

BIOMETRÍA VOCAL: LA HUELLA DIGITAL VOCAL

Por Jordi Renú, Perito Judicial del Audiovisual. Colegiado CPAC nº 0206

¿Qué es la Biometría Vocal? Partiendo de la base de que cada voz tiene una característica personal única, diferente para cada persona, la Biometría nos permite estudiar dónde residen las diferencias entre las voces que una muestra de población es capaz de generar, y desarrollar herramientas que permitan la identificación de personas a través de su voz. Ésta técnica se basa en los elementos físicos de cada individuo y que difieren entre personas. Difieren aspectos como la vibración de las cuerdas vocales, que generan ondas que resuenan en el tracto respiratorio y que serán modificadas por parámetros fisiológicos en su camino hacia el exterior. Intervienen elementos como el cuello, la boca, la cavidad buconasal, la lengua y los dientes; elementos físicos que determinan unas características únicas. Además, las resonancias internas generan armónicos, que enriquecen la voz y la personalizan, haciéndola única. Todos estos aspectos, únicos en cada individuo, son los que nos permiten distinguir una voz de las demás.

La Biometría vocal no se basa sólo en el tono de la voz, en la forma de hablar y el comportamiento del hablante, elementos que un buen imitador domina con destreza y que nos puede divertir, pero que en ningún momento puede servir para suplantar la voz original. La Biometría la descarta.

Sabemos que la voz cambia, no es la misma a primera hora de la mañana que al final de una larga jornada de trabajo. Cambios de temperatura y humedad provocan cambios en la voz; lo mismo que le ocurre a un instrumento de cuerda. Por tanto, la Biometría de voz es más compleja que la huella dactilar, dado que ésta es siempre la misma.

Pero gracias al modelado estadístico somos capaces de separar esta parte variable de la parte fija que nos identifica, que nos distingue. De aquí su fortaleza.

¿Qué es la Huella Vocal?

Es la base de muestras de voz de un individuo que nos permiten diferenciarlo del resto. Para generar una Huella vocal realizaremos una fase de registro o entrenamiento, grabando la voz de la persona en un entorno controlado en la que dirá frases y números que nos permitirán realizar un mapa de su voz. Esta muestra constituye la huella vocal.



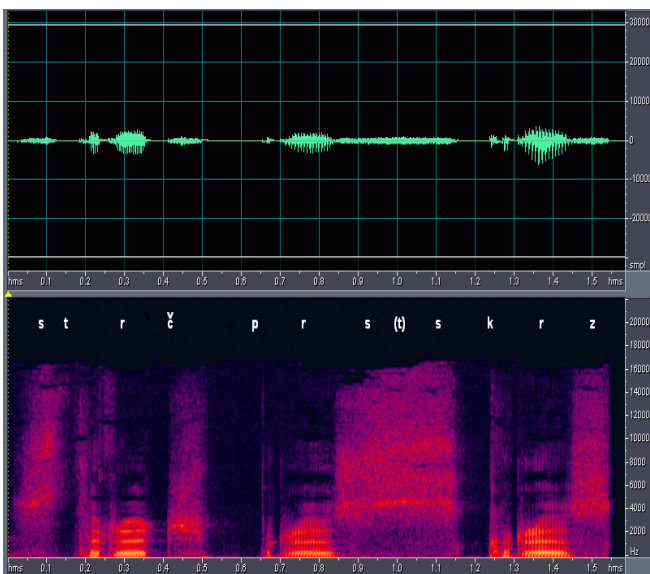
El siguiente paso es verificar que dicha huella es suficientemente segura. Para ello comparamos la muestra obtenida con un banco de muestras de voz de otros cientos o miles de usuarios del mismo sexo, franja de edad, idioma y de la misma zona de influencia. Una vez superada la prueba, dispondremos de una huella vocal verificada.



