

## Reconocimiento de voz en rueda: una vuelta de tuerca

Speech Recognition Parade:  
One Step Further

**ELENA GARAYZÁBAL HEINZE**  
*Universidad Autónoma de Madrid*  
elena.garayzabal@uam.es  
<https://orcid.org/0000-0001-7534-9141>

**MERCEDES REIGOSA RIVEIROS**  
*Cuerpo Nacional de Policía*  
mercereigosa@gmail.com

**Resumen:** La complejidad de poner en marcha una rueda de reconocimiento de voz durante un proceso judicial es indudable, teniendo en cuenta que es una prueba perceptiva que parte de la recuperación de un recuerdo auditivo de la víctima o testigo de un hecho delictivo grave, y que debe realizarse tiempo después de que este se haya producido atendiendo a criterios referidos a la cualidad tímbrica de la voz del agresor. Actualmente, los estudios realizados en los distintos campos implicados en esta técnica forense señalan su rigurosidad. Sin embargo, para que esa evidencia pueda ser considerada lo bastante robusta, debe producirse un engranaje perfecto entre las distintas piezas que componen los actores intervinientes: el eje formado por víctima y sospechoso, y el equipo de expertos, y el formado por las autoridades judiciales, pues será el juez quien tome la decisión final.

**Palabras clave:** Identificación auditiva, prueba perceptiva, patrón-timbre, adecuación perceptiva, distractores.

**Abstract:** The complexity of starting a speech recognition parade during a judicial process is unquestionable, considering that it is a perceptual examination starting from the recovery of an auditory memory by the victim or witness of a serious criminal act, and it must be carried out some time after it has occurred, considering criteria referred to the timbre quality of the aggressor's voice. Currently, the research in the different fields involved in this forensic technique points to a rigorous approach. But for this proof to be considered solid enough, there must be a virtuous circle of different actors: the axis victim-suspect and the team of experts, and the one formed by the judicial authorities, since the judge will make the final decision.

**Keywords:** Earwitness identification, perceptual evidence, timbre-pattern, adaptation of speech perception, foil speakers.

*La duda es un estado incómodo,  
pero la certeza es un estado ridículo.*

VOLTAIRE

## 1. Introducción

Hablar de la reciente disciplina de la lingüística forense y sus novedosas técnicas dentro de la lingüística aplicada resulta cuando menos llamativo, porque claramente ni la disciplina es reciente ni sus métodos novedosos. Lo cierto es que los lingüistas forenses llevan ya muchos años poniendo su conocimiento especializado al servicio de la justicia en numerosos países en sus tres áreas de aplicación que, en líneas generales, se podrían condensar en tres según la clasificación realizada por Gibbons y Turell en 2008 (*apud* Garayzábal Heinze, Queralt Estévez y Reigosa Riveiros, 2019): el lenguaje de la ley, el lenguaje judicial y el lenguaje probatorio o evidencial. Es en este último donde se encuadra la técnica identificativa conocida como reconocimiento de voz en rueda o, lo que es lo mismo, rueda de reconocimiento de voz.

Si sencillamente se hace una revisión de la amplia terminología empleada para denominar dicha técnica en lengua inglesa, dependiendo de dónde y quién la refiera, ya da a entender la falta de homogeneidad de criterios respecto a la misma y de valoraciones dispares que esta técnica se arroga: *speech or voice recognition parade*, *earwitness or voice lineup*, *voice identification parade*, *speech identification or recognition in lineup*. Lo que es innegable es que sigue suscitando interés entre los expertos de los distintos campos de conocimiento que confluyen en ella, ya que continúan realizándose investigaciones y publicaciones que contribuyen a reforzar el uso de la misma en contextos tan particulares como el judicial y el policial, eso sí, siempre y cuando se den una serie de condiciones que minimicen al máximo la posibilidad de que se produzca un error en la identificación de un no culpable, además de cumplir con la legalidad.

¿Por qué entonces ese rechazo a la utilización de técnicas forenses en las que el lenguaje en sus distintas formas es parte del conjunto de pruebas en un procedimiento judicial? Igual que el estudio de la huella de una zapatilla deportiva en la escena del crimen, el análisis de la lengua, tanto en su

manifestación oral como escrita (y como herramienta básica en el proceso comunicativo de la interacción humana), aporta información relevante cuando el autor de un delito deja su rastro lingüístico durante su comisión. Es posible argumentar un sinfín de razones por las que no se recurre a este tipo de técnicas identificativas englobadas en la lingüística forense, pero la verdad es tan simple como reconocer que sigue existiendo una absoluta falta de conocimiento y formación, y quizá de interés, en estas técnicas en general, y más concretamente en la identificación de voz en rueda, por parte de algunos de los sectores implicados. Todo ello contribuye a limitar la disponibilidad de expertos competentes en su elaboración, y en el desarrollo de una reglamentación específica sobre su uso en nuestro sistema judicial, donde se equipara sin criterio alguno a la rueda visual. El reconocimiento de voz en rueda es una técnica forense más con valor de prueba que puede apoyar o rebatir la tesis del ministerio fiscal. Indudablemente, por sí sola no representa un valor absoluto, ya que se trata de una prueba auditivo-perceptiva donde se exige a la víctima o testigo de un hecho delictivo que ponga en marcha no solo su buen oído, sino una memoria auditiva y una cierta templanza para recordar lo acontecido en un momento de máximo estrés. No obstante, los expertos reconocen que la puesta en marcha de una rueda de voz bien estructurada y controlada por parte de personal cualificado puede resultar ciertamente eficaz.

Es evidente que los reconocimientos *in situ* afirmando «esa es su voz» meses, incluso años después, no ofrecen ninguna seguridad jurídica, pero las investigaciones realizadas en este campo hasta la fecha, así como los últimos casos en los que se ha recurrido a la realización de una identificación de voz en rueda, están demostrando que, con rigurosidad en el cumplimiento de las estrategias reguladas para cada fase del proceso y su adaptación en cada caso específico, pueden ser garantía suficiente para poder sacarle el mayor partido cuando se den las circunstancias para ello. Son muchas las partes implicadas en la realización de esta técnica, no solo víctima o testigo y sospechoso, y los expertos que implementan la técnica, sino también las autoridades judiciales, quienes requerirán la realización de la misma, y que no solo estarán informadas del seguimiento en sus distintas fases, sino también presentes en alguna de ellas, pero, sobre todo llegado el momento de la vista oral, será el juez quien evalúe el resultado y quien finalmente tome una decisión sobre su valor como prueba.

Mucho se ha escrito sobre la conveniencia o no de aplicar esta técnica de identificación como parte de un procedimiento judicial, la mayoría de las ve-

ces para echar por tierra su utilización con argumentos de lo más variopinto: la memoria de los testigos no es de fiar, los investigadores orientan al testigo, un par de palabras no es suficiente, no existe familiaridad con la voz o con el idioma, etc. Quizá sea el momento de conocer un poco más en detalle en qué consiste esta técnica, haciendo hincapié en que su realización la lleven a cabo expertos bien entrenados que aporten confianza en su desarrollo, ya que deberán ser precisos en el seguimiento de las distintas fases y saber cuándo y por qué puede resultar o no adecuada en determinados casos. Será parte de la labor de estos expertos intentar minimizar sus riesgos mediante el control de todas las variables que intervienen en el proceso, y conociendo que no siempre tiene que ser una prueba incriminatoria sino también exculpatoria. Es igualmente importante saber cuál es el papel que cada uno de sus participantes tiene en las distintas partes del proceso, porque cada caso, cada testigo y cada circunstancia particulariza cada una de las fases que componen esta técnica identificativa.

Aunque de forma breve, se abordará no solo lo concerniente a la descripción de esta técnica, cuáles son los inconvenientes que presenta y cuáles las posibles ventajas, también se mencionarán algunas de las investigaciones que se siguen realizando y su consideración dentro de distintos sistemas judiciales. Asimismo, se podrá comprobar que la precisa estructuración en la metodología aplicada a la hora de llevar a cabo las ruedas de identificación auditiva dentro de un procedimiento judicial puede asegurar un evidente valor probatorio de la misma; y tal vez sea el momento de dar un paso adelante y plantear su uso en el procedimiento judicial de forma más rutinaria.

## 2. Fundamentos teóricos

La utilización de diferentes técnicas destinadas a la identificación de los autores de hechos delictivos ha sido una constante en el proceso de una investigación policial y, por ende, a lo largo de un procedimiento judicial. La complejidad de estas ha llevado a la necesidad de establecer una serie de directrices que garanticen un proceso riguroso y una conclusión con el mayor grado de fiabilidad, minimizando los riesgos de un resultado incorrecto o *falso positivo*, es decir, la posible identificación de un no culpable.

