

JOSÉ JUAN LUCENA MOLINA  
Comandante de la Guardia Civil.  
Jefe del Departº de Acústica e Imagen  
del Servicio de Criminalística de la Guardia Civil

## LA ACÚSTICA FORENSE

### INTRODUCCIÓN.

La acústica forense es una parte de la criminalística que engloba la aplicación de técnicas desarrolladas por la ingeniería acústica para el esclarecimiento de los delitos y la averiguación de la identidad de quienes los cometen. No existe más límite, en la selección y uso de esas técnicas, que el impuesto por la casuística pericial. Por tanto, los apartados que se exponen en este trabajo responden a los requerimientos judiciales por orden de importancia o, al menos, de frecuencia en las peticiones de informes periciales.

### RECONOCIMIENTO DE LOCUTORES POR LA VOZ.

#### Las primeras técnicas aplicadas en el ámbito forense [1].

La primera tecnología que recibió un nombre propio dentro del ámbito policial y forense la encontramos en los Estados Unidos en la década de los sesenta. La denominada técnica del “voiceprint” estuvo unida a un instrumento de medida, el espectrógrafo, cuya aparición tuvo lugar en 1941, empleándose en investigaciones de habla y música relacionadas con sistemas de comunicaciones. Este instrumento fue usado por Lawrence G. Kersta en los Laboratorios Bell, a principios de la mencionada década, para lograr identificar a las personas por la voz. El instrumento generaba un gráfico de la señal de voz teniendo en cuenta la información frecuencial, temporal y energética que se denominó espectrograma o sonograma. Kersta comenzó sus investigaciones partiendo de la hipótesis de que la voz de cada persona es tan única como la huella dactilar, pudiéndose determinar lo que él llamaba “huella acústica” utilizando el análisis espectrográfico.

El término “voiceprint identification” fue acuñado por Kersta y, dado su parecido con el término “fingerprint identification”, se difundió erróneamente la analogía de significado.

La pretendida unicidad de la voz surgió al pensar en el mecanismo fisiológico que produce esa señal acústica, así como en el proceso de aprendizaje del habla. Kersta sostuvo que la individualidad de la voz se fundamentaba en los anteriores procesos. En concreto, afirmaba que las cavidades vocales y los articuladores son aquellas partes del tracto vocal que determinan la individualidad. Las primeras son resonadores y se caracterizan por sus dimensiones y forma de acoplarse, resultando despreciable la probabilidad de que dos personas tengan idénticas estas propiedades. Los segundos (labios, dientes, lengua, paladar blando y músculos mandibulares) son controlados dinámicamente por cada persona y permiten que se puedan emitir los sonidos consonánticos y vocálicos. La probabilidad de encontrar a dos personas con idéntico patrón dinámico en el uso de los articuladores es extremadamente remota.

Aunque la identidad de un locutor está fuertemente correlada con las características fisiológicas y de comportamiento (hábitos lingüísticos, entonación de las frases, etc.) del locutor: variaciones en los tamaños de las cavidades del tracto vocal dan lugar a diferencias en las resonancias; variaciones en el tamaño y elasticidad de las cuerdas vocales se traducen en cambios en el valor medio de la frecuencia fundamental de los sonidos sonoros; variaciones en el paladar blando y el tamaño de la cavidad nasal producen diferencias espectrales en los sonidos nasales; la presencia de variaciones anatómicas atípicas (disposición y estado de los dientes, estructura del paladar, etc.), pueden dar lugar a sonidos atípicos o una nasalidad anormal. El problema de la detección de las características presentes en la señal de voz que individualicen al hablante no ha sido un problema fácil.

Los expertos en procesamiento de señal de habla subrayan que la información inherente al locutor se encuentra completamente imbricada con el resto de las informaciones presentes en la señal de voz y, en cierto modo, enmascaradas por ella. Si analizamos el proceso de la señal de voz captada por nuestro oído o por un transductor: paso de flujo de aire sobre el tracto vocal del locutor, radiación de dicho sonido al exterior, propagación acústica hasta el oído o un transductor, y en este último caso, paso de la señal de voz a través de sistemas eléctricos y electrónicos con sus respuestas en frecuencia y fase (amplificadores, filtros, canales telefónicos, conversiones analógico/digitales y digitales/analógicas, procesos de codificación, etc.), observamos que toda esa cadena de elementos introducen sus propios efectos, claramente perjudiciales para la identificación. Sin embargo, aún con todo lo anterior, los seres humanos somos capaces de identificar a las personas.