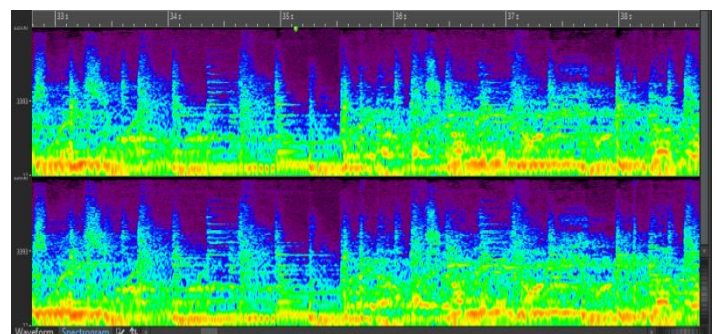
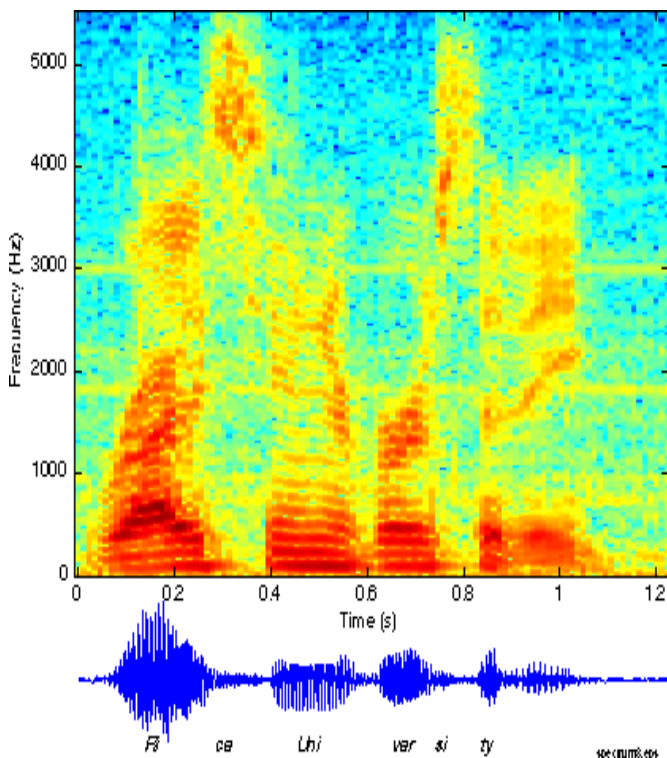
Historia:

Hace algunos años que vivimos rodeados de sistemas de identificación de voz, con sistemas que se han ido desarrollando por diversos fabricantes de aplicaciones para automoción, teléfonos móviles, contestadores de atención al cliente, y sistemas de ayuda para el domicilio. Son sistemas que han crecido a nuestro alrededor y con los que nos hemos acostumbrado a interactuar. Todo y que se trata de sistemas en constante evolución, no siempre son capaces de identificar determinadas palabras entre las diferentes voces, por lo que tienen márgenes de error.

Hace años que vemos películas y series donde las voces se suplantan, donde investigadores son capaces de identificar voces aislando ruidos externos y obteniendo muestras que utilizan para comparar con sus archivos sonoros e identificar sospechosos. Hasta hace poco los procedimientos utilizados en casos civiles eran bastante arcaicos, con intervención humana y técnicas de análisis totalmente subjetivas, con técnicas comparativas.



Entonces tiempo atrás, ¿cómo se podían identificar las voces? Supongamos un caso de duda con dos voces grabadas en soporte audiovisual; una dubitada y una indubitada: los peritos audiovisuales e Ingenieros de sonido utilizaban espectrómetros y monitores de onda para visualizar la forma física la voz. Cada palabra, cada frase, contiene unas frecuencias y armónicos únicos, formando ondas complejas. Además, las personas nunca hablan exactamente igual, así que era difícil que dos grabaciones distintas de la misma persona, las ondas apareciesen exactamente iguales. Comparando subjetivamente las muestras, los resultados eran aproximados.



En la actualidad, la Biometría vocal ha dado un giro espectacular a la labor pericial, eliminando la opinión subjetiva con sistemas automatizados de inteligencia artificial, que son capaces de realizar los cálculos matemáticos necesarios y las comparativas entre miles de voces de forma autónoma y absolutamente objetiva, y en un breve espacio de tiempo. El resultado de los análisis ya no es la opinión de un perito o Ingeniero; son datos empíricos.

Operaciones comerciales realizadas con la voz:

Hace años que realizamos contratos de servicios autorizando los cargos por voz. Las empresas con las que podemos contratar sus servicios con una simple llamada telefónica guardan las grabaciones del contrato como fe de que la operación ha sido contratada y autorizada por quién dice ser; que no siempre ha de ser quien dice que es, puesto que, por poner otro ejemplo, es habitual que los hijos de personas mayores les ayuden contratando los servicios de agua, luz, telefonía y seguros, por citar algunos, haciéndose pasar por ellos. Disponiendo de los datos de la persona, DNI, dirección y número de cuenta bancaria, cualquier persona puede suplantar a otra; con o sin consentimiento del titular.

