

Rueda de reconocimiento de voz

PROTOCOLO C. G. POLICÍA CIENTÍFICA

Earwitness Identification

VOICE LINE-UP PROTOCOL, SPANISH SCIENTIFIC POLICE

MARTA ALONSO CARROZA

*Laboratorio de Acústica Forense
Comisaría General de Policía Científica
cpg.acustica@policia.es*

CARLOS DELGADO ROMERO

*Laboratorio de Acústica Forense
Comisaría General de Policía Científica
carlos.delgado@policia.es*

Resumen: Este artículo presenta el procedimiento establecido por el laboratorio de Acústica Forense de la Comisaría General de Policía Científica para la práctica de Ruedas de Reconocimiento de Voz (RRV). Esta técnica es utilizada en aquellos hechos delictivos en los que la víctima o testigo no pudo ver el rostro del sujeto agresor pero sí tiene el recuerdo de su voz. Se trata de una tarea de discriminación perceptivo-auditiva cuya finalidad es la asociación de estímulos de habla que la víctima o testigo evoca de su memoria a corto plazo con otros del supuesto agresor/a obtenidos mediante una grabación. En dicha tarea también se incluyen registros de voz de otros sujetos de la misma comunidad lingüística y similar característica vocal (timbre, tono, plano expresivo, etc.) que el locutor investigado.

Palabras clave: rueda de reconocimiento de voz (RRV), estímulo, memoria a corto plazo, timbre.

Abstract: This lecture introduces the procedure followed by the Forensic Acoustics Laboratory of the Spanish Scientific Police for the practice of recognition skills using voice line-ups. This procedure is used in those criminal acts in which the victim or witness could not see the face of the aggressor but has the memory of his/her speech. It is a job of aural perceptual discrimination whose purpose is the association of speech stimuli that the victim or witness elicits from his or her short-term memory with others of the presumed aggressor obtained through a recording. This task also involves voice recordings of other subjects of the same linguistic community and similar vocal characteristics (timbre, tone, manner of speaking, etc.) as the speaker under investigation.

Keywords: voice line-up, stimulus, short-term memory, timbre.

I. Introducción

...«Si volviese a escuchar su voz la reconocería sin lugar a dudas»...

Frases o expresiones de este tipo suelen ser un común denominador entre las víctimas o testigos que la Autoridad Judicial o los investigadores policiales consideran posibles candidatos a una Rueda de Reconocimiento de Voz (RRV).

Esencialmente, la técnica de reconocimiento de voz mediante rueda pretende objetivar las decisiones de la víctima o testigo a través de la emisión de múltiples juicios de asociación entre el registro que tiene en su memoria de la voz producida por el agresor en la escena del crimen y los diferentes estímulos auditivos que se le presentan en grabaciones de cadenas de audio editadas convenientemente. En dichas cadenas se ubican alternativamente la voz del supuesto agresor/a y otras voces perceptivamente similares que en adelante denominaremos *voces contraste*.

Una rueda de reconocimiento de voz podría considerarse un enfoque pragmático de la Psico Acústica, cuyo objeto es el «estudio de las respuestas subjetivas a estímulos acústicos, especialmente en tareas de detección y discriminación... el último estadio es la identificación, que requiere relacionar el estímulo que se nos presenta con una representación que tenemos en la memoria» (Marrero Aguiar [2001: 28]).

Para encontrar los antecedentes más próximos a los fundamentos de esta práctica hemos de acudir a las ruedas de reconocimiento visual. Por ello es labor esencial para los fonetistas forenses conocer en profundidad los trabajos de campo realizados por psicólogos especializados en materia de memoria visual y auditiva como Wagenaar, Wells y Yarmey.

Muy probablemente el elemento de mayor complejidad relacionado con la técnica presentada es el propio proceso de percepción auditiva, un proceso de construcción activo y desde luego más complicado que la simple audición. Percibimos tanto de forma categorial, mediante la etiquetación y categorización de diferentes estímulos, como de forma continua, constatando las diferencias relevantes del proceso de coarticulación del habla.

Las principales variables que intervienen en el desarrollo de la percepción auditiva son:

- El sonido, aquí entendido como la sensación que recoge el oído producida por estímulos de cambio de presión de las moléculas del aire.
- El oído, órgano transductor que transforma las vibraciones del aire

en energía eléctrica, dirigiéndose esta al cerebro a través del nervio auditivo.

- La psicoacústica, disciplina que básicamente se ocupa de la relación entre las características del sonido y la sensación auditiva que provoca.
- Descodificación de los sonidos lingüísticos en la que participan la memoria a corto, medio y largo plazo.