## 2.1 En qué consiste un reconocimiento de voz en rueda

Una de esas técnicas es el conocido reconocimiento en rueda que, aunque mayoritariamente se conoce en su aplicación visual, tanto a nivel legislativo (art. 369 de la LECrim) como jurisprudencial y práctico, tiene también una realización quizá menos conocida y mucho más compleja en su elaboración y puesta en marcha, las denominadas ‘ruedas de identificación de voz’, un reconocimiento de voz en rueda usado en el ámbito policial y judicial, cuya finalidad es la posible identificación del autor de un delito por parte de un testigo que manifiesta ser capaz de reconocer la voz de su agresor a pesar de no haber podido visualizar sus rasgos faciales; es decir, las pistas verbales se convierten en una de las claves en la identificación de los autores de delitos que revisten cierta gravedad, como es el caso de los robos con violencia e intimidación o las agresiones sexuales en las que el autor cubre su rostro o actúa en entornos con escasa iluminación.

Cuando tiene lugar un hecho delictivo de estas características, se hace necesario que el testigo recupere de su memoria la voz del sospechoso en distintas partes del proceso de elaboración de la rueda de voz, y en su última fase deberán distinguir esa voz de entre todas las que se le presentan en unas cadenas habladas que se editan a tal fin. Y aquí parece encontrarse un primer escollo. El reconocimiento de voz en rueda implica la ejecución de dos funciones y procesos cognitivos diferentes: discriminar una voz y reconocerla de entre varias ofrecidas (Kreiman y Papcun, 1991); es decir, supone reconocer una voz e identificar al hablante, por lo que se trata de una tarea que consiste básicamente en la emisión de juicios por parte del testigo, que, a la sazón, no es un experto en la tarea de reconocer voces, de ahí que su testimonio se considere poco fiable; pero ¿realmente lo es?

A través de la voz se obtiene información del habla, por medio de rasgos inherentes y constantes del locutor, como el timbre, el sexo o la edad, pero también información que es cambiante y que, en buena medida, tiene que ver con la variedad dialectal que utiliza, su procedencia geográfica o su pertenencia a un grupo social concreto (Ladefoged, 1975; Laver y Trudgill, 1979; Belin, Fecteau y Bédard, 2004). La variabilidad puede proporcionar información relevante que afecta a la fluidez de la codificación y a la eficiencia del procesamiento perceptivo. Los hallazgos de Pisoni (1993) evidenciaron que las categorías del habla son altamente sensibles al contexto y que algunos detalles de la entrada de la señal no se pierden ni se filtran como consecuencia del análisis perceptivo. Esto es consistente con la idea de la implicación

de sistemas de memoria múltiple y del papel de los sistemas de representación perceptual en la memoria, que permiten dar cuenta de la variabilidad y los efectos del contexto en la percepción del habla. Estudios más recientes constatan que, independientemente del contenido, nuestro cerebro es capaz de identificar aspectos invariables de la voz para su reconocimiento (Zäske, Hasan y Belin, 2017). En muchas ocasiones, los testigos son capaces de reconocer el tono, la cadencia, la velocidad, la intensidad o la duración del habla, incluso el estado emocional de quien efectúa las emisiones, como, por ejemplo, el enfado, la urgencia o el miedo (Doddington, 1985). Es importante recordar aquí que no existen dos voces absolutamente iguales, ni tan siquiera entre gemelos, dado que tanto la anatomía como la fisiología difieren de un hablante a otro y, en consecuencia, también la acústica del habla (Nolan y Oh, 1996).

Una prueba de identificación es producto de un proceso cognitivo complejo de por sí, pero lo es especialmente en aquellos casos en los que un testigo oye, pero no ve al sospechoso. El valor de una identificación de voz en rueda va a estar directamente relacionado con la meticulosidad y precisión con que se presente la tarea, para la que, en general, los expertos defienden el uso de una construcción múltiple. Reproducir la grabación de un sospechoso y realizar al testigo una pregunta cerrada del tipo «¿es esta la voz que oyó?» no ofrece demasiadas garantías a la hora de evitar falsas identificaciones. Con el fin de eludir posibles sesgos, como que el testigo responda afirmativamente en un intento de ser cooperativo o que pretenda hacer una identificación cuando presenta ciertas dudas, la colocación del sospechoso dentro de un grupo de voces presentadas sucesivamente permite, probabilísticamente al menos, la protección del posible inocente si la rueda auditiva de selección múltiple está adecuadamente elaborada (Nolan, 2001).

## ***2.2 Investigación sobre el reconocimiento de voz en rueda y su problemática***

La principal dificultad en la realización de esta técnica de identificación viene dada por la cantidad de variables que intervienen en el proceso de reconocimiento perceptivo, que tienen que ver con la habilidad y la memoria auditiva de la víctima, su edad, las circunstancias que rodearon la emisión del discurso y el tiempo de exposición al mismo durante la comisión del delito, la familiaridad o no con la voz, el tiempo transcurrido desde la emisión, etc.

(Garayzábal Heinze, Queralt Estévez y Reigosa Riveiros, 2019). Todos esos factores contribuyen a mantener las dudas sobre su fiabilidad y conducen a una falta de consenso entre los profesionales que trabajan, interpretan y sentencian sobre la realización de una rueda de reconocimiento de voz.

En 1997 Laubstein planteaba los problemas existentes respecto al reconocimiento con segundas lenguas o en relación con grupos raciales que parecen identificar *mejor* a los de su propio grupo. Respecto al número de componentes de la rueda auditiva, algunos expertos estimaron que a mayor número de distractores, mayor el grado de credibilidad del testigo, llegando a proponer una rueda con 20 hablantes distintos. Yarmey (1995), por el contrario, lo consideró impracticable y no necesariamente efectivo, y propuso un intervalo comprendido entre 6 y 12 voces distintas.

En España se vienen realizando trabajos de investigación cada vez más interesantes sobre este tipo de rueda de identificación, que refutan o apoyan las hipótesis de estudios anteriores. Uno de estos es el trabajo realizado por Valiente (2011) sobre cómo el factor ‘número de distractores’ influye en el reconocimiento de voces en rueda. En su trabajo parte de la premisa de que a mayor número de distractores menor precisión en los aciertos, de lo que se deduce que un menor número de distractores facilita la tarea de identificación; sin embargo, el resultado obtenido tras los experimentos fue justo el contrario: aquellos expuestos a un mayor número de distractores «excluyen de forma más resuelta las muestras» al aumentar su posibilidad de comparación; mientras que el otro grupo que se enfrentaba a un menor número de distractores parecían dudar más a la hora de identificar la voz diana.

### 2.2.1 Memoria auditiva–prueba perceptiva

Varios son los problemas que se aducen para que se otorgue la oportuna confianza a este tipo de procedimiento; la mayoría de ellos se encauzan en la figura de los testificantes y la fiabilidad de su respuesta en relación con la mayor o menor familiaridad de la voz, la capacidad de reconocimiento auditivo, la memoria para la voz y el periodo de latencia. En este sentido, los estudios tampoco son claros. Ya en 1980, Saslove y Yarmey sugirieron que la identificación de hablantes a largo plazo debía ser considerada con precaución por el sistema de justicia penal. Clifford (1980: 377) también señala al respecto de la buena memoria para la voz en situaciones forenses que:

[...] depends upon so many factors, some of which have already been alluded to briefly -the time between initially hearing a voice and eventually attempting to recognize it; the relative accuracy of identifying a voice heard live or from a tape; the number of foils employed at recognition; the relative accuracy of males and females, or one individual as opposed to another; the differential identifiability of male and female voices and of one voice as opposed to another. Of critical importance also would seem to be the duration or quality of the speech sample heard. In trying to establish base-line performance it would also seem that the age of the witness must be taken into account.

Con posterioridad, Read y Craik (1995) señalan como variables que afectan a la habilidad de reconocer una voz tras una exposición relativamente breve la preparación del testigo para la prueba de memoria o la duración de la muestra de voces que se le van a presentar, entre otras. En un estudio sobre la precisión en la identificación y valoración de la voz en una prueba incidental de memoria de voces desconocidas, Smith y Baguley (2014) recuerdan que el factor tiempo puede ser motivo de distorsión del recuerdo de la voz.

Es en este momento donde es necesario introducir el concepto de *psicoacústica*, que ayuda a explicar el proceso de identificación de voces, ya que se centra en el estudio de la audición a través de las respuestas subjetivas a los estímulos acústicos, especialmente en tareas de detección y discriminación (Marrero, 2001). Lo importante es descubrir cómo la señal percibida permite recuperar la señal emitida, y para ello se distinguen tres fases en el proceso: (i) detección, que conlleva notar la presencia o ausencia de un estímulo en función de su duración, intensidad y frecuencia, sin llegar a identificarlo; (ii) discriminación, es decir, establecer una comparación para buscar diferencias entre estímulos próximos; (iii) identificación, que requiere relacionar el estímulo que se presenta con una representación existente en la memoria.