Los ingenieros han dividido los factores que afectan a la variabilidad intra locutor en dos grandes grupos:

- Variabilidad inherente a la propia señal de voz: derivada de factores directamente achacables al locutor, ya sean voluntarios o involuntarios (denominados factores intrínsecos: edad, estado emocional, estado físico, estar sometido a estrés, velocidad de articulación o tipo de habla -leída, susurrada, conversacional, etc.-; o bien, de circunstancias externas al locutor (llamadas factores extrínsecos: dispositivos de adquisición y transmisión de la señal de voz, ancho de banda, distorsión de canal, reverberación, ruido aditivo, etc.).
- Variabilidad debida al paso del tiempo.

A petición del FBI, un equipo de investigación dirigido por el Dr. Richard H. Bolt, y entre los que se encontraba Oscar I. Tosi, uno de los padres del “voiceprint”, llevó a cabo un trabajo consistente en evaluar la técnica del “voiceprint” desde los aspectos científico, técnico y jurídico. El estudio se denominó “On the theory and practice of voice identification” y fue publicado en febrero de 1979.

El informe sostuvo que la ciencia y la práctica de la identificación de voz se encontraban, entonces, en un estado claramente imperfecto, al igual que ocurría en el ámbito jurídico. Destacaba, entre otros, los siguientes puntos oscuros no resueltos en el ámbito de la ciencia: las características básicas que distinguen una voz de otra, la distribución de esas características dentro de grandes poblaciones, la susceptibilidad de la voz al control voluntario, como en la mímica, el disfraz, etc. En cuanto a la práctica, denunciaba la falta de fundamentos sólidos en los procedimientos de operación, de los métodos de aprendizaje y de las aseveraciones sobre el grado de exactitud de las conclusiones. De tal forma que -expresaba-, no era de extrañar que se argumentara en las sentencias sobre la inexistencia de principios claros que guiaran a los Tribunales a la hora de evaluar la validez de la metodología del “voiceprint”. Por tanto, a causa de lo que calificaba como estado relativamente inmaduro de la ciencia y técnica en este ámbito de

la criminalística, la aplicación del método del “voiceprint” en años anteriores había conducido a confusiones y controversias.

### **La aportación de los filólogos en el reconocimiento de locutores por la voz [2].**

La fonética, sintaxis, morfología, lexicología, dialectología, sociolingüística o la psicolingüística, por citar algunas de las materias filológicas más relacionadas con nuestro tema, aportan conocimientos nada despreciables en el ámbito forense. La determinación del origen geográfico de un hablante, edad, ámbito social en el que se desenvuelve, el descubrimiento de rasgos de su personalidad o hábitos de procedencia patológica o aprendida que lo caractericen de alguna manera, son informaciones muy útiles. También los filólogos han intentado aplicar sus conocimientos para resolver el problema de la identificación de la voz.

La fonética articuladora y, posteriormente, la fonética acústica, han partido del estudio del mecanismo de producción del habla para explicar la naturaleza del lenguaje humano. Para esas ciencias, ha sido también esencial descubrir los átomos del lenguaje denominados rasgos distintivos. Por tanto, estudian lo que hay de común en todos los hablantes.

Partiendo del apriorismo de aceptar la teoría de la unicidad de la voz, muchos filólogos han intentado aplicar sus conocimientos para detectar semejanzas y desemejanzas entre voces de diferentes locutores para llegar a realizar inferencias de identidad. Sus estudios suelen seguir la siguiente pauta:

- Determinación y medición de los parámetros acústicos más invariables en un locutor.
- Utilización de la teoría fonética a la hora de realizar un estudio auditivo de las voces.
- Análisis instrumental (espectrográfico) y estudio de los rasgos segmentales y suprasegmentales.
- Utilización de sus conocimientos de sintaxis, semántica, morfología, lexicología, dialectología, sociolingüística o psicolingüística para subrayar características que contribuyan a individuar al locutor.

La IAFPA (International Association for Forensic Phonetics and Acoustics) es la más activa asociación internacional en Europa que engloba a destacados expertos lingüistas en el campo del reconocimiento de locutores por la voz. Fue constituida formalmente en York (Inglaterra) en 1991.

### **Los avances de la ingeniería en el reconocimiento de locutores por la voz [3].**

Desde que la ciencia experimental empezó a crear modelos ideales convincentes de la producción del habla, los ingenieros desarrolladores de sistemas semiautomáticos o automáticos de reconocimiento de locutores por la voz han tratado de mejorarlos impulsando la investigación en parametrización relacionada con información individualizadora del locutor, englobando técnicas encaminadas a conseguir robustez paramétrica, es decir, a la extracción de parámetros característicos de la señal de voz que sean inmunes a la presencia del ruido contaminante. Igualmente, han tratado de mejorar el propio proceso de modelado, han avanzado en la normalización de las verosimilitudes y en el procedimiento estadístico de evaluación de la evidencia.