El punto de partida de una RRV es el acto perceptivo, es decir, un ejercicio subjetivo donde el material de estudio está formado por sensaciones interpretadas. En consecuencia, se valoran, por un lado, aquellas variables intrínsecas a las características de la víctima o testigo del hecho, que tienen que ver con sus destrezas auditivas, cognitivas y de atención, su capacidad de memoria y con la situación de estrés acaecida en el transcurso de los hechos. Por otro, las variables externas al receptor del mensaje relacionadas con el entorno acústico, tanto las que puedan afectar a las condiciones de registro del mensaje, como las relacionadas con la forma de emisión del mensaje (voz susurrada, gritada, simulada, etc.).

Considerando la diversidad de factores que intervienen en el procedimiento resulta más que conveniente que el mismo sea llevado a cabo por personal especializado en lingüística, fonética y psicología, todo ello en aras de aportar el mayor grado de fiabilidad posible a la prueba (Künzel, 1994).

Más allá del relato recogido en el libro del Génesis donde un ya ciego Isaac realiza un exitoso reconocimiento de su hijo impostor Jacob, los comienzos más formales del uso del reconocimiento de voz en rueda como vestigio ante los tribunales, datan de pasada la mitad del siglo xvii en Inglaterra. Desde entonces hasta nuestros días hemos asistido a un desfile de iniciativas de materialización de la prueba.

Con el transcurso del tiempo se ha podido constatar que la falta de familiaridad con el procedimiento por parte de los órganos judiciales, así como su limitada aplicabilidad e intrínseco carácter subjetivo de la cuestión, generan incertidumbre y falta de claridad en su papel probatorio. Por todo ello los científicos forenses han realizado un esfuerzo continuado para reducir el umbral de certeza de la prueba mediante el establecimiento de pautas y criterios comunes.

En 1995 la Asociación Internacional de Fonética Forense (IAFP) se comprometió a revisar los elementos específicos de la prueba para su correcto desarrollo. Para ello se constituyó un comité de trabajo encargado de desarrollar unos criterios comunes de actuación ante la aplicación de esta técnica,

siendo los miembros del comité A. P. A. Broeder, P. French, H. Holien, R. Huntley y H. Künzel.

2. El Laboratorio de Acústica Forense del Cuerpo Nacional de Policía

A principios del presente milenio, la Sección de Acústica Forense de la Comisaría General de Policía Científica, como miembro del Grupo de trabajo para el Análisis Forense de Habla y Audio (WGFSAA) de la red ENFSI, crea un protocolo RRV siguiendo las recomendaciones formuladas al respecto por nuestra comunidad científica y en línea con las directrices generales establecidas en el seno de la mencionada Red Europea de Institutos de Ciencias Forenses.

Llegado este punto, resulta pertinente reseñar que la eficacia de una rueda de reconocimiento siempre vendrá influenciada por las circunstancias particulares del caso (uso de máscaras, ambiente acústico, factor canal, variedad lingüística, etc.) así como el propio desarrollo de la RRV (disponibilidad de *voces contraste*, lapso temporal transcurrido, etc.). Por tales razones el protocolo es diseñado *ad hoc* en el sentido de procurar una óptima adecuación al contexto de cada escenario.

2.1 En la práctica

El Laboratorio de Acústica Forense recibe por parte de juzgados o grupos operativos policiales requerimientos compatibles con la realización de una RRV. Con carácter general, el procedimiento para llevar a cabo una rueda de voz comprende las siguientes etapas:

- Realización por parte de personal cualificado del laboratorio de una entrevista personal con la víctima o testigo del hecho en sede judicial con la finalidad de recabar aquellas informaciones de interés que, desde una óptica perceptiva, puedan afectar a dicha persona. A este respecto, se tienen en cuenta una serie de premisas encaminadas a conocer o explorar:
 - o El tiempo transcurrido entre el suceso y la prueba de reconocimiento por parte de la víctima, ya que en términos generales

- la memoria auditiva va disminuyendo con el tiempo (Hollien, Huntley, Künzel y Hollein, 1995).
- Las frases, locuciones y características del discurso proferido por el sujeto agresor, atendiendo al plano expresivo, semántico, características idiolectales y sociolectales, referencias de edad, ubicación diatópica o diastrática, etc.
 - Si se trata de una voz familiar, conocida o por el contrario es una voz desconocida.
 - El entorno acústico asociado a la escena del crimen durante el transcurso de los hechos.
 - La posible influencia de agentes tóxicos externos en el habla del sospechoso.
 - Las capacidades perceptivas del testigo o la víctima relacionadas con la discriminación de las cualidades de la voz.
- Evaluación de la entrevista de cara a estimar la viabilidad de la prueba, confeccionando en su caso, un cuerpo de voz con aquellas frases que la víctima o testigo rememoró durante dicha entrevista, las cuales fueron supuestamente emitidas por la persona investigada durante la comisión del hecho delictivo.
 - Realización de una toma de muestras de voz de la persona a quien la Autoridad Judicial atribuya la autoría del hecho delictivo. La grabación consistirá en la repetición de las frases contenidas en el cuerpo de voz previamente elaborado, teniendo en consideración las circunstancias acústicas y expresivas descritas por la víctima o testigo. De la misma forma, se registrará una conversación espontánea con el objetivo de obtener una muestra más natural del habla.