### 2.2.2 *Recreación controlada de ruedas identificativas de voz*

El estudio empírico de la identificación de voces sigue siendo escaso y complicado de llevar a cabo; el principal motivo es que no se puede realizar en situaciones reales, por lo que se intenta reproducir en situaciones de experimentación con voluntarios. El objetivo de la mayor parte de estas investigaciones (Yarmey, 1991; Pisoni, 1993; Orchard y Yarmey, 1995) se centra en cuestiones tales como: conocer la fiabilidad en el ejercicio del reconocimien-



to de voces, conocidas y desconocidas, dentro de una cadena donde hay que identificar la voz diana; testear la memoria para voces, o lo que es lo mismo, la precisión en el reconocimiento de una voz ya escuchada y presentada después de un tiempo; determinar a qué características de la voz se presta más atención; o concretar cuánto tiempo de exposición a la voz es necesario para poder memorizarla y recuperarla con posterioridad, etc.

Estos son solo algunos de los enfoques planteados en los estudios realizados en este campo, que, aunque interesantes, distan ciertamente de la realidad, por lo que están lejos de representar una certeza conclusiva. Lo que se pretende en cada estudio es controlar todas las circunstancias que rodean a los voluntarios participantes, en algunos casos en número reducido, en el proceso de escucha y discriminación de las muestras de voz; motivo por el que los resultados podrían considerarse orientativos no preceptivos. La situación de estrés vivida por la víctima o el testigo de un acto delictivo grave presenta una amplia casuística, por lo que hay que tener en cuenta no solo cómo este les afecta sino en qué condiciones se encuentran para afrontar un complejo proceso de identificación cuando la voz constituye una evidencia inculpatoria. La recreación de un entorno situacional de este tipo parece estar fuera de lugar, por lo que las condiciones personales de los participantes nunca se ajustarán a las que presente víctima o testigo en una rueda real.

Otra cuestión importante es que en algunos de los estudios realizados (Sarwar, Allwood, Martin y Zetterholm, 2014; Bartle y Dellwo, 2015; Perrachione, Furbeck y Thurston, 2019) las muestras se hacen sobre la lectura de un texto balanceado o un texto narrativo, lo que se aleja de la espontaneidad del habla y, por ende, del ámbito delictivo, donde tampoco se suele producir un discurso extenso, sino más bien breve y con locuciones muy precisas.

A pesar de que gran parte de la bibliografía sobre identificación auditivo-perceptiva de hablantes no se centra exclusivamente en las ruedas de reconocimiento de voz, las investigaciones en este campo han ido avanzando a buen ritmo desde que los sistemas judiciales de distintos países pusieran en práctica las primeras ruedas de identificación auditiva múltiple para la identificación de un delincuente por su voz, una vez que la víctima manifestara la posibilidad de poder reconocerla. Sin embargo, continúan planteándose dudas respecto a su aplicación, sobre todo por la falta de confianza expresada respecto a su incierto valor probatorio ante un tribunal, aunque dentro del contexto judicial la finalidad del proceso de identificación no es únicamente la de identificar al autor de un delito, también prevenir sentencias inculpatorias a personas inocentes. Conviene recordar que la rueda de reconocimiento

de voz resultará eficaz siempre y cuando se dirija de forma adecuada, mitigando su fragilidad mediante la organización de un método preciso y bien planificado. Ya en 1995, Hollien y sus colaboradores manifestaban que para dar fiabilidad y validez al reconocimiento de voz en rueda había que mejorar los procedimientos y establecer rigurosos criterios de uso. La creación de estándares y procedimientos respecto a su diseño y posterior desarrollo aportarían mayor consistencia.

### 2.2.3 Rueda visual vs. rueda auditiva

Inicialmente, los expertos argumentaban el uso de la rueda de voz siguiendo el modelo de la rueda visual (Broeders y Rietveld, 1995; Yarmey, 1995) dado que en ambas se ponen en juego procesos similares: modalidades sensoriales, memoria, situaciones de confrontación, la selección de una persona entre varias, etc. No obstante, es obvio que ambas situaciones presentan diferencias en cuanto a cómo se procesa la memoria auditiva y la visual, cómo se estructuran ojos y voces, y cómo el miedo o el enfado afectan a ambos procesos de identificación. Las habilidades individuales permiten a ciertas personas tener una mejor memoria visual, mientras otras presentan una mayor capacidad auditiva. Estas y otras diferencias muestran que las identificaciones visuales y auditivas no son equivalentes (Künzel, Masthoff y Köster, 1995; Hollien, 2002). Por ello es importante centrarse en los fundamentos del proceso de reconocimiento auditivo y cómo puede estructurarse. Para Ballesteros (1997), la diferencia entre la percepción visual y la auditiva es que esta se extiende en el tiempo, es decir, tiene un carácter sucesivo y no simultáneo, lo que permite que la duración de los estímulos auditivos almacenados sea mayor que la de los estímulos visuales; por lo tanto, parece cuestionable *a priori* que el reconocimiento visual y auditivo sean comparables, y menos en situaciones forenses.

Una dificultad añadida es que las investigaciones centradas en la percepción visual abundan en comparación con las dedicadas a la percepción auditiva; quizá se deba a que en la práctica forense se realiza un mayor número de ruedas de reconocimiento visual que auditivo y se ha hecho una extrapolación simplificada del procedimiento de las ruedas de reconocimiento visual a las de reconocimiento auditivo, cuando las disimilitudes son considerables.

Una cuestión relevante en la elaboración de las ruedas es la elección de los distractores que acompañarán al sospechoso. Los criterios de selección

varían según la técnica de identificación elegida, como se podrá comprobar más adelante:

[...] while the visual line-up can capitalise on the surplus of information in the mental trace of the culprit in the witness's memory, over and above that contained in the verbal description of the culprit's voice by the witness, the voice line-up necessarily has to settle for a high degree of suspect similarity (Broeders y van Amelsvoort 2001: 241)

### ***2.3 Dificultades a las que hacer frente en la construcción de una rueda de voz***

En situaciones de normalidad, los estudios apuntan a que somos capaces de recordar aspectos concretos de la fuente de voz de un hablante, dado que estas características se almacenan a largo plazo (Pisoni, 1993). Yarmey (2007) observa que la tasa de error en el reconocimiento de voces desconocidas es elevada y que los errores superan el 50%, algo que no parece ocurrir con voces familiares, aunque es también asumible que pueden producirse errores de identificación. Cook y Wilding (1997) muestran que en el reconocimiento de una voz escuchada una única vez la duración del enunciado tiene efectos de mayor trascendencia que la variación fonemática, tanto en la memoria a corto como a largo plazo. También Legge y sus colaboradores (1984) encontraron que la longitud de la emisión era decisiva en la identificación de hablantes. En línea con estos dos últimos estudios, Kerstholt y colaboradores (2004) observaron que la fiabilidad en el reconocimiento de voces desconocidas, incluso en ambientes ruidosos, era mejor cuanto más larga fuera la muestra de habla.

Son muchos y variados los factores que pueden afectar a la creación de una rueda de reconocimiento de voz, entre ellos estarían el tipo y las circunstancias del delito, las condiciones físicas y psicológicas de la víctima, las características del intercambio comunicativo, la familiaridad o no de la voz, el período transcurrido entre el hecho delictivo y el momento de realizar la identificación e incluso el lapso temporal de exposición a la voz del agresor por parte de víctima o testigo, que según Broeders y van Amelsvoort (1999) debe ser algo más de 30 segundos. Dada la cantidad de variables que intervienen en el procedimiento, es evidente que no podrán ser controladas en su

totalidad, pero sí atenuadas. La complejidad de su elaboración debe ser supervisada por personal especializado, con formación específica en lingüística y psicología (Künzel, 1994), lo que aportará un mayor grado de fiabilidad a la rueda auditiva. Su viabilidad estará determinada por el rigor en el seguimiento de los adecuados procedimientos en su construcción y su administración, que deben estar bien estructurados y ser lo suficientemente precisos para asegurar que sea justa para las partes.

### 3. Cómo recoge la jurisprudencia el reconocimiento de voz en rueda

Desde que Lindbergh identificara por su forma de hablar, y especialmente por su fuerte acento alemán, al secuestrador de su hijo casi dos años y medio después de oír brevemente su voz durante la entrega del dinero del rescate en un cementerio, ha pasado mucho tiempo y han cambiado muchas cosas, pero algunas relacionadas con la actuación de los tribunales respecto al uso de esta técnica de identificación sin seguir criterio técnico alguno, parecen no haber variado. En el caso Lindbergh el tiempo de exposición a la voz desconocida fue de unos 2 o 3 segundos. La identificación de la voz se produjo 29 meses después en dos ocasiones, primero en la fiscalía del distrito donde el aviador se ocultaba bajo un disfraz y luego en la propia sala del tribunal, y fue lo que determinó la inculpación de Hauptmann; no se tuvo en cuenta la base articulatoria del acusado al pronunciar en lengua inglesa; en ningún momento se optó por realizar una rueda de elección múltiple con hablantes similares, ni se contó con el testimonio de expertos en la identificación de hablantes (Solan y Tiersma, 2003).