¿Se puede dar el caso de fraude? Por supuesto. En 2018 se me citó en el Juzgado para un caso de Peritaje audiovisual en el que el Juez pedía que se identificase a una persona que contrató los servicios de una empresa energética por teléfono. Dicha persona no pagó ni uno solo de los recibos emitidos alegando que no fue ella quien realizó el contrato; 14 meses después. Si con los datos básicos de una persona se puede contratar un servicio, ¿quién garantiza que esa persona es quién dice ser? A priori nadie.

¿Qué herramientas tenemos para actuar contra el fraude? *La huella Vocal*.

La Voz en los servicios donde se requiere un alto grado de seguridad:

Hay operaciones de especial importancia donde es imprescindible contar con sistemas de extrema fiabilidad, por ejemplo, con las cuentas bancarias o transacciones comerciales realizadas a distancia, donde la Biometría Vocal es el único sistema que puede garantizar la identidad de una persona a través de su voz. No puede haber margen de error. Los bancos ya están trabajando en sistemas donde sus clientes pueden identificarse y operar por medio de la voz. Hasta no hace mucho, cada usuario disponía de una clave alfa numérica de identificación y se le asignaba una tarjeta con unos códigos numéricos y unas coordenadas, con 100 opciones posibles que se utilizaban para autorizar la operación. Dando los datos correctos se autorizaban las transacciones. Este sistema ha quedado muy obsoleto y los teléfonos móviles han tomado ventaja, recibiendo actualmente por SMS las claves únicas e intransferibles para autorizar las operaciones. Sin embargo, lo más simple es utilizar la voz para todos los propósitos y con total seguridad; siempre que se disponga de huella vocal y equipos de control. Hay momentos en que no se puede estar recibiendo códigos y mirando pantallas de datos, como por ejemplo conduciendo, todo y que siempre es aconsejable duplicar por diversos medios las autorizaciones.

Otras aplicaciones de la Biometría Vocal:

Existen otro tipo de controles donde es mucho más sencillo para los usuarios el acceso, no por ello menos crítico. Por ejemplo, en las empresas donde los trabajadores realizan trabajos manuales como el campo o la minería, muchas veces no es posible utilizar la huella digital para accesos restringidos o el fichaje en el acceso al puesto de trabajo y a la salida. Muchos operarios no disponen de huella digital y en estos casos no es habitual utilizar tarjetas de acceso, que podrían dar pie a fraude de identidad, como se ha dado en casos de trabajadores del campo, donde un trabajador documentado se daba de alta en diversas empresas y enviaba a trabajar a indocumentados en su lugar, que se hacían pasar por el contratado. El problema era finalmente para el empresario, que en el hipotético caso de que alguno de éstos trabajadores se hiciese daño en la jornada laboral, se encontraba con que tenía trabajando en su empresa a un trabajador sin permiso de trabajo, aunque se tratase de una suplantación de identidad. La Ley laboral es muy precisa en el hecho de la

obligación del empresario a controlar el acceso a su empresa exclusivamente a las personas contratadas. Con los sistemas de acceso y fichaje con huella vocal estos problemas han desaparecido.

Los asistentes de voz también han llegado para quedarse. No es necesario teclear los datos en un ordenador para pedir un viaje, hotel, una canción, el tiempo, etc. Cada vez son más precisos, pero su ámbito es meramente personal, a través del ordenador, tablet, móvil, coche o asistente doméstico.

Por ejemplo, hay Hoteles que están implementando controles de acceso a áreas y habitaciones por medio de la voz debido a su robustez. Se acabaron las pérdidas de llaves por parte de los clientes, que ya pueden acceder con su voz.

Las aplicaciones de biometría vocal alcanzan al ámbito Jurídico-Legal con sello biométrico vocal: pueden utilizarse con plena garantía en la lectura de cláusulas a los clientes por parte de empleados de banca, los abogados pueden realizar contratos verbales, los Notarios pueden utilizarlo en la lectura de cláusulas de contratos o testamentos, en el ámbito hospitalario en el consentimiento a intervenciones o cuidados que requieren la firma del enfermo o de un familiar directo, y que se podría realizar a distancia por parte del familiar o por el propio enfermo aunque no pudiese firmar físicamente en ese momento, podría autorizarlo con su voz. Se trata de intervenciones en este caso que pueden llegar a salvar vidas. O por ejemplo, un enfermo al cual se va a intubar en estado de consciencia: se podría dar la posibilidad de que antes de ese momento desease dejar sus últimas voluntades a un médico, a una enfermera, a un Notario, quien se podría presentar con un dispositivo con sistema de huella vocal y grabar sus últimas voluntades por si no superase el tratamiento; incluso telefónicamente. Por desgracia, en éstos últimos años, hemos vivido momentos muy duros y por falta de éste tipo de herramientas, se han imposibilitado muchas de las acciones aquí descritas.