Si el equipo científico considera que la grabación de la persona investigada reúne las exigibles referencias de calidad, se procede a efectuar un análisis sobre los componentes fundamentales y características individuales de la voz objeto de estudio, de cara a la posterior obtención de muestras en condiciones similares provenientes de otros hablantes (*voces contraste*).

Una vez registradas las *voces contraste* se procede a la confección definitiva de una «cadena editada de voces-problema» que contendrá concatenados fragmentos de discurso de la persona/s sospechosa/s y de las *voces contraste*. Dichos fragmentos se distribuyen de forma aleatoria y se etiquetan con códigos correlativos (V1, V2, V3...).

Las «voces contraste» serán elegidas buscando cierta compatibilidad perceptiva con el habla del investigado, considerando tanto rasgos lingüísticos

y paralingüísticos de la voz, como pueden ser la variedad dialectal, el tono, la velocidad discursiva y el proceso de articulación (Smith y Baguley, 2014).

En cuanto al número de «voces contraste» a utilizar, se considera adecuado que este se sitúe en torno a cinco hablantes (Broeders y van Amelsvoort, 1999).

La obtención de las «voces contraste» será llevada a cabo por personal especializado del laboratorio adaptándose a las circunstancias del contexto en que se produjeron los hechos. Tales circunstancias, así como las características relativas al criterio seguido en la selección de las *voces contraste* y el estudio de la voz sospechosa, se hacen constar en el informe de resultados.

Como ya se ha referido anteriormente, a cada una de las voces contraste o «impostores» se les asigna un número de identificación: V1, V2, V3, V4, y V5. Los cinco locutores leerán las frases del cuerpo de voz previamente confeccionado, intentando adecuarlas a los rasgos prosódicos descritos por la víctima o testigo.

Una vez grabadas las emisiones de los «locutores contraste» y las del investigado, se procede a la confección de las cadenas editadas de voz o *cadenas problema*. Una cadena tipo —depende de cada caso— puede incluir en torno a seis estímulos auditivos, cinco voces contraste más la voz de la persona investigada. La secuencia de aparición de los distintos estímulos se realiza de forma aleatoria en cada cadena, procurando evitar la repetición consecutiva de la ubicación del mismo locutor y de los locutores vecinos. Finalmente, se confeccionan al menos tres series de respuestas, donde cada una de ellas contiene diferentes *cadenas problema* para así ofrecer múltiples opciones de decisión a la víctima o testigo.

Ejemplo de hoja de respuestas a los estímulos presentados:

HOJA DE RESPUESTAS RUEDA DE RECONOCIMIENTO DE VOZ

Serie número: 1

Muestras frase número 1:



Imaginemos un caso en el que el sujeto agresor profiriese cinco frases. La víctima o testigo habría de efectuar quince tomas de decisión.

Todas las emisiones de habla grabadas utilizadas en la prueba se remiten a la Autoridad Judicial competente junto con el informe de resultados de la prueba.

La rueda de reconocimiento de voz se efectúa de acuerdo al procedimiento legal que la Autoridad Judicial estime oportuno teniendo en consideración las particularidades de cada caso. En este sentido, los miembros actuantes del laboratorio asesoran al juez o su representante sobre las condiciones más idóneas para llevar a cabo la prueba, tanto en lo concerniente al ambiente acústico, concentración del sujeto objeto de la prueba, etc.

Una vez conseguidas las circunstancias más favorables se procede a la práctica de la prueba, la cual, enfrentará a la víctima o testigo a diferentes «cadenas problema» debiendo señalar aquellos estímulos que considere atribuibles al sospechoso/a.

2.2 La tormenta perfecta: **invidentes/percepción auditiva**

En 2019 un Juzgado de Instrucción de Córdoba solicitó la colaboración del laboratorio de Acústica de la Policía Científica para la realización de una rueda de reconocimiento de voz.

Se trataba de un hecho delictivo en el que miembros invidentes de la ONCE habían sido atracados por un mismo sospechoso en sus puestos de

trabajo, en diferentes localizaciones de Córdoba capital. Llevadas a cabo las necesarias entrevistas con diferentes víctimas, finalmente se decide realizar la prueba con dos de ellas. Las víctimas narraron los hechos delictivos acaecidos detallando las frases emitidas por el sospechoso. Como resultado de ello se pudieron seleccionar cinco frases diferentes por cada una de las víctimas.