Casi noventa años después, y todavía sin estándares de buenas prácticas para la aplicación de esta técnica, los tribunales tienden a actuar erróneamente justo en el sentido contrario al cuestionar de forma sistemática que una víctima sea capaz de recordar la voz de su agresor. Si se revisan algunas de las sentencias de los tribunales españoles de los últimos años donde se menciona como parte del procedimiento judicial un reconocimiento de la voz del sospechoso por parte de una víctima o testigo, parece que se repitan como en una especie de mantra antiguas sentencias en las que se aceptan o rechazan las pruebas orales con las mismas argumentaciones, dejando de lado algo tan relevante como el avance del conocimiento, las recientes investigaciones realizadas, la mejora tecnológica y la rigurosidad en la aplicación

de las técnicas actuales en lo referente a los análisis de los expertos en las pruebas periciales.

### 3.1 *El reconocimiento de voz en rueda en España*

El sistema judicial español ha recurrido a este tipo de diligencia probatoria en contadas ocasiones, y una gran parte de ellas se han producido en circunstancias ciertamente poco rigurosas, es decir, sin locutores de referencia, con escaso discurso o sin identificación presencial, etc. según se recoge en algunas sentencias. Si se repasa la jurisprudencia más reciente, se observa que la reseña de sentencias anteriores es un procedimiento jurídico muy recurrente, lo que en sí no es cuestionable, pero si se hace un examen más minucioso respecto a aquellos procesos en los que se requiere la realización de una práctica identificativa de la voz de un investigado, se reiteran argumentaciones que, analizadas con el conocimiento actual sobre el tema, resultan un tanto inadecuadas. Sirva como ejemplo la STSJ LR 12/2020 donde se puede leer en diferentes ocasiones que «no se alega la concurrencia de causa alguna de incredibilidad subjetiva respecto a la identificación de voces por parte de uno de los testigos», un agente de la Benemérita quien tras visionar la grabación «reconoció sin ningún género de dudas a ambos acusados, por la voz»; o que la identificación de uno de los investigados fue realizada por parte del propio Tribunal «tras escucharle personalmente en el acto plenario». Todo ello sin que conste en ningún momento opinión o constatación realizada por parte de un experto acerca de en qué consiste la similitud entre las voces.

Mucho más recientemente, en la STSJ EXT 289/2022, se sigue reconociendo que «la identificación de la voz como la de los acusados puede ser apreciada por el Tribunal en virtud de su propia y personal percepción y por la evaluación ponderada de las circunstancias concurrentes». Tampoco la credibilidad de los policías a cargo de las comunicaciones intervenidas durante la investigación ofrece dudas, «dado el acostumbramiento a su audición, acento o entonación». En la misma se hace referencia a otra sentencia de hace más de 20 años (STS 2384/2001) en la que se recoge que:

[...] la intermediación no es solo estar presente, sino entender, percibir, asimilar, verificar, en definitiva formar opinión en conciencia y en el conjunto sobre todo lo dicho, notando las reacciones y gestos de todos, singularmente de los inculcados, por ello, lo que se critica supone

precisamente la manifestación más propia de la intermediación judicial como es verificar que la voz escuchada en una cinta, coincide con lo escuchado directamente de una persona en el Plenario, y concluir con la afirmación de pertenecer a la misma persona.

Ante esto, quizá haya llegado el momento de plantearse si son oportunas exposiciones tan taxativas a la hora de establecer conclusiones respecto a cuestiones tan altamente especializadas. Afirmar con rotundidad que dos emisiones de habla corresponden a la misma persona solo con la apreciación perceptiva de un lego en lo que respecta al conocimiento de los rasgos que caracterizan una voz y demás circunstancias que rodean las emisiones orales es, incluso para un juez, un tanto osado, y eso sin entrar en valoraciones legales al respecto. Tal vez convendría repasar la Ley de Enjuiciamiento Civil 1/2000 y recordar que en su artículo 335.1 expresa la posibilidad de solicitar la ayuda de expertos con conocimientos especializados «para valorar hechos o circunstancias relevantes en el asunto o adquirir certeza sobre ellos».

Como muestra de esa inadecuación temporal, en esa misma sentencia se puede observar el uso del término ‘cinta’ hasta en nueve ocasiones, para referirse a las grabaciones que forman parte de las evidencias, casi la mitad de las cuales se refieren a los dispositivos de almacenamiento de registros de habla en formatos digitales. Esto da una idea del desajuste no solo terminológico sino conceptual, teniendo en cuenta el avance de la tecnología, en general, y de las técnicas de identificación forense, en particular.

Es cierto que en otros casos (SAP M 8008/2021) las referencias sí están más actualizadas, pero se sigue manteniendo la misma idea: «Y es que los mensajes en cuestión, resultan ser audios y se ha tenido la oportunidad de escucharlos en el juicio, pudiendo concluirse que la voz del ahora acusado coincide con la de los mensajes en cuestión». En este caso, el Tribunal concluye, tras una escucha perceptiva en la sala y sin aportar criterio lingüístico alguno, que la voz del acusado es la misma que emitió los mensajes. Cabría preguntarse entonces cuál puede ser el motivo de que a nivel jurisprudencial se sigan obviando los avances en el desarrollo de ciertas técnicas forenses. La falta de conocimiento o de interés no puede justificar el empleo de argumentaciones anacrónicas que no contribuyen en nada a la modernización del sistema judicial.

Entre los distintos motivos que se alegan en la STSJ CAT 3835/2021 para revocar parcialmente una sentencia previa es que la víctima de una agresión sexual se desdijo de su reconocimiento *in situ*, tal como consta en la

sentencia, de la voz de uno de los investigados, cuyos rasgos identificativos «pueden concordar con un amplio sector de la población de origen magrebí de esa edad». En la SAP M 6239/2019, donde concurren varios delitos (robo con violencia e intimidación, lesiones y detención ilegal), se lee que la perjudicada reconoce la voz del agresor e identifica al hombre que había estado en su casa arreglando la antena, y lo vuelve a identificar en una rueda de voz que se practicó judicialmente, en la que se le dio a elegir entre varias voces indubitadas. Sin embargo, en la sentencia, el Tribunal observa tanto limitaciones en la identificación de la víctima como deficiencias en el desarrollo y la documentación de la rueda. Se pone en tela de juicio el testimonio de la víctima, señalando, por un lado, su falta de familiaridad con la variedad del español hablado por personas de origen marroquí; y por otro, la disminución de la memoria auditiva por el lapso transcurrido entre el primer contacto y la rueda de voz. A esto se añade, además, que, a pesar de todo lo que había dicho el autor durante la comisión del atraco, la víctima dijo reconocerlo por «el modo en que decía el nombre de su jefe: Isidoro». También se aducen consideraciones en relación con la metodología empleada, como que la rueda de voces no ha sido grabada ni el Letrado de la Administración de Justicia ha diligenciado:

[...] las semejanzas o diferencias entre los tonos de voz escuchados por la testigo, tampoco modo de articular, dialecto, velocidad del habla u otras particularidades tanto de la voz como de los integrantes de la rueda, tales como su edad, por lo que no se puede afirmar que las voces fueran semejantes, lo que es exigible por igualdad de criterio por lo establecido en el art. 369 de la LECrim relativo a las ruedas por reconocimiento personal.

En este caso el fallo de la sentencia fue absolutorio ante la falta de existencia de otras pruebas o indicios que acreditasen la autoría del investigado, en cuyo caso no habría nada que objetar al prevalecer el principio de presunción de inocencia. Pero ¿y si la mayor parte de esas irregularidades burocráticas, y sobre todo técnicas, fueran resueltas con la aplicación de unos estándares administrados con precisión por parte de lingüistas o fonetistas forenses?

Respecto a la jurisprudencia que recoge el valor de prueba de las técnicas identificativas, una reciente sentencia (SAP M 13864/2019), relativa a un robo con intimidación en un salón de juegos, recuerda lo que dos años antes dijo el Tribunal Supremo sobre que la prueba de reconocimiento en rueda

es propia de la instrucción sumarial y, por tanto, puede ser realizada en sede judicial, lo que no quita valor al hecho de que sea realizada por la policía, aunque luego deberá ser ratificada por el testigo en sede judicial. Lo destacable aquí es que, aunque la referencia procesal a la técnica de identificación es la rueda visual, en la sentencia se incluye una mención al reconocimiento «también por su voz», al considerarla parte del conjunto de rasgos identificativos (altura, complexión, indumentaria, etc.) que se pueden observar en el autor del delito a pesar de llevar su cara tapada.

### **3.2 Legislación y protocolos sobre el uso del reconocimiento de voz en rueda en algunos países de la Common Law**

Una aproximación al uso de las ruedas de reconocimiento de voz por parte de distintas jurisdicciones en las Islas Británicas muestra la disparidad de criterios respecto al uso de esta técnica identificativa y el modo en que cada comunidad legislativa afronta un proceso tan complejo, en el que «el testigo oye, memoriza, compara y recupera la voz de otros» (Robson, 2018), y cómo esa diversidad de enfoques puede producir diferencias en los resultados. Robson recoge la crítica a los tribunales australianos, canadienses y estadounidenses por sus enfoques *ad hoc* a la identificación de voz en rueda y señala que, mientras que para las ruedas visuales existen procedimientos perfectamente establecidos, en las ruedas auditivas no se dispone de criterios registrados que homogeneicen su elaboración en cada caso, como, por ejemplo, en lo que respecta al proceso de selección de distractores o la duración de las muestras; critica también que no se tengan en cuenta el resto de disciplinas que se entrecruzan en este proceso de identificación de la voz en el ámbito forense.

