esto en el formulario de "Mantenimiento Preventivo y/o Verificación de Software" y mediante la entrega del informe por parte realizado por el técnico que realiza el servicio.

De no contar con contrato para el mantenimiento y verificación, realice verificación del funcionamiento del software Nuance Forensics al menos una vez al año con muestras de voz indubitadas, deje registro de esto en el formulario de "Mantenimiento Preventivo y/o Verificación de Software", indicando cuáles muestras se utilizaron y el resultado del mismo, anote además si el resultado es el esperado, en caso de no serlo verifique nuevamente. Si el resultado sigue siendo distinto al esperado pese a realizar varias veces la verificación coloque un rótulo de "Equipo en mal estado" hasta que se logre determinar por medio del soporte técnico especializado el motivo del resultado distinto y corregirlo posteriormente.

Nota 2: Los archivos de audio que se sometan a este procedimiento de comparación, deben cumplir con los parámetros establecidos en el Procedimiento Análisis Preliminar de Muestras de Voz para Comparación de voces y estar consignado por medio del formulario de trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica".

7.1 Prepare la muestra en el programa Adobe Audition o similar para ser ingresada al software de comparación, con las siguientes características digitales:

-Formato WAV con codificación lineal PCM sin compresión.

-Frecuencia de muestreo 8 KHz.

-Resolución 16 bits.

-Canales: mono.

-Se deben eliminar vocalizaciones que no son del habla, como por ejemplo: risas, tos, carraspeos y clics, o estilos de habla inusuales, susurros, gritos, entre otros. De igual forma se deben eliminar pausas.

-En caso de que un archivo de audio presente más de una voz, se debe proceder a separar las mismas, de manera que se genere un archivo de audio para cada uno de los hablantes.

7.2 Deje registro de todo lo indicado en este punto en el formulario de trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica".

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 02	PAGINA: 4 de 8
Comparación de voces mediante software Nuance Forensics	P-DCF-ECT-ISF-36	

7.3 Accese al software forense de comparación de voces Nuance Forensics y cree un nuevo perfil del hablante que utilizará como modelo (speaker) y cargue el archivo correspondiente.

7.4 A partir del modelo o speaker creado genere un caso nuevo (identifique el caso en el sistema con el número único y/o número de orden de trabajo), agregue el audio contra el que desea hacer la comparación, la población de referencia a utilizar y si tiene algún comentario que agregar realícelo en esta misma ventana.

Nota 3: Dicha población de referencia no se considera una colección de referencia debido a que estos audios son utilizados para realizar los procesos de normalización recomendados para llevar a cabo tareas de reconocimiento del software, no para ser utilizados como audios de test o modelo, por lo tanto no se requiere una identificación única para los audios de esa población ya que no se van a realizar identificaciones con la misma.

7.5 Abra el reporte HTML que genera el software.

7.6 Observe y analice el resultado que genera el programa, utilizando como referencia los cuadros 1 y 2 indicados en el punto 10.2.

7.7 Convierta el reporte de resultados en formato PDF y agréguelo al legajo digital del SADCF.

7.8 Deje registro de cada uno de los resultados en el formulario de trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica", incorpore el mismo al legajo digital del caso en el SADCF.

7.9 Indique en el formulario de trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica" el valor de EER observado.

8 Criterios de Aceptación o Rechazo de Resultados:

8.1 Los criterios de aceptación o rechazo se definen en el procedimiento de "Análisis Preliminar de Muestras de Voz y Comparación Biométrica".

9 Cálculos y evaluación de la incertidumbre:

N/A.

10 Reporte de Análisis y Resultados:

Nota 4: El LR es una relación de dos probabilidades, que la voz del sospechoso sea la del dubitado o que la voz del sospechoso pertenezca a un tercero cuya grabación se hubiera realizado en condiciones acústicamente similares, es decir que no corresponda a la del dubitado.

Si el resultado del LR es igual a 1, la voz del sospechoso pertenece al dubitado, con una baja probabilidad. Cuanto más aumente el número del LR, mayor será la probabilidad de que la voz del sospechoso pertenece al dubitado. Cuando el resultado es superior a 1, la línea de evidencia, queda dentro de la curva de intravariabilidad. Indica que es más probable que el audio modelo sea la misma persona respecto al audio cuestionado, y muestra cómo puntúa el sospechoso contra sí mismo.

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 02	PAGINA: 5 de 8
Comparación de voces mediante software Nuance Forensics	P-DCF-ECT-ISF-36	

Si el LR es menor a 1, y cuanto más bajo sea, aumenta la probabilidad de que los audios en cuestión no corresponden a la misma persona. Cuando el resultado del LR es inferior a 1, la línea de evidencia, queda dentro de la curva de intervariabilidad. Indica que es más probable que el audio modelo no sea la misma persona respecto al audio cuestionado.

10.1 En el dictamen pericial indique como resultado el valor numérico que da el software, el mismo puede ser mostrado mediante la imagen de los resultados brindada por el software.

10.2 En el Dictamen Pericial, indique como interpretación si existe o no correspondencia entre las voces comparadas, explicando lo que significa en términos de probabilidad el valor de LR otorgado por el software y brinde un apoyo a una hipótesis según corresponda, siguiendo como referencia la escala verbal para expresar LR adaptada de las directrices de ENFSI para informes forenses, según lo indicado en los cuadros 1 y 2.