Cotejo Vocal:

Todos disponemos de un equipo de grabación: un teléfono móvil. Con él podemos realizar una grabación en cualquier parte, básicamente una grabación sonora, dado que es mucho más fácil mantener el dispositivo oculto o disimulado cerca del punto de la conversación sin necesidad de tener que estar enfocando a la otra persona con la cámara, como sería en el caso de una grabación de vídeo. Y evidentemente sería más fácil que el interlocutor se diese cuenta de que le están grabando, por lo que se eliminaría el grado de confidencialidad y confianza de la conversación, en la que se pueden confesar cosas o hechos que de otra manera no se realizarían.

Este hecho no contraviene la Ley Orgánica 3/2018 de Protección de datos personales, y existe jurisprudencia que le da soporte legal siempre que no se contravengan una serie de circunstancias.

En esa grabación se puede conseguir el reconocimiento de una deuda pendiente, de una falta de responsabilidad o promesa incumplida, en una conversación donde la otra persona puede reconocer hechos que en otro momento podría negar. Esa grabación, como por ejemplo en el reconocimiento de una deuda, puede ser admitida como prueba siempre que no exista coacción. Esa muestra será llamada prueba “dubitada”, puesto que no se podrá determinar fehacientemente la identidad de los dos locutores por más que quede grabado su nombre; podrían alegar suplantación. Lo habitual en estos casos es requerir un peritaje vocal para determinar la identidad de los hablantes. El perito designado verificará en primer lugar

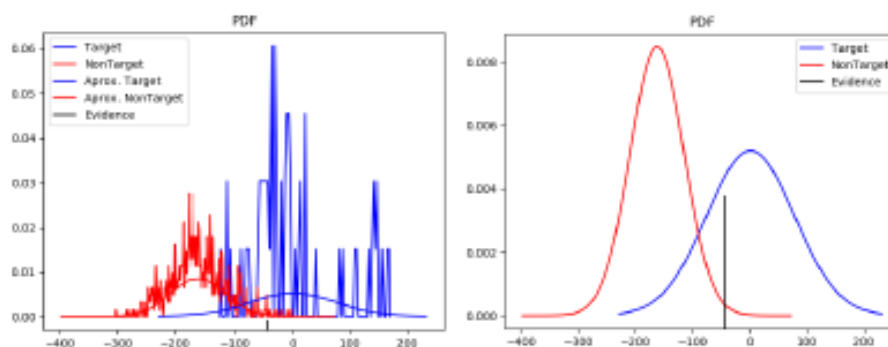
que las muestras no han sido manipuladas ni tratadas, que realmente se trata de muestras originales, y requerirá a los individuos en el Juzgado para realizar una muestra vocal que le permita identificar a los hablantes comparando las voces. Estas muestras serán las “indubitadas”. El siguiente proceso es realizar una comparativa simple y subjetiva por parte de un técnico o ingeniero, o designar la labor a una empresa que disponga de sistemas de Biometría Vocal Automática, con capacidad de generar huellas vocales a partir de las muestras indubitadas, y que realice la comparativa de muestras sin saber de los interlocutores nada más que su idioma, edad y lugar de residencia, para la selección de las muestras de voz para el banco de muestras. La utilización de sistemas automatizados es la forma más segura de realizar cotejos de voz, dado que no interviene la subjetividad humana. En primer lugar, se obtiene la huella vocal de cada uno de los sospechosos. En segundo lugar, se compara con las cientos o miles de voces anónimas almacenadas en la base de datos para ver el grado de fortaleza de la muestra, obteniendo unos valores de verosimilitud que, en caso de ser muy bajos respecto a las muestras anónimas, se determina que esa voz es única. Por último se coteja contra ella misma, para obtener un resultado de identidad confirmada entre la dubitada y la indubitada. El resultado es de **n** número de veces de igualdad entre la dubitada y la voz indubitada, trasladándose en porcentajes que suelen arrojar valores de más del 85% en el caso de nuestras en las que se determina de que se trata de la misma persona.