En Inglaterra y Gales, la mayor parte de los procesos que requieren una identificación de voz consideran que es una prueba más en el conjunto de evidencias aportadas. En la preparación del procedimiento existente, que tiene su base legal en el *Code D of the Police and Criminal Evidence Act* (Home Office, 2003), que incorporó explícitamente las ruedas auditivas en 2005, participaron expertos en fonética como MacFarlane y Nolan, pero dada la complejidad en el proceso de constitución de la rueda auditiva, la mayoría de los cuerpos policiales se niegan a utilizarla, lo que no parece ser visto por las autoridades judiciales como una brecha en el *Code of Practice*. Algunas de



las directrices establecidas son las siguientes: (i) cada muestra tiene una duración de 1 minuto; (ii) el testigo tiene que oír las un cierto número de veces; (iii) ha de registrarse habla espontánea del sospechoso; (iv) hay que localizar interrogatorios grabados de al menos 20 hablantes con similitudes respecto a edad, etnia, dialecto y sociolecto; (v) el tiempo transcurrido entre el hecho delictivo y la identificación se establece entre 4 y 6 semanas.

Desde un punto de vista práctico, el ensamblaje de la rueda lleva mucho tiempo, por lo que aumenta el plazo previsto para que el testigo realice la identificación; además, la excesiva duración de las muestras supone un esfuerzo cognitivo enorme para el testigo a la hora de realizar la tarea identificativa, y puede provocar una pérdida de atención por su parte (Robson, 2017). Y aunque el procedimiento ha sido revisado, se continúan realizando estudios que ponen en duda su eficacia, como parece quedar patente en el escaso número de veces que se recurre a su utilización:

Not only is there scope to reduce the length of the overall parade, but accuracy is higher when the voices are presented in a different way (Smith, Bird, Roeser, Robson, Braber, Wright y Stacey, 2020: 14).

En el caso de Escocia, un sospechoso no puede ser condenado solo a partir de la identificación de su voz por parte de un testigo. Aunque no existe instrucción alguna sobre la realización de ruedas identificativas de voz, que se realizan del mismo modo que las visuales y las hace fáciles de organizar, la falta de evidencia científica que apoye la fiabilidad del método utilizado debilita su eficacia. Está permitido realizar la identificación del acusado en vivo en la misma sala, donde se le sitúa entre voluntarios con *voz y acento similar*, amén de cuestiones relacionadas con edad, género y aspecto físico semejantes, y a quienes se les pide que hablen; después se pregunta al testigo si reconoce alguna voz en particular. En ningún momento es requerida la presencia de un experto para verificar esa similitud perceptivo-auditiva, ni tampoco existe un método estándar para describir la voz del sospechoso o de los distractores.

En Irlanda no existen directrices que regulen los procedimientos de la identificación de voz en rueda, pero el Tribunal de Apelaciones de Crowe determinó que la falta de un procedimiento similar a una rueda de reconocimiento puede afectar a la imparcialidad del juicio (Robson, 2018). Robson reconoce la potencial importancia de esta evidencia, que en muchos casos conlleva una identificación positiva, pero hace un llamamiento al sistema

legal: que sea especialmente sensible a los cambios implicados y que asegure que el proceso de la recopilación de evidencias hasta la celebración del juicio minimice los riesgos de distorsión y maximice la posibilidad de que las variables que impacten sobre la fiabilidad sean identificadas y abordadas. No tiene duda de que esos factores se identifican mejor en colaboración con otras disciplinas. La idea no es nueva y, aunque él la circunscriba a jurisdicciones propias del sistema británico de la *Common Law*, es perfectamente extrapolable a cualquier sistema judicial.

En general, podría resumirse que el proceso de preparación de una rueda de identificación de voz cuesta tiempo y dinero, por lo que mientras que en algunos países la policía hace uso de ellas solo cuando lo consideran realmente oportuno, en otros, simplemente, existen normas en contra de su utilización. En cualquier caso, la recomendación unánime de las autoridades en todos ellos es que cuando el tribunal se tenga que enfrentar a una identificación de voz en rueda no considere su admisibilidad de forma aislada, sino dentro del conjunto de pruebas acusatorias.

#### **4. Revisión metodológica en la elaboración de las ruedas de reconocimiento de voz**

Las revisiones metodológicas, así como el cuestionamiento del modo en que este tipo de ruedas identificativas se han venido realizando desde hace tiempo en distintas partes del mundo, no son planteamientos nuevos. Queda claro que no existe homogeneidad de criterios para establecer unas estrictas normas de cumplimiento que garanticen la infalibilidad de las decisiones tomadas por los testigos de un hecho delictivo en esta fase del proceso, pero lo que sí resulta patente es la necesidad de establecer una serie de directrices que refuercen su uso en el ámbito judicial como apoyo al conjunto evidencial que compone cada procedimiento. En línea con las investigaciones que se vienen realizando en distintos campos de conocimiento en relación con esta técnica y la experiencia obtenida en los tribunales, conviene plantear el establecimiento de estándares metodológicos que sirvan de guía, tanto en lo que respecta al diseño como al desarrollo de la prueba, que permitan aportar una mayor consistencia y fiabilidad a la misma.

El diseño de la rueda identificativa de voz conlleva una serie de pasos o etapas que no son de obligado cumplimiento, pues los expertos responsables de la misma pueden considerar en un determinado punto que no es viable

su continuidad, para lo que motivarán las causas que les hacen llegar a esa conclusión y mantendrán informada a la autoridad judicial competente en el caso. Pero ¿cómo se diseña una rueda auditivo-perceptiva y qué factores han de tenerse en cuenta para que resulte lo más fiable y segura posible para todas las partes intervinientes? En este punto, el control de las diferentes variables por el equipo de expertos minimizará los posibles errores que puedan darse en la identificación del posible sospechoso. La elección del método empleado es relevante en la resolución del proceso identificativo, y parece que el procedimiento de elección múltiple ofrece un mayor número de aciertos (Salmón Torralbo, 2019).

#### *4.1 Fases de desarrollo en la elaboración de un reconocimiento de voz en rueda*

##### *4.1.1 Obtención de datos mediante una entrevista*

Una vez que la autoridad judicial acuerda la realización de esta práctica identificativa, la primera fase se centrará en la obtención de la mayor cantidad de información posible sobre las circunstancias que, desde una óptica perceptiva, pudieran haber afectado o condicionado a la víctima o testigo, tanto en el transcurso como tras la comisión del delito, para lo que esta deberá entrevistarse con el equipo de especialistas (policías, lingüistas y psicólogos con experiencia forense) encargado de poner en marcha la rueda auditiva.

Con el fin de obtener todos los datos posibles respecto al hecho delictivo y la percepción que los testigos han podido tener durante el mismo respecto a la emisión oral de su autor, parece conveniente la realización de dos tipos de entrevista (Arce y Fariña, 2006): la narrativa, que aporta más información, y la interrogativa, siguiendo ese mismo orden a fin de evitar ofrecer información adicional en el planteamiento de las preguntas que pueda condicionar de algún modo al entrevistado. La evocación, nuevamente, de un hecho tan estresante puede causar intranquilidad, por lo que la pericia del entrevistador ayudará a establecer un ambiente tranquilo y relajado, que contribuirá a la buena disposición del testigo para llevar a buen término la entrevista.

No solo la pericia del entrevistador y la buena disposición del entrevistado harán que la entrevista fluya en la buena dirección, sino que son varios los factores que habrá que tener en cuenta para que dicha prueba se resuelva del mejor modo. Una variable fundamental es el tiempo que transcurre en-

tre el suceso y el momento en el que el afectado tiene que reconocer la voz del sospechoso, ya que la memoria auditiva tiende a disminuir con el paso del tiempo (Hollien, Huntley, Künzel y Hollien, 1995). En el ámbito forense se considera conveniente no exceder los tres meses entre los dos acontecimientos (Broeders y van Amelsvoort, 1999), lo que resulta complicado cuando los sistemas judiciales tienden a dilatar en demasía sus procedimientos. Además de recabar información relativa a ciertas circunstancias personales del entrevistado (edad, capacidad memorística, situación psico-emocional, entorno familiar y profesional, coherencia narrativa, etc.), deben valorarse otra serie de factores de mayor complejidad que tienen que ver con sus habilidades auditivo-perceptivas; por ejemplo, su capacidad discriminativa en la obtención de referencias sociolingüísticas, o qué datos concretos puede aportar sobre la persona a quien va a tratar de identificar (edad aproximada, físico, procedencia geográfica, variedad socio y geolectal, etc.); también características específicamente relacionadas con la emisión de su discurso (frases o enunciados y modo en el que han sido expresados, rasgos idiolectales y sociolectales, tiempo de exposición al discurso y longitud del mismo, etc.). Como complemento a toda esta información, resulta interesante la obtención de los datos situacionales que rodearon el hecho delictivo, el entorno acústico y aquellos estímulos acústicos que se produjeron de forma simultánea a la emisión hablada del autor en el momento de la agresión.