Los sistemas de reconocimiento automático que han conseguido mayor éxito con parametrización acústica han sido los basados en modelos de mezclas gaussianas (GMM). Son modelos basados en una suma ponderada (mezcla) de funciones densidad de probabilidad gaussianas de modo que la distribución de los vectores de parámetros

extraídos a partir de las alocuciones producidas por un determinado locutor quedan adecuadamente modeladas.

Han existido dos motivos principales para usar GMMs en el reconocimiento de locutores en el ámbito forense:

- La idea intuitiva de que las componentes individuales de una densidad multimodal son capaces de modelar las clases acústicas subyacentes en el proceso de identificación; esto es, que el espacio acústico que caracteriza la voz de un individuo se puede aproximar mediante un conjunto de clases acústicas (que representan conjuntos amplios de eventos acústicos) como pueden ser las vocales, las consonantes nasales o fricativas, etc. Estas clases acústicas denotan dependencia respecto a las configuraciones del tracto vocal específicas de cada locutor, siendo de gran utilidad a la hora de caracterizar a un hablante.
- La constatación empírica de los resultados alcanzados por estos sistemas en las evaluaciones internacionales llevadas a cabo por el NIST (National Institute of Standards and Technology), perteneciente al Departamento de Comercio de los Estados Unidos, y en la evaluación realizada en el año 2003 por TNO-NFI (The Netherland Forensic Institute) con voz forense.

[4]

#### **Estado de la técnica en reconocimiento automático de locutores con fines forenses.**

La multimodalidad biométrica, es decir, la existencia de numerosos sistemas de identificación biométricos orientados a la identificación de personas, no sólo cabe concebirla entre datos relacionados con diferentes órganos de los seres humanos, de los que cabe obtener información identificativa de un individuo (iris + huella dactilar + voz + etc.). También es posible emplear la fusión aprovechando toda la información que la naturaleza de la muestra analizada nos proporciona. En el caso de la voz, los diferentes niveles de información que se entremezclan en el habla humana son fuente de tratamiento biométrico separado, pudiendo llegar a demostrarse la independencia entre esos niveles, y puede culminarse el proceso creando sistemas multimodales exclusivos de la voz. Se trata de la metodología de análisis más moderna y fructífera de cuantas hoy se conocen para mejorar las tasas de rendimiento de los sistemas automáticos de reconocimiento de locutor.

La fusión multimodal vía SVM (Support Vector Machines) se está empleando con gran éxito en numerosos campos de la biometría, también en voz. En la actualidad se trabaja con sistemas de fusión independientes de usuario, dependientes de usuario y adaptados al usuario. Los mejores resultados se alcanzan con esta última opción.

Las diferentes propuestas tecnológicas de los sistemas de reconocimiento automático de locutor que actualmente compiten en NIST son, resumidamente, las siguientes:

- Sistemas de reconocimiento fonético de locutores por modelado estadístico del lenguaje (con n-gramas) en el tiempo.
- Sistemas basados en GMMs parametrizando el pitch y la energía.
- Sistemas NERF (Non-Uniform Extraction Region Features): las regiones de extracción de parámetros son segmentos de la forma de onda limitados por ciertos eventos; en cada región los parámetros describen pausas (número y longitud máxima), F0 (máximo, rango y pendiente) y Energía (máxima, mínima, rango y pendiente).
- Sistemas idiolectales: se reconoce el uso particular, por cada hablante, de los elementos del lenguaje.

- Sistemas de reconocimiento prosódico de locutor por tokenización de pendientes (F0, E).
- Sistemas de reconocimiento fonético-acústico basados en GMMs o SVMs.

Existe una sustancial potencialidad de mejora de los sistemas automáticos de reconocimiento combinando los anteriores sistemas. No obstante, siguen siendo los mejores sistemas individuales los sistemas fonético-acústicos GMM o SVM con parámetros MFCC.

### **La formulación de conclusiones en informes de voz realizados con sistemas automáticos [5].**

Siempre han existido intentos de establecer convenciones en torno a los niveles de certeza sobre las opiniones vertidas en los informes forenses. Si queremos ser científicamente serios, cualquier convención sobre la forma de expresar conclusiones en los informes debe estar firmemente enraizada en la lógica formal.

Durante muchos años, la forma tradicional de expresar las conclusiones en criminalística ha estado ligada a la formulación de escalas verbales de probabilidad. Sin embargo, la amplísima crítica llevada a cabo contra ellas en foros científicos en todo el mundo ha hecho que se tienda paulatinamente a implantar la inferencia bayesiana incorporando información subjetiva en el problema identificativo que se analiza.