La experiencia en estos casos es muy positiva por el hecho de que ya no interviene la opinión de ninguna persona que pueda determinar la similitud entre espectrogramas y formas de onda de sonido, y por el hecho de que el procedimiento ha superado intentos de impostación de voz o modificación de la misma por parte de algunos sospechosos. Indicar que los valores obtenidos han estado muy cercanos al 90% en el peor de los casos, llegando incluso al 99,994%.

Se trata de un sistema muy preciso y que en poco tiempo lo tendremos completamente implementado en sistemas cotidianos donde la voz será nuestra llave: única y exclusiva.

Metodología analítica:

La consistencia, pulcritud, metodología y procedimiento científico de los informes se puede comprobar en la labor que realiza la empresa Biometric Vox en sus informes, gracias a la utilización de inteligencia artificial:



Esta analítica comprensible sólo por técnicos e ingenieros de sonido, se traslada a indicadores y datos empíricos, a porcentajes que pueden ser utilizados en una sentencia procesal.

MUESTRA	PUNTUACIÓN	LLR	PROBABILIDAD (%)
PROTOLS_TEXTO_01.WAV	-53,75	6,539	86,735
PROTOLS_TEXTO_02.WAV	-45,42	10,719	91,467
TASCAM_TEXTO_01.WAV	-48,39	8,967	89,967
TASCAM_TEXTO_02.WAV	-26,99	35,207	97,238
MEDIA	-43,54	12,018	92,318

Jurisprudencia:

La jurisprudencia que valida una grabación para ser lícita o ilícita estará en función de quién la realiza y de su participación en la misma. Siempre propongo al Juzgado realizar las muestras de voz indubitadas a las dos partes para evitar impugnaciones, basándome en sentencias como la STC 114/1984 de 29 de noviembre en la que se denegó el recurso de amparo, señalando que, *“El derecho al secreto de las comunicaciones, salvo resolución judicial, no puede oponerse, sin quebrar el sentido constitucional frente a quien tomó parte en la comunicación,... el bien constitucional protegido es la libertad de las comunicaciones, abarcando no solo el proceso de comunicación, sino también el mensaje, no solo el contenido de la conversación sino también la identidad subjetiva de sus interlocutores...”*

“...quien emplea durante su conversación un aparato que permita copiar aquella conversación, no está violando el secreto de comunicaciones, salvo que se entrara en la esfera íntima del interlocutor.” (recogido en el Art. 18.1 de la CE).

Esta sentencia se ve reforzada por otras del Tribunal Supremo: de 11 de mayo de 1994 y de 30 de mayo 1995 entre otras.

El Tribunal Supremo en STS de 20 de noviembre de 2014 declara respecto a la grabación de una conversación entre dos personas, siendo una de ellas quien estaba grabándola y por tanto participando, que *“...no puede considerarse referida a un ámbito propio y reservado que no puedan conocer los demás, porque las dos personas estaban hablando en el ámbito de su actividad, sin que hubiera nada referente a la vida íntima de ninguno de los dos.”*

Otra sentencia del Tribunal Supremo, la 3585/2016, así lo acredita. *“La aportación al proceso de grabaciones de conversaciones particulares realizadas por uno de sus protagonistas no vulnera el derecho al secreto de las comunicaciones, pues este derecho no puede esgrimirse frente a los propios intervinientes en la conversación”.*

Lo que en ningún caso se puede hacer es grabar conversaciones ajenas, salvo que se disponga de autorización judicial para ello. Sin autorización, se trataría de una grabación ilícita.

Informes periciales presentados a la Justicia:

NOTIFICACIÓN JUDICIAL - 2021104194880 - 22-06-2021
ANNA CLUSELLA MORATONAS

011010110110110

NOTIFICACIÓN JUDICIAL - 2021104194880 - 22-06-2021
ANNA CLUSELLA MORATONAS

Juzgado de Primera Instancia nº 7 de Sabadell
Avenida Francesc Macià, 34-36, Torre 3 - Sabadell - C.P.: 08208
TEL: 937400239
FAX: 937376645
EMAIL: juzgado7.sabadell@j1.gencat.cat
N.I.G.: 9818742130580696440
Procedimiento ordinario 319/2018 -2B
Matrícula Judicial número: 0266-16-00001

Partes demandante/representada: MARIA LOURDES RUIZ ROMERA
Procuradora: Anna Clusella Moratonas
Agradecido: JUAN ANTONIO SERRANO MONTSERRAT

Partes demandado/representado: ANTONIO ESPINOLA PADILLA
Procurador: Ramon Jofresa Lluch
Agradecido:

SENTENCIA Nº 125/2021

En Sabadell, a dieciocho de junio de dos mil veintuno.