Con toda esta información, se valorará si la cantidad de frases emitidas por el agresor que la víctima es capaz de recordar, su extensión y el modo en que se produjeron (vociferadas, murmuradas, etc.) hacen posible seguir adelante con la elaboración de la rueda. En ese caso se creará un cuerpo de texto que se ajuste a la descripción realizada por el entrevistado, con una duración aproximada de 30 segundos siguiendo las recomendaciones de los expertos. La experiencia muestra que, en general, una gran parte de las víctimas son capaces de recordar las frases emitidas por su agresor, que normalmente son breves e incluyen en la mayoría de los casos expresiones groseras que quedan grabadas en su memoria de forma contundente. Será necesario tener en cuenta también el tipo de delito en el que se hayan producido y la interacción entre sus protagonistas, dado que no es lo mismo que se produzca un robo con intimidación en la vía pública, donde el autor obtenga rápidamente su botín, que si este se produce en un establecimiento con rehenes donde la víctima oponga cierta resistencia y se produzca un intercambio comunicativo entre ambos.

#### 4.1.2 Toma de muestras de habla del sospechoso

Tras considerar que las circunstancias expresadas por la víctima para la posible identificación del sospechoso son las adecuadas para continuar con el procedimiento, ante la presencia de la autoridad judicial se procederá a realizar la toma de muestras de voz a la persona interesada, respetando en todo momento el debido cumplimiento de los requisitos legales pertinentes. El investigado debe ser informado del motivo por el que se encuentra en sede judicial y dar su consentimiento para la realización de la prueba; está en su derecho negarse a realizarla, lo que supondría dar por concluido el proceso. En el caso de que quiera colaborar, el equipo de expertos le describirá en qué va a consistir el proceso de registro de su voz, para cuya grabación se utilizará un dispositivo de uso profesional que cumpla con los requisitos de calidad e idoneidad.

La grabación se adecuará en lo posible a las circunstancias acústico-expresivas descritas por el afectado durante la entrevista. En la sesión se realizarán varias tomas de muestras de voz donde el locutor participará de forma activa emitiendo las frases seleccionadas en la forma que el experto le solicite, y conforme a las variaciones tonales que la víctima describió según las recordaba para adecuar en lo posible el contexto comunicativo en el que se produjeron:

If possible, verbatim speech as produced by the perpetrator during the incident (according to the witness) should also be recorded. If these utterances are very short, they should be repeated a number of times. [If these verbatim utterances are (likely to be) perceived as threatening by the witness or the suspect, or are otherwise deemed prejudicial to the exercise, they are better omitted] (Broeders y van Amelsvoort, 1999: 1374)

En el caso de que la víctima hubiera manifestado que el agresor cubría su boca o su rostro al completo con algún tipo de elemento, es conveniente que el mismo locutor reproduzca nuevamente las frases con una prenda similar, para que esa adecuación resulte, si cabe, más fidedigna. Una vez efectuada la grabación, de la que tendrá copia la autoridad judicial, se lleva a cabo el estudio de los rasgos de caracterización de la voz del sospechoso, es decir, se establecerán aquellos componentes fundamentales y característicos que individualizan su voz (por ejemplo, si es voz joven o más bien adulta, tipo de voz y de articulación, frecuencia fundamental, etc.) para obtener información lo más precisa posible sobre la cualidad tímbrica de la voz registrada y

conseguir establecer de ese modo un patrón-timbre que caracterice la voz del sospechoso, que será decisivo para llevar a cabo una de las fases más complicadas del procedimiento: la selección de los modelos de ‘voces contraste’, o voces auxiliares, que compondrán la rueda junto con la del sospechoso.

#### 4.1.3 Selección de modelos de voces contraste

La ejecución de esta fase es la más complicada. El objetivo es obtener muestras de voz similares a la del sospechoso provenientes de otros hablantes modelo que servirán para confeccionar una cadena editada de voces que contendrá fragmentos de discurso, aleatoriamente distribuidos e identificados con códigos, correspondientes tanto a la voz del sospechoso como a la de las voces contraste. La recogida de estas muestras debe ser realizada por el mismo equipo de especialistas responsables de la elaboración de la rueda, siguiendo el criterio de adecuación perceptiva aportado por la víctima, pero teniendo especialmente en cuenta los rasgos de la cualidad de la voz del sospechoso; es decir, será el perfil lingüístico el que marque las pautas para seleccionar distintos rasgos asociados al habla, tanto de tipo biológico, como social y discursivo (timbre, tono, articulación, etc.), que se ajusten al patrón-timbre extraído del registro de su voz para localizar las voces modelo de los locutores auxiliares o distractores. No resulta sencillo describir correctamente una voz, así que en las ruedas auditivas el criterio que prima a la hora de seleccionar a los locutores auxiliares es el de similitud con el sospechoso, es decir, la selección de voces contraste se realizará fundamentalmente teniendo en cuenta las características lingüísticas y expresivas del sospechoso. En caso de observar en la base articulatoria del agresor rasgos marcados de un área específica, las voces de los impostores deberán recogerse en la misma zona:

[...] for the earwitness line-up anyway, foils should not only meet the description of the voice by the witness. They should also meet the speech profile of the suspect as made by a competent phonetician, and be matched for features such as biological and social gender, perceived age, accent, dialect and the like (Broeders y van Amelsvoort 2001: 241)

#### *Nivel de similitud de las voces*

¿Y cuál será el nivel de similitud que deben presentar estas voces respecto a la del sospechoso? Pues justamente el adecuado. Por un lado, un alto grado de similitud podría incrementar la posibilidad de seleccionar un falso culpable; por otro, si la diferencia es muy obvia (por ejemplo, incluir voces auxiliares con variables geolingüísticas de la sierra granadina si la voz del sospechoso presenta una base articulatoria de habla dialectal del norte de Marruecos) se estaría señalando claramente en una dirección. Para mantener la eficacia e imparcialidad en la construcción de la rueda auditiva, el equilibrio estaría en que una de las voces auxiliares presente rasgos diferenciadores perceptibles para un oído no entrenado, que otra se asemeje lo máximo posible a la voz del sospechoso, mientras que el resto se mantendría en un nivel intermedio de similitud perceptiva (Hollien, 2012). Se ha de recordar que en este proceso selectivo es necesario tener en cuenta las variables básicas relativas al autor que la víctima aportó durante su entrevista (sexo, edad, variedad dialectal, etc.), pero por encima de todo el grado de similitud con la voz del sospechoso:

[...] what needs to be assessed to ensure the fairness of a voice parade is not «voice quality», but «speaker similarity» —of which, voice quality may be only one component (Nolan, 2007: 123)

Por lo tanto, lo relevante a la hora de seleccionar las voces auxiliares es centrarse en delimitar el posible rango de edad, la localización geográfica, el timbre, la línea entonativa y conseguir una adecuación perceptiva en los registros de todas las emisiones. La selección de locutores puede llevarse a cabo en el entorno laboral de los propios especialistas con voluntarios, siempre que no exista ningún tipo de relación o vinculación con el hecho delictivo ni con ninguno de los implicados. Se les grabará con el mismo equipo utilizado para el registro de la voz del sospechoso y en circunstancias similares, es decir, se les hará repetir varias veces las frases seleccionadas, prestando especial atención a la línea entonativa de cada emisión, y haciendo uso de un cubrebocas en caso de que fuera necesario. En ningún momento las muestras registradas se podrán asociar a ninguna persona, ya que se les asignará un número de identificación de modo que en ningún momento puedan ser reconocidas, al igual que se hace con el locutor sospechoso.

*Número de distractores*

Respecto al número total de distractores tampoco parece haber consenso, pero sí respecto a la cantidad mínima considerada aceptable para poder elaborar las cadenas de voz, que Broeders y van Amelsvoort (1999) establecieron en cinco. Sobre el perfil de los distractores en lo que respecta a su edad, estudios o contacto de lenguas, no hay que mostrar rigidez, no es un factor invalidante *per se* el que se amplie el intervalo de edad o que alguno de los distractores haya vivido durante años en una zona geográfica de la que no es originario o que tenga estudios universitarios o no. En lo que sí parece no haber duda es que en la selección de los distractores debe primar la similitud perceptiva de sus voces con la del sospechoso, teniendo en cuenta tanto rasgos lingüísticos como paralingüísticos, variedad dialectal, tono o *pitch*, velocidad discursiva, frecuencia de las pausas o articulación (Smith y Baguley, 2014).