Debe tenerse en mente que “la metodología bayesiana no pretende calcular probabilidades ‘verdaderas’: se trata de un método eficaz para analizar, criticar y chequear la coherencia de las opiniones de la gente, ayudándoles a revisar sus opiniones de forma coherente. Nada más, y nada menos.” (Taroni y otros, 2001).

Existe una interpretación de la probabilidad consistente en considerarla como una medida de creencia, lo cual es una propiedad de un individuo concreto. Esta interpretación mantiene que la probabilidad es subjetiva. Por el contrario, la definición frecuentista de probabilidad está basada en interpretarla como la propiedad de una secuencia. Todo aquél que observe esa secuencia estará de acuerdo con su valor. Se trata de una interpretación objetiva, si bien no está exenta de cierta idealización porque es imposible en la práctica realizar un número infinito de pruebas, en idénticas circunstancias y del mismo modo, siendo entre sí independientes.

Lindley ha aportado un excelente argumento en defensa de la probabilidad subjetiva. Comenta que mientras la objetividad es considerada como un marchamo de lo que puede llamarse ciencia, la subjetividad es algo que el científico debe evitar a toda costa. Sin embargo, observa, si una hipótesis se cree a priori digna de tener en cuenta en la ciencia experimental pero suscita controversia, al mismo tiempo, se realizan experimentos repetidos hasta que exista un acuerdo generalizado. Al igual que pasa con los científicos, los miembros de un Jurado pudieran tener diferentes creencias apriorísticas.

Consideremos el siguiente ejemplo: se tiene una información inicial  $I$  y dos personas tienen diferentes creencias sobre la veracidad de un evento  $G$ , es decir, sus probabilidades  $P(G/I)$  no son iguales. Se produce una evidencia  $E$  y se calculan las probabilidades a posteriori de cada uno de ellos  $P(G/E, I)$ . Se puede demostrar que la evidencia  $E$  acerca los valores de sus probabilidades a posteriori. A medida que fuera mayor la evidencia  $E$ , la cercanía entre los valores de las probabilidades a posteriori también sería mayor. Lindley señala que es eso lo que exactamente ocurre en un Tribunal. Los miembros del Jurado llegan a una convicción común sobre lo que se dilucida gracias a lo que oyen en el juicio oral. Nada les fuerza a que lleguen a un acuerdo y, sin embargo, la experiencia demuestra que ese acuerdo común es generalmente alcanzado.

De Finetti, escribió en 1952 lo siguiente: “aceptamos la definición de probabilidad subjetiva como el grado de creencia que alcanza un individuo respecto a la ocurrencia de un cierto evento ... Vale la pena considerar el significado preciso, técnico, sobre la diferencia entre lo objetivo y lo subjetivo. De hecho, estimo que muchos errores de interpretación y muchas discusiones acaloradas derivan de una interpretación más o menos inconsciente y vaga sobre lo objetivo, en cuanto sinónimo de ‘fundado, razonable, serio’, mientras que solemos llamar subjetivo a un juicio ‘no fundado, precipitado, improvisado sacado de una chistera’. Nada puede estar más lejos de las intenciones de la teoría de la subjetividad: su finalidad se centra en el estudio y promoción de evaluaciones sobre la probabilidad con la misma profundidad con que lo hacen quienes defienden evaluaciones objetivas y, si fuera necesario, hasta con un mayor sentido de responsabilidad respecto a no hacerse falsas ilusiones de que pudieran obtenerse probabilidades de naturaleza objetiva. Aquellos a los que no les agradan o desprecian las probabilidades subjetivas y creen que lo remedian utilizando probabilidades objetivas, no alcanzarán mejores resultados...”.

Una de las críticas más comunes a la aproximación bayesiana para evaluar la evidencia (y también para otras áreas de análisis estadístico) es el uso de probabilidades subjetivas. Sin embargo, la subjetividad no puede interpretarse como arbitrariedad. Un punto de vista repetido es que si la probabilidad es una medida de creencia, entonces tiene que ser arbitraria, porque una persona pensará una cosa y otra persona otra cosa diferente. Sin embargo, el que el grado de creencia sea personal no significa que sea arbitrario. La probabilidad puede representar en qué medida una persona cree en algo como cierto, pero eso está basado en todo lo que esa persona conoce al respecto. Esa información suele ser distinta de una persona a otra, de forma que es fácil que tengan grados de creencia distintos sobre lo mismo. Esa diferencia no es una diferencia arbitraria. La implicación consiste en que el grado de creencia está condicionado a lo que conoce cada persona. Por tanto, todas las probabilidades son condicionales y esas condiciones deben ser explícitamente establecidas. También puede decirse que todas las probabilidades son personales.