Se elaboraron las pertinentes cadenas de habla con cinco *voces contraste* más la voz del investigado. La prueba deparó los siguientes resultados:

- A la víctima n.º 1 se le plantearon diez tomas de decisión. Respondió a todas ellas y reconoció la voz del investigado en las diez ocasiones.
- La víctima n.º 2 se enfrentó también a diez tomas de decisión. Al igual que la víctima n.º 1 respondió a todas ellas y reconoció la voz del investigado en nueve de las diez ocasiones.

A priori, el número de tomas de decisión previstas era de quince, pero, en un momento dado, la juez actuante consideró oportuno no continuar con series adicionales dada la tremenda contundencia y seguridad con la que las víctimas llevaban a cabo sus decisiones.

3. El valor probatorio de las RRV

Sin lugar a dudas, la experiencia del Laboratorio de Acústica Forense de la Policía Científica en relación con las RRV ha de calificarse como muy positiva. A diferencia de lo que ocurre con otras técnicas practicadas en el Laboratorio (Análisis Comparativo de Habla, Pasaporte Vocal, etc.) en las Ruedas de Reconocimiento de Voz el factor subjetividad que caracteriza la discriminación perceptiva, condiciona claramente la apreciación de la prueba de cara a los Tribunales de Justicia.

Los resultados aportados por la víctima o testigo no son interpretados por los científicos forenses en lo concerniente a una posible identificación o descarte de autoría. Es la autoridad judicial quien, en última instancia, evalúa las tomas de decisión de la prueba, si bien, puede solicitar aclaraciones al respecto por parte de los científicos.

En términos generales, una RRV es apreciada como un indicio adicional de investigación pues existen demasiadas variables a controlar. Algunas de ellas vinculadas con el sujeto objeto de la prueba (habilidades auditivas, per-

ceptivas o asociadas a la memoria, estrés, etc.) y otras, relacionadas con diversos factores extrínsecos (ambiente acústico de los escenarios, adecuación de voces contraste, etc.)

En definitiva, nos encontramos ante una herramienta de incuestionable valor, en la que la visión del especialista forense juega una baza trascendental a la hora de velar por que su utilización sea ejecutada con la debida pericia y sus resultados sean contemplados con la necesaria cautela. De la misma forma, resulta determinante que los expertos de nuestro ámbito -al margen del debido auxilio a la autoridad judicial en cualquiera de sus requerimientos- ejerzan un papel protagonista en las diferentes fases del proceso: valoración de los escenarios, tasación del nivel de competencia de las víctimas o testigos, análisis fonético lingüístico de las emisiones de habla interesadas en la prueba, etc.

Rigor, experiencia y objetividad. Estas son las coordenadas que, como ocurre en cualquier otra técnica de investigación científico-forense, enmarcan y definen la liturgia del reconocimiento perceptivo mediante rueda de voz. Una especie de antídoto para una técnica pericial en la que el factor subjetividad se presenta como principal desafío de adversidad.

Bibliografía

- Broeders, A. P. A. y van Amelsvoort, A.G. (1999):** «Lineup construction for forensic ear-witness identification: a practical approach», en John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville y Ashlee C. Bailey (ed.), *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco, pp. 1373-1376, <http://www.internationalphoneticassociation.org/icphs/icphs1999>.
- Hollien, Harry (2012):** «On earwitness lineups», *Investigative Sciences Journal*, 4, 1, www.investigativesciencesjournal.org.
- Hollien, Harry; Huntley, Ruth; Künzel, Hermann; y Hollien, Patricia A. (1995):** «Criteria for earwitness lineups», *The International Journal of Speech, Language and the Law*, 2, 2, pp. 143-153.
- Künzel, Hermann (1994):** «On the problem of speaker identification by victims and witnesses», *The International Journal of Speech, Language and the Law*, 1, 1, pp. 45-57, <https://doi.org/10.1558/ijssl.viii.45>.
- Marrero Aguiar, Victoria (2001).** *Fonética perceptiva*. Addenda, material complementario de la asignatura Fonética y Fonología de la Lengua Española, Madrid, Departamento de Lengua Española y Lingüística General de la Facultad de Filología (UNED).
- Smith, Harriet Mary Jessica, y Baguley, Thom (2014):** «Unfamiliar voice identification: Effect of post-event information on accuracy and voice ratings», *Journal of European Psychology Students*, 5, 1, pp. 59-68, <http://dx.doi.org/10.5334/jeps.bs>.