#### *Selección de tramos*

Una vez obtenidos todos los registros de las voces auxiliares se realizará un estudio perceptivo de las muestras para seleccionar aquellos tramos que más se ajusten al modelo emitido por el sospechoso. Tras la realización de una prueba control en que se reproducen repetidas veces las locuciones registradas de los distractores y del sospechoso, se toma una decisión sobre aquellas frases consideradas más precisas en su emisión ateniéndose a las circunstancias que a nivel perceptivo fueron experimentadas por la víctima, obteniendo un discurso de cada uno de los hablantes y procediendo a armonizar la longitud de las muestras y la intensidad de la señal.

#### *4.1.4 Elaboración de las cadenas de voces*

Es conveniente tener siempre presente que el objetivo último de esta prueba es que el testigo realice una serie de tomas de decisión para determinar o descartar la posible identificación de la voz de un sospecho, y que, si se dilatan en exceso, pueden llegar a provocar agotamiento físico y auditivo.

Tratando de buscar el equilibrio entre las recomendaciones ofrecidas por los expertos, sería aconsejable preparar dos o tres bloques de respuesta, cada uno de los cuales contendría entre 5 y 7 locuciones seleccionadas y repetidas por el total de locutores que compondrán cada cadena, incluido el sospechoso, si bien en alguna ocasión se puede tomar la decisión de no incluir a



este en alguna de las cadenas. La secuenciación de las voces incluidas en las cadenas será aleatoria, evitando la repetición consecutiva en la ubicación de las muestras, así como la cercanía en su emparejamiento.

Para terminar de estructurar esas cadenas será necesario editar las muestras de audio procediendo a intercalar un estímulo auditivo o ‘voz control’, normalmente del sexo opuesto a las voces que componen la rueda (Broeders y van Amelsvoort, 1999), que introduzca tanto cada una de las frases como cada una de las muestras de voz, controlando, además, que la duración de las pausas entre uno y otro estímulo sean de algo más de 1 segundo en todos los casos.

Una vez finalizada esta tarea se elaborará un formulario de respuestas en el que la víctima recogerá cada toma de decisión que vaya realizando a la hora de identificar o no la voz del sospechoso entre los locutores que componen cada una de las cadenas.

#### 4.1.5 Rueda de reconocimiento de voz

El último paso es la realización propiamente dicha del reconocimiento de voz en rueda, en el que en presencia de la autoridad judicial la víctima realizará la tarea de identificar una voz de entre el grupo de voces que se le presentan. Esta fase conlleva una toma de múltiples decisiones dirigida a reducir, en la medida de lo posible, los elementos de aleatoriedad y subjetividad inherentes a los procesos de asociación cognitiva. Con el fin de que la práctica de la prueba se desarrolle en las circunstancias más favorables es conveniente anticipar el cuidado de ciertas cuestiones, entre las que cabe destacar no solo la concentración de quien va a realizar la prueba, sino también el aislamiento acústico-ambiental del lugar donde se va a realizar. A ello hay que añadir la utilización de un buen equipo de reproducción y emisión de audio, que deberá estar correctamente ubicado respecto al receptor, ajustando tanto intensidad como frecuencia de salida de la señal sonora.

Una vez todo dispuesto, se le explicará a la víctima a qué se va a enfrentar, subrayando que va a colaborar en un proceso de reconocimiento de voz mediante percepción auditiva, para lo que tendrá que emitir una decisión de identificación cada vez que crea reconocer la voz del presunto agresor entre las cadenas de voces que se le presenten, y de lo que dejará constancia en el formulario que se le haya entregado. Como parte del proceso de adecuación, resulta conveniente que con tacto se le pidiera recordar la voz objeto del re-

conocimiento, y que intente recrear mentalmente las circunstancias de la escena del delito justo en el momento en que aconteció. No hay que olvidar que se exponen al recuerdo de algo vivido como traumático, y es normal que puedan mostrar una cierta incomodidad, además de sentirse responsables de sus posibles respuestas.

Para la realización de esta prueba, la víctima ha de estar lo más tranquila posible, centrando su atención en los estímulos sonoros que se le vayan presentando, tomándose el tiempo que considere necesario para emitir una respuesta y, si lo necesitara, pedir que se repita alguna de las secuencias, al menos una vez. Se le recordará que si no está segura no tiene obligación de dar una respuesta. Otros detalles acerca de la prueba que deben aclarársele serían, por ejemplo, los relacionados con el diseño de la rueda o que la voz del sospechoso puede no encontrarse en alguna de las cadenas, si ese fuese el caso. Para que el proceso quede claro, es recomendable realizar una prueba de adecuación perceptiva en el que se reproduce una de las cadenas seleccionada aleatoriamente.

Se reitera a la víctima que debe escuchar atentamente las grabaciones y, si reconociera la voz del agresor entre las voces que se le presentan en cada cadena, deberá marcar con una cruz en la letra que corresponda del formulario entregado previamente. Se le insistirá en que su respuesta debe ser sincera, señalando únicamente en caso de reconocer la voz de su agresor. Será también informada de que puede solicitar un receso si lo necesitara. Normalmente, a medida que avanza la presentación de las cadenas, la víctima presenta un mayor grado de exactitud en sus respuestas y un mayor convencimiento en sus contestaciones al ir disipando sus posibles dudas iniciales.

#### **4.2 Expresión de resultados y conclusiones**

Se documentará todo lo que ocurra en el propio acto de la identificación, desde la explicación acerca de en qué consiste el proceso que se va a llevar a cabo hasta la hora de inicio o de finalización; así como el tiempo que se tardó en dar una respuesta, o si se plantearon dudas, incluso si se solicitó la repetición de alguna cadena o un breve descanso en la realización de la tarea, y cualquier circunstancia que se considere relevante. Con todos esos datos, y tras un detallado análisis de los resultados obtenidos que se incluirán en el correspondiente informe, la expresión de las conclusiones se realizará de forma clara y concisa, sin valoraciones, consideraciones u opiniones perso-

nales, sino ciñéndose a la presentación objetiva de los resultados y la actitud o disposición de la víctima o testigo en el momento de la identificación, pues corresponderá a la autoridad judicial valorar en última instancia las decisiones alcanzadas por estos.

## 5. Conclusiones

La desconfianza ante el uso de esta técnica como parte del proceso judicial se centra sobre todo en la poca seguridad que ofrece la memoria de una víctima tras haber sufrido una agresión que le ha supuesto, en la mayor parte de los casos, una grave situación de estrés durante el tiempo en que se produjo y, en algunos casos, con posterioridad al mismo. Frente a esta desconfianza en las capacidades de una víctima, a quien algunos estudios consideran capaz de retener detalles muy precisos sobre lo que le ha ocurrido (que, además, tiene el aliciente de poder resarcirse, al menos emocionalmente, con una resolución judicial que cierre el caso), se encuentra la otra cara de la moneda, la ancestral costumbre de ciertos participantes en el proceso de considerarse más que capaces de realizar valoraciones técnicas, aunque sin conocimiento especializado, porque así se les presupone. Quizá entre una y otra consideración esté el equilibrio: contemplar la identificación de voz en rueda como una prueba más entre el conjunto de pruebas que se presentan en cada proceso judicial aporta más peso a dicho conjunto evidencial, lo que en sí mismo es positivo, y recurrir a ella toda vez que sea necesario y siempre bajo supervisión profesional.

No es solo cuestión de cautela respecto al recuerdo de la víctima, sino que es necesario considerar todas las variables posibles de forma precisa estableciendo límites en su desarrollo en el caso de que se pudiera producir algún *fallo* en su implementación: tal vez el tiempo transcurrido haya sido excesivo, a lo mejor la víctima no se encuentra en condiciones óptimas para aportar detalles sobre el agresor, quizá este se niegue a participar en el montaje de la rueda, o no se puedan conseguir distractores suficientes con el dialecto propio del sospechoso... No es preceptivo llegar hasta el final de la misma, el equipo de expertos deberá tomar una decisión al respecto cuando así lo considere. Pero si ese es el caso, y tras la presentación del informe de resultados y su posterior ratificación ante el tribunal, queda un último y decisivo trámite: la valoración final de esos resultados quedará en manos del juez, quien será el que considere, o no, que esa prueba de identificación tiene

todas las garantías procesales para formar parte del conjunto de evidencias que componen el procedimiento.

Sería sencillo que en el informe se recogiese una conclusión contundente que asegurase que el nivel de identificación fuera del 100 %, facilitando así la labor del juez. Pero ese no es el caso. Pocas veces se puede hablar de certeza absoluta y menos en cuestiones tan comprometidas como la validez procesal de la recuperación del recuerdo de una voz desconocida, oída en una única ocasión en circunstancias muy estresantes, con el fin de identificar al autor de un delito para el que está previsto una corrección penal, porque, reconozcámoslo, ni las matemáticas son verdades absolutas. Sirva como ejemplo un caso en que el testigo tuviera que responder en 18 ocasiones si ha reconocido o no la voz de su agresor entre las cadenas presentadas, sobre la premisa de que en todas ellas consta incluido el sospechoso, y los resultados que se muestran son: en 3 ocasiones decide no responder ante la duda, en 4 selecciona un distractor y en las 11 restantes identifica correctamente la voz del sospechoso. ¿Qué valoración se haría sobre ese resultado? ¿Cuál en caso de que las 18 ocasiones hubiese discriminado correctamente la voz diána? ¿Este último resultado tendría un valor absoluto respecto al primero? Es cierto que en estos casos no se recurre a los números para la expresión de las conclusiones, aunque sí en demasiadas ocasiones al socorrido recurso de utilizar la expresión «sin ningún género de duda» por parte de los especialistas forenses; y esa es una máxima absoluta. Para Verdú Pascual (2012), que la considera una frase desafortunada en el lenguaje pericial, no es labor del perito asumir el papel del juzgador porque le impedirá a este último cumplir con algo consustancial a la administración de justicia: la valoración de la viabilidad o no de su consideración como prueba.