En la actualidad se intenta expresar la conclusión de un informe de voz realizado con sistemas automáticos mediante el uso de la relación de verosimilitudes o LR. Además se incorpora al informe una curva Tippett que refleja el comportamiento del sistema ante voces con características acústicas similares a las utilizadas como voces dubitadas e indubitadas. Al mismo tiempo, se avanza en la determinación del grado de fiabilidad de la valoración de la evidencia hecha mediante LR gracias al empleo de técnicas de calidad.

## **AUTENTIFICACIÓN DE GRABACIONES [6].**

### **Terminología básica en autenticación de grabaciones.**

La Sociedad Americana de Acústica cuenta con un Grupo de Trabajo (WG-12 Working Group on Forensic Audio, 1991) que fue publicando estándares que pudieran servir de guía para una buena práctica técnico-científica relacionada con la elaboración de informes periciales de acústica forense. Dentro de los objetivos perseguidos por el citado Grupo de Trabajo se encuentra la autenticación de grabaciones de audio y se cita expresamente en el prólogo del documento denominado AES27 (1996) que se siguen los criterios expuestos en un trabajo realizado en 1974 para el Tribunal de Distrito del Estado de Columbia de los Estados Unidos.

Se detallaron las normas técnicas de referencia que se tuvieron en cuenta en la elaboración del documento y se definieron una serie de términos que son de especial relieve para la correcta intelección de los trabajos desarrollados por los peritos: grabación

segura, grabación autenticada, análisis de autenticidad, magnetófono original, grabación original y grabación cuestionada, entre otras.

El documento AES43 (2000) publicado en noviembre de 1999, fue el fruto de un trabajo liderado por A. Pellicano en el seno del Grupo de Trabajo WG-12. Recogió un procedimiento técnico para autenticar grabaciones de audio analógicas inspirándose en las publicaciones de autores expertos en la materia.

Las definiciones básicas de este último documento completaron las primeramente expuestas en el AES27, las cuales, por supuesto, asume.

En el seno del Grupo de Trabajo “Speech and Audio Analysis” de ENFSI se constituyó un área de interés común en esta materia en la reunión anual de 1999 celebrada en La Haya (Holanda). La Guardia Civil expuso un documento que trató de resumir los principales avances tecnológicos y de procedimiento que, hasta esa fecha, habían tenido lugar en el mundo.

En la reunión anual del año 2005 celebrada en Wiesbaden (Alemania) se presentó un dossier conjunto, elaborado por los representantes del NFI (The Netherland Forensic Institute), BKA (Bundeskriminalamt) y Guardia Civil donde se abordó, de forma expresa, la problemática inherente a la terminología en los informes de autenticación de grabaciones de audio.

La Guardia Civil emplea la siguiente terminología, de uso para cualquier tecnología de audio:

*Grabación original*: cualquier soporte de grabación o parte del mismo que contenga, de forma permanente, los sucesos acústicos capturados por un sensor y registrados en tiempo real, justamente en los instantes en que la señal acústica se iban produciendo.

*Grabación íntegra*: cualquier soporte de grabación o parte del mismo que contenga, permanentemente y de forma continua, los sucesos acústicos capturados por un sensor y grabador.

*Grabación precisa*: cualquier soporte de grabación o parte del mismo que contenga, permanentemente, los sucesos acústicos capturados por un sensor y grabador, desviándose ligeramente o dentro de los límites de tolerancia de un estándar tecnológico.

*Grabación auténtica*: grabación que merece ser considerada verosímil a juicio de la Autoridad Judicial. Se trata de un concepto que estimamos que está fuera del alcance de un experto forense porque su naturaleza es decisoria y sobre una cualidad de algo sobre la que sólo la Autoridad Judicial puede pronunciarse. Solamente puede llegarse a ello tras un proceso de decisión lógica en el que resulta indispensable, en aras a ser calificado de respetuoso con la lógica racional, cumplimentar las leyes de la probabilidad. En este sentido, nos encontramos ante un proceso de decisión semejante al de la identificación de una persona a partir de un vestigio.

### **Examen de autenticidad.**

Cuando un Juez solicita de un perito que examine si una grabación puede ser calificada de auténtica, lo que en nuestra opinión es importante es preguntarse por los siguientes elementos:

- Indicios de originalidad.
- Precisión.
- Integridad.
- Ausencia de alteraciones inexplicables, entendiéndose por tales las manipulaciones o los fenómenos cuya explicación escape al control técnico de los peritos, teniendo en cuenta su formación y experiencia profesional. Lo primero entrañaría

rechazar la autenticidad de aquellos pasajes afectados por las alteraciones y lo segundo produciría duda.