Por ejemplo: "Según el resultado del análisis y el LR obtenido, la comparación entre la voz contenida en la grabación del archivo X y la voz contenida en la grabación del archivo Y es LR 0,00047 (sin redondear 0,000470655546406432), lo que significa que la probabilidad de que ambas grabaciones pertenezcan a la misma persona es 0,00047.

Tomando como referencia el LR obtenido y la escala verbal para expresarlo, el perito que firma este dictamen brinda un apoyo fuerte a la hipótesis de que la voz en el archivo "X" no corresponde con la voz en el archivo "Y".

Cuadro 1

Cuando el valor del LR es superior a 1, se utiliza la escala verbal del siguiente cuadro:

Soporte a hipótesis	Escala verbal	LR
Voz dubitada corresponde con la del sospechoso	Apoyo limitado	$1 < LR \leq 10$
	Apoyo moderado	$10 < LR \leq 100$
	Apoyo moderadamente fuerte	$100 < LR \leq 1000$
	Apoyo fuerte	$1000 < LR \leq 10000$
	Apoyo muy fuerte	$LR > 10000$

Cuadro 2

Si el valor del LR es inferior a 1, se utiliza la escala verbal del siguiente cuadro:

Soporte a hipótesis	Escala verbal	LR
Voz dubitada no corresponde con la del sospechoso	Apoyo limitado	$0.1 \leq LR < 1$
	Apoyo moderado	$0.01 \leq LR < 0.1$
	Apoyo moderadamente fuerte	$0.001 \leq LR < 0.01$
	Apoyo fuerte	$0.0001 \leq LR < 0.001$
	Apoyo muy fuerte	$LR < 0.0001$

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 02	PAGINA: 6 de 8
Comparación de voces mediante software Nuance Forensics	P-DCF-ECT-ISF-36	

10.3 Indique en el Dictamen Pericial la población de referencia utilizada y sus características en el análisis de comparación de voces.

11 Medidas de Seguridad y Salud Ocupacional:

11.1 Nunca coloque los audífonos sobre los oídos antes de que el audio esté sonando, si el sonido inicia de golpe muy fuerte puede lastimarse los oídos.

12 Simbología:

dB: decibeles.

EER: Equal error rate.

ENFSI: European Network of Forensic Science Institutes.

Kbps: Kilobit por segundo.

KHz: KiloHertz.

Hz: Hertz.

N/A: No aplica.

PDF: Portable Document Format.

PNG: Portable Network Graphics.

SADCF: Sistema Automatizado Departamento de Ciencias Forenses.

13 Terminología:

Audios de entrenamiento: Los audios de entrenamiento son aquellos audios con los que se generará un modelo del sospechoso que se utilizará en un cálculo biométrico.

Audios de impostor: Impostores son aquellos audios que no pertenecen al sospechoso y que han sido grabados bajo las mismas condiciones que los audios de test. Al igual que las poblaciones de referencia, los impostores se utilizan para llevar a cabo procesos de normalización para realizar tareas de reconocimiento fiables.

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.

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 02	PAGINA: 7 de 8
Comparación de voces mediante software Nuance Forensics	P-DCF-ECT-ISF-36	

Curva de intervariabilidad: representa el rango de puntuaciones del audio de test cuando se enfrenta a una población de referencia (modelos que no son la misma persona que el audio de test).

Curva de intravariabilidad: estimará como suele puntuar el sospechoso contra sí mismo.

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

Entrenamiento de modelo: Es el proceso biométrico en el que se extraen las características de la voz de una persona y se genera a partir de toda esta información un modelo biométrico.

Escala verbal para expresar LR: equivalente verbal utilizado en análisis científicos cuya unidad de medida es el likelihood ratio (LR).

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.

Falsa aceptación: Situación de error que ocurre cuando un hablante auténtico es rechazado por impostor.

Falso rechazo: Situación de error que ocurre cuando se acepta un impostor como un verdadero hablante.

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.

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

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.

LR: Una relación de verosimilitud, o "likelihood ratios", es un valor probabilístico obtenido por aproximación Bayesiana. Un LR es una relación de probabilidades: por un lado la probabilidad de que la voz desconocida o audio de test pertenezca al sospechoso y por otro estará la probabilidad de que la voz pertenezca al resto del mundo (es decir, que no pertenezca al sospechoso).

Normalización: Proceso que se aplica en los cálculos de reconocimiento para corregir los desalineamientos que se producen en las puntuaciones debido a la naturaleza de los audios.

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.

PCM: Se refiere a un método para representar una forma de onda de audio analógica con una serie de valores de muestra digital cuantizados.

Población de referencia: Una población de referencia es un grupo de modelos utilizados para realizar los procesos de normalización recomendados para realizar tareas de reconocimiento fiables. Cada modelo de sospechoso implicado en una tarea de reconocimiento tendrá asociada una población de referencia y dichos modelos de referencia deberán tener exactamente las mismas características acústicas que las del modelo de sospechoso.

DEPARTAMENTO DE CIENCIAS FORENSES	VERSIÓN 02	PAGINA: 8 de 8
Comparación de voces mediante software Nuance Forensics	P-DCF-ECT-ISF-36	

Tasa de error igual: Punto en el que se encuentran las tasas de falsa aceptación y falso rechazo.

14 Anexos

N/A.

COPIA NO CONTROLADA