D. Àngel Eduard Vacas Chalfoun, Juez del Juzgado de Primera Instancia n.º 7 de Sabadell en funciones de sustitución ordinaria, ha visto los autos de procedimiento ordinario registrados con el número 319/2018, promovidos por D.ª MARÍA LOURDES RUIZ ROMERA, representada por la Procuradora D.ª ANNA CLUSELLA MORATONAS y asistida por el Letrado D. JUAN ANTONIO SERRANO MONTSERRAT, contra D. ANTONIO ESPINOLA PADILLA, representado por el Procurador D. RAMÓN JUFRESA LLUCH y asistido por el Letrado D. VICENÇ LÓPEZ FIGUERAS, sobre acción de reclamación de cantidad derivada de contrato atípico.

ANTECEDENTES DE HECHO

PRIMERO.- El/la Procurador/a de la actora, en el nombre y representación que

metadatos del fichero de audio aportado en la demanda (los cuales se anexan al informe), y de la correspondencia de las voces grabadas con la de los litigantes no deja lugar a dudas. El perito expresa que el grado medio de coincidencia de la voz tomada en la muestra indubitada de la actora con la muestra dubitada (esto es, con la grabación aportada en autos) es del 98,511%, con un pico del 98,924%, mientras que, en el caso del demandado, el grado medio de coincidencia con la muestra dubitada es del 94,619% con un pico del 97,238%. El sr. RENÚ ESPADA equipara estos resultados con la plena correspondencia de las voces (señala que así sucede con datos superiores al 92% y que es muy raro que existan grados de coincidencia superiores al 99,5% porque una misma persona no siempre habla igual en todas las ocasiones, y que el menor grado de correspondencia en el caso del demandado obedece a una maniobra de impostación de la voz en el momento de captación de la muestra indubitada).

Tales conclusiones han de considerarse válidas. El rigor metodológico de la pericial ha de calificarse, a falta de otra prueba, de impecable: en particular, debe resaltarse la captación de muestras indubitadas con tres aparatos distintos, la composición de dichas muestras indubitadas de una lectura de un texto según un guion elaborado con expresiones similares a las de la muestra dubitada (a grabación aportada a los autos, como ya se ha dicho) y de una lectura de un texto sin relación con el pleito (que el sujeto grabado afrontaría con mucha mayor relajación) a fin de reducir el impacto de un posible falseamiento de la voz (en particular, por parte del demandado) que pudiera comprometer el resultado de la pericia, y la confrontación de las muestras indubitadas captadas con la dubitada mediante un proceso biométrico automatizado, sin intervención humana, en aras de la mayor objetividad. Es indicativo del grado de acierto de la pericial el hecho de que la defensa letrada del demandado, en el turno de conclusiones, apenas formulara objeciones a los resultados del informe.

b) Una conducta posterior, como son los pagos realizados por el sr. ESPINOLA PADILLA, por importes de 198,78 € y 274,39 € en los meses de agosto y septiembre de 2016, oportunamente documentados (docs. 3 y 4 de la demanda,

Las sentencias agradecen la metodología utilizada y la precisión de los datos expuestos. Se trata de una gran ayuda para que la justicia pueda utilizar estas herramientas para aportar a sus designas validez científica, donde los datos presentados no dan pie a la interpretación. En los últimos 3 casos realizados no se ha citado al perito por considerar las partes que los datos eran irrefutables. Y así lo ha expuesto algún juez en la sentencia:

La contundencia del informe pericial de sonido elaborado por el sr. RENÚ ESPADA en cuanto a las circunstancias de la conversación, a partir de los

(esto es, con la grabación aportada en autos) es del 98,511%, con un pico del 98,924%, mientras que, en el caso del demandado, el grado medio de coincidencia con la muestra dubitada es del 94,619% con un pico del 97,238%. El sr. RENÚ ESPADA equipara estos resultados con la plena correspondencia de

Tales conclusiones han de considerarse válidas. El rigor metodológico de la pericial ha de calificarse, a falta de otra prueba, de impecable: en particular, debe resaltarse la captación de muestras indubitadas con tres aparatos distintos, la

Sin lugar a dudas estamos ante la herramienta definitiva en el ámbito de la Biometría vocal.

Contacto para información y asesoramiento: Jordi Renú Espada, colegiado nº 0206 del CPAC

peritaudiovisual@gmail.com