Las etapas que, de forma general, se siguen en un estudio de autenticidad en grabaciones de audio en soportes magnéticos, son las siguientes:

- Examen físico del soporte de la grabación.
- Escucha crítica.
- Análisis de forma de onda.
- Análisis frecuencial.
- Análisis espectrográfico.
- Análisis espacial.

La señal de video analógica precisa, además, de un análisis específico de la señal relacionada con las imágenes y la de sincronización, en función de la norma técnica utilizada por el equipamiento que se perite.

La autenticación de grabaciones digitales tiene una peculiaridad especial debido a la naturaleza de la señal grabada en los soportes. La autenticación digital necesita de la incorporación de sistemas de seguridad informáticos como la firma electrónica, la técnica del “timestamping”, los algoritmos “hashing” u otros sistemas de análogas características.

### **Conclusiones en autenticación de grabaciones.**

Hace muchos años que los expertos forenses están convencidos de que no es posible conseguir la certeza absoluta a la hora de abordar de forma objetiva un examen de autenticación de grabaciones.

Si ningún examen pericial de esta naturaleza puede llegar a la certeza absoluta, resulta muy importante conocer el grado de confianza que podemos alcanzar sobre una particular creencia. Aquí juega un papel esencial la aproximación bayesiana, puesto que considera a las probabilidades como medidas de creencia sobre la ocurrencia de un evento particular (a estas probabilidades se las denomina probabilidades subjetivas).

La aproximación bayesiana permite combinar probabilidades objetivas, basadas en datos, con probabilidades subjetivas, para lo que la formación y experiencia de los expertos resulta relevantes.

Los Jueces también necesitan utilizar probabilidades subjetivas relacionadas con la credibilidad de los testimonios de testigos oculares.

Por tanto, las conclusiones deberían exponerse siguiendo la pauta propuesta por la valoración bayesiana de la evidencia. Por ejemplo: “una vez examinada la evidencia con los análisis establecidos por la instrucción técnica del Departamento, lo observado es más probable encontrarlo si la hipótesis de que la grabación es auténtica es cierta que si la hipótesis alternativa y, en este caso, complementaria, lo fuera”.

### **LIMPIEZA DE GRABACIONES [7].**

La limpieza de grabaciones forense tiene una perspectiva propia que la diferencia, en algún sentido, de cualquier trabajo análogo. No interesa resaltar, ni siquiera respetar, la calidad acústica de la voz. Lo que interesa es mejorar su inteligibilidad, por tanto, las técnicas de filtrado han de aplicarse bajo ese prisma. En otras ocasiones el objetivo principal no es la voz sino el ruido de fondo.

En el contexto de grabaciones en soporte magnético, para una correcta aplicación de las técnicas se comprueba previamente el estado de los tornillos de las carcasas, la



longitud de la cinta, el estado de las bobinadoras y el de la cinta en sí misma. Todo ello para asegurarse de que no habrá problemas en la reproducción de la señal. Si pudiera haberlos, se realiza un transporte de la cinta a otra carcasa. Para evitar borrados accidentales de la grabación original se rompen las pestañas que previenen contra dichos borrados en los formatos de cinta que las posean.

Con ayuda de técnicas de microscopía con luz polarizada y cristales reveladores, u otras de análoga eficacia, se visualizan las pistas de la grabación sobre el soporte. Este análisis sirve para elegir la mejor opción en orden a reproducir la señal grabada. Resulta de gran importancia fijarse si las pistas pudieran tener un desplazamiento severo respecto a la cabeza o cabezas que tengan que reproducir su contenido y tener en cuenta si la grabación es monoaural o estereofónica.

El ajuste de azimuth consiste en ajustar la cabeza del magnetófono que reproduce al ángulo formado por los dominios respecto a la línea vertical respecto a los bordes de la cinta, de tal modo que la respuesta en altas frecuencias sea la máxima posible. Como criterio general, a menor velocidad y mayor anchura de pista se obtienen mayores pérdidas por desajuste de azimuth.

Se realiza un análisis de la velocidad de reproducción siguiendo la evolución de un tono grabado sobre el soporte magnético. Los tonos más frecuentes son los de red (50 Hz y armónicos), los telefónicos o algunos tonos procedentes de interferencias eléctricas. Podemos medir así qué porcentaje de deriva, respecto a la velocidad oficial de reproducción del equipo, es la que se está produciendo.

Una de las etapas más importantes del análisis es la escucha crítica, que sirve para especificar el tipo de grabación (microfónica, telefónica, etc.) y localizar los tramos que requieren mejora de la inteligibilidad. El contenido de lo grabado se etiqueta para aplicar distintos filtros en función del ruido que haya que eliminar o reducir.

A continuación, se realiza un análisis en frecuencia antes de aplicar filtros digitales de limpieza, llevándose a cabo un promediado exponencial de la FFT en zonas de silencio y en zonas con voz. Gracias a ese examen podemos determinar: rango de frecuencia del habla; relación señal/ruido; presencia de tonos discretos; posible presencia de ruido en determinadas bandas; y efectos de convolución.

Resulta también muy importante controlar dinámicamente la señal de habla y limitar el rango de frecuencias al canal de transmisión o grabación utilizando un limitador que impida saturaciones, un filtrado del ancho de banda del canal, y un compresor/expansor.

Los distintos tipos de ruido como las distorsiones no lineales ocasionada por niveles de grabación fuera del rango tolerable, el uso de sistemas de transmisión o recepción de escasa calidad, fallos de dispositivos electrónicos, cambios convolucionales producidos por alteraciones de frecuencia lineales del sistema de grabación, canal de transmisión, o de entorno acústico como una habitación reverberante, ruido de sistema como el producido por el equipo reproductor y sistemas de transmisión (por ejemplo, el zumbido de la red o el lloro y centelleo), ruido de entorno como el aditivo a la señal de voz antes de que sea captada por un micrófono (por ejemplo, ruido de televisión, radio, ventilación, tonos, manejo del micrófono, etc.), se combaten con filtros digitales especialmente adaptados a los fenómenos descritos.

## **MEDIDAS ACÚSTICAS FORENSES [8].**

Desde que el delito medioambiental está tipificado en el Código Penal español, la contaminación acústica pasa a ser una tarea técnica forense. No sólo se trata de medir los índices de contaminación o aislamiento acústicos siguiendo las recomendaciones técnicas

internacionales (normas ISO, UNE, IEC) y las dispuestas en la Directiva 2002-49-CE del Parlamento Europeo, de 25 de junio de 2002, sobre evaluación y gestión del ruido ambiente, la Ley 37/2003, de 17 de noviembre, del Ruido, y en las Órdenes de las Consejerías de Medio Ambiente de las Comunidades Autónomas fundamentalmente, sino que el análisis de los resultados permite evaluar en qué medida el ruido afecta a la salud de las personas. La Autoridad Judicial no está resolviendo, en esos casos, un expediente abierto por una supuesta infracción administrativa.

Las técnicas de acústica arquitectónica, medioambiental, industrial o en campo libre encajan perfectamente en la acústica forense.

#### **Técnicas de acústica arquitectónica.**

Se realizan mediciones de aislamiento acústico en interiores de edificios. Se suele utilizar una fuente sonora en el lugar de donde procede el ruido y se miden los tiempos de reverberación e índices de reducción sonora aparente. Con ello se pretende averiguar si un sonido procedente de un habitáculo puede ser percibido en otro contiguo, o incluso, en el caso de la voz, si pudiera ser la conversación inteligible midiendo el índice RASTI. Siempre que intervengan seres humanos, como actores principales en el planteamiento de la pericia, han de ser sometidos a una audiometría.

#### **Técnicas de acústica medioambiental e industrial.**

Normalmente se suelen emplear sonómetros integradores que miden diferentes niveles de presión sonora equivalente y con ponderación A, aunque en ocasiones se necesitan medir niveles relacionados con ruidos impulsivos. Se suelen realizar también mapas de ruido cuando existen áreas habitadas afectadas por ruido aéreo: inmediaciones de aeropuertos o infraestructuras ferroviarias, principalmente.

#### **Técnicas de acústica en campo libre.**

Cuando el ruido es provocado por un disparo o una explosión y la Autoridad Judicial requiere una pericia sobre la posibilidad de su audición por un testigo, además de la audiometría se hacen necesarias una serie de pruebas acústicas que intenten objetivizar, en la medida de lo posible, la posibilidad real de oír esos sonidos a la distancia y en las circunstancias descritas en su testimonio.

### **CONCLUSIONES.**

La acústica forense es ya una disciplina de la criminalística que ha alcanzado un grado de madurez científica suficiente como para considerarla imprescindible en todo laboratorio especializado. La casuística pericial obliga, a los integrantes de esos laboratorios, a tener amplios conocimientos de fonética, tecnología del sonido y del habla en particular, patologías del habla y procesado de señal.

Donde aún se tiene que experimentar un crecimiento notable es en la práctica pericial de muchos laboratorios hoy día constituidos. Esa falta de experiencia se vislumbra en la práctica inexistencia de protocolos internacionales que armonicen los métodos y procedimientos de trabajo pericial en los distintos laboratorios. No existe consenso todavía en aspectos, tan básicos, como en terminología en autenticación de grabaciones. Sin embargo, no faltan propuestas y foros de discusión, por lo que el futuro es esperanzador.

Debido al constante y creciente desarrollo tecnológico en acústica aplicada, los avances en cada una de las áreas como reconocimiento de locutores por la voz, limpieza de grabaciones, autenticaciones de grabaciones o medidas acústicas relacionadas con el

ruido, por ejemplo, obligan a los peritos a tener una necesidad imperiosa de actualizar sus conocimientos y a fomentar la investigación criminalística.

También se ha constatado la importancia, cada vez mayor, de incorporar técnicas de inferencia estadística en la formulación de conclusiones periciales, especialmente los relacionados con identificación de personas por la voz y en autenticación de grabaciones.

Por último, dado que la criminalística tiene una perspectiva propia a la hora de emplear la tecnología y los conocimientos científicos, y su incidencia social es cada vez más relevante y frecuente, se vislumbra un futuro en el que será necesario dar vida a proyectos como la creación de centros específicos de formación criminalística o de titulaciones especiales.

### **Bibliografía.**

[1] “*On the theory and practice of voice identification*”, Richard H. Bolt y otros. Editado por National Academy of Sciences, Washington, D.C., U.S. Department of Commerce. National Technical Information Service (NTIS). Febrero de 1979.

Especialmente útiles los comentarios en el trabajo titulado “*Scientific Evidence in Civil and Criminal Cases*”, por Andre A. Moenssens y otros. Westbury, New York. The Foundation Press, Inc. 1995.

[2] *Fonética acústica de la lengua española*. Dr. D. Antonio Quilis Morales. Biblioteca Románica Hispánica. Editorial Gredos, Madrid. 1981.

Como publicaciones de referencia internacional: *Forensic Phonetics*, John Balwin & Peter French, Pinter Publishers, 1990 Londres y *Forensic Speaker Identification*, Philip Rose, Taylor and Francis Forensic Science Series, 2002.

[3] “*Influencia y compensación del entorno acústico en sistemas de reconocimiento automático de locutores*” y “*Técnicas de mejora de voz aplicadas a sistemas de reconocimiento de locutores*”. Tesis Doctorales dirigidas por el Dr. D. Luis A. Hernández Gómez, cuyos autores son el Dr. D. Joaquín González Rodríguez y el Dr. D. Javier Ortega García, respectivamente. Presentadas en la ETSIT de la Universidad Politécnica de Madrid.

[4] *NIST Speaker Recognition Evaluations*. Web: [www.nist.gov/speech/tests/spk/2004](http://www.nist.gov/speech/tests/spk/2004). Pueden consultarse resultados de años anteriores en páginas similares.

[5] “*Statistics and the Evaluation of Evidence for Forensic Scientists*”, C.G.G. Aitken y F. Taroni, John Wiley & Sons, 2ª edición, 2004.

[6] “*Authentication of Forensic Audio Recordings*”, Bruce E. Koenig, J.Audio Eng.Soc., Vol. 38, No. 1/2, 1990 January/February. Federal Bureau of Investigation, Engineering Research Facility, Lorton, VA 22079, USA.

[AES27-1996] J. Audio Eng. Soc., Vol. 44, No. 4, 1996 April.

AES recommended practice for forensic purposes - Managing recorded audio materials intended for examination. Document developed by WG-12 (Working Group of Forensic Audio).

[AES43-2000] J. Audio Eng. Soc., Vol. 47, No. 10, 1999 October.

AES standard for forensic audio - Criteria for the authentication of analog tape recordings. Document developed by WG-12 (Working Group of Forensic Audio).

[7] “*Enhancement of Forensic Audio Recordings*”, Bruce E. Koenig, J.Audio Eng.Soc., Vol. 36, No. 11, 1988 November. Federal Bureau of Investigation, Engineering Research Facility, Lorton, VA 22079, USA.

“*Panorámica de los esquemas de mejora de voz en presencia de ruido*”, Joaquín González Rodríguez, ATVS (UPM), Libro de Actas del Primer Congreso de la Sociedad Española de Acústica Forense, Madrid 5-6 octubre 2000.

[8] “*Manual de Medidas Acústicas y Control de Ruido*”, Cyril M. Harris, Editorial McGraw-Hill, 3ª edición, Diciembre 1995.