La existencia de unos estándares relativos a la composición de la rueda y su puesta en marcha, el concienzudo trabajo de lingüistas y psicólogos con especialización en el campo del mundo forense, el intercambio de información respecto a la elaboración de la rueda auditivo-perceptiva con los implicados en cada parte del proceso, la continua actualización del conocimiento respecto a nuevos datos obtenidos en rigurosas investigaciones, la ratificación tanto de los expertos como de los testigos en la vista oral, entre otras cuestiones, pueden contribuir a que el proceso de identificación de voz en rueda se convierta en una evidencia con coherencia suficiente que fortalezca su eficacia dentro del procedimiento judicial. La utilización de esta técnica en aquellos casos donde se requiera, bajo supervisión de expertos forenses siguiendo una estandarización metodológica, ayudará a que la valoración fi-

nal tenga una buena fundamentación técnica que sustente la decisión última del tribunal.

## Bibliografía

- Arce, Ramón; y Fariña, Francisca (2006).** «Psicología del testimonio: evaluación de la credibilidad y de la huella psíquica en el contexto penal», en José Antonio Ballester Pascual (dir.), *Psicología del testimonio y prueba pericial*, Madrid, Consejo General del Poder Judicial, pp. 39-103, [https://www.usc.es/export9/sites/webinstitucional/gl/servizos/uforense/descargas/Psicologia\\_del\\_testimonio\\_evaluacion\\_de\\_la\\_credibilidad\\_y\\_de\\_la\\_huella.pdf](https://www.usc.es/export9/sites/webinstitucional/gl/servizos/uforense/descargas/Psicologia_del_testimonio_evaluacion_de_la_credibilidad_y_de_la_huella.pdf).
- Ballesteros Jiménez, Soledad (1997).** *Psicología General. Aprendizaje y memoria. Memorias sensoriales*, Madrid, Editorial Universitas.
- Bartle, Anna; y Dellwo, Volker (2015).** «Auditory speaker discrimination by forensic phoneticians and naive listeners in voiced and whispered speech», *The International Journal of Speech, Language and the Law*, 22, 2, pp. 229-248, <https://doi.org/10.1558/ijsll.v22i2.23101>.
- Belin, Pascal; Fecteau, Shirley; y Bédard, Catherine (2004).** «Thinking the voice: neural correlates of voice perception», *Trends in Cognitive Sciences*, 8, pp. 129-135, <https://doi.org/10.1016/j.tics.2004.01.008>.
- Broeders, A. P. A.; y van Amelsvoort, A. G. (1999).** «Lineup construction for forensic earwitness 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>.
- **y van Amelsvoort, A. G. (2001).** «A practical approach to forensic earwitness identification: Constructing a voice lineup», en Józef Wójcikiewicz (ed.), *Proceedings of EAFS Meeting. Problems of Forensic Sciences*, Cracovia, vol. XLVII, pp. 237-245.
- Broeders, A. P. A.; y Rietveld, A. C. M. (1995):** «Speaker Identification by Earwitnesses», en Angelika Braun y Jens-Peter Köster (eds.), *Studies in Forensic Phonetics*, 64, pp. 24-40.
- Clifford, Brian R. (1980).** «Voice Identification by Human Listeners: On Earwitness Reliability», *Law and Human Behavior*, 4, 4, pp. 373-394.
- Cook, Susan; y Wilding, John (1997).** «Earwitness testimony: Never mind the variety, hear the length», *Applied Cognitive Psychology*, 11, 2, pp. 95-111, [https://doi.org/10.1002/\(SICI\)1099-0720\(199704\)11:2<95::AID-](https://doi.org/10.1002/(SICI)1099-0720(199704)11:2<95::AID-)

- ACP429>3.o.CO;2-0.
- Doddington, G. R. (1985).** «Speaker recognition—Identifying people by their voices», *Proceedings of the IEEE*, 73, 11, pp. 1651-1665, <https://doi.org/10.1109/PROC.1985.13345>.
- Garayzábal Heinze, Elena; Queralt Estévez, Sheila, y Reigosa Riveiros, Mercedes (2019).** *Fundamentos de la lingüística forense*, Madrid, Síntesis.
- Hollien, Harry (2002).** *Forensic voice identification*, San Diego, Academic Press.
- (2012). «On earwitness lineups», *Investigative Sciences Journal*, 4, 1, pp. 1-17.
- 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, <https://doi.org/10.1558/ijll.v2i2.143>.
- Kerstholt, J. H.; Jansen, N. J.; van Amelsvoort, A. G; y Broeders, A. P. A. (2004).** «Earwitnesses: Effects of speech duration, retention interval and acoustic environment», *Applied Cognitive Psychology*, 18, 3, pp. 327-326, <https://doi.org/10.1002/acp.974>.
- Kreiman J.; y Papcun G. (1991).** «Comparing discrimination and recognition of unfamiliar voices», *Speech Communication*, 10, 3, pp. 265-75, [https://doi.org/10.1016/0167-6393\(91\)90016-M](https://doi.org/10.1016/0167-6393(91)90016-M).
- 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/ijll.v1i1.45>.
- Künzel, H. J.; Masthoff, H. R.; y Köster, J. P. (1995).** «The relation between speech tempo, loudness, and fundamental frequency: an important issue in forensic speaker recognition», *Science & Justice: Journal of the Forensic Science Society*, 35, pp. 291-295, [https://doi.org/10.1016/S1355-0306\(95\)72683-1](https://doi.org/10.1016/S1355-0306(95)72683-1).
- Ladefoged, Peter (1975).** *A Course in Phonetics*, Nueva York, Harcourt Brace Jovanovich.
- Laubstein, Ann Stuart (1997).** «Problems of voice line-ups», *International Journal of Speech, Language and the Law*, 4, 2, pp. 262-279.
- Laver, John; y Trudgill, Peter (1979).** «Phonetic and linguistic markers in speech», en Klaus R. Scherer y Howard Giles (eds.), *Social Markers in Speech*, Cambridge, Cambridge University Press, pp. 1-3.
- Legge, Gordon E.; Grossmann, Carla; y Pieper, Christian M. (1984).** «Learning unfamiliar voices», *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10, 2, pp. 298-303, <https://doi.org/10.1037/0278-7393.10.2.298>.
- 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).
- Nolan, Francis (2001).** «Speaker identification evidence: Its forms, limitations and roles», en Tarja Salmi-Tolonen, Richard Foley e Iris Tukiainen (eds.), *Proceedings from the conference «Law and Language – Prospects and Retrospects»*, Rovaniemi, Universidad de Laponia.
- (2007). «Voice quality and forensic speaker identification», *GOVOR: Caso-pio Za Fonetiku*, 24, 2, pp. 111-128.
- Nolan, Francis; y Oh, Tomasina (1996).** «Identical twins, different voices», *The International Journal of Speech, Language*

- and the Law, 3, 1, pp. 39-49, <https://doi.org/10.1558/ijssl.v3i1.39>.
- Orchard, Tara L.; y Yarmey, A. Daniel (1995).** «The effects of whispers, voice-sample duration, and voice distinctiveness on criminal speaker identification», *Applied Cognitive Psychology*, 9, 3, pp. 249-260, <https://doi.org/10.1002/acp.2350090306>.
- Perrachione, Tyler K.; Furbeck, Kristina T.; y Thurston, Emily J. (2019).** «Acoustic and linguistic factors affecting perceptual dissimilarity judgments of voices», *Journal of the Acoustical Society of America*, 146, 5, pp. 3384-3399, <https://doi.org/10.1121/1.5126697>.
- Pisoni, David B. (1993).** «Long-term memory in speech perception: Some new findings on talker variability, speaking rate and perceptual learning», *Speech Communication*, 13, 1-2, pp. 109-125. [https://doi.org/10.1016/0167-6393\(93\)90063-q](https://doi.org/10.1016/0167-6393(93)90063-q).
- Read, Daniel; y Craik, Fergus I. M. (1995).** «Ear-witness identification: some influences on voice recognition», *Journal of Experimental Psychology: Applied*, 1, 1, pp. 6-18, <https://doi.org/10.1037/1076-898X.1.1.6>.
- Robson, Jeremy (2017).** «A Fair Hearing? The use of voice identification parades in criminal investigations in England and Wales», *Criminal Law Review*, 1, pp. 36-50.
- (2018). «'Lend me your ears': An analysis of how voice identification evidence is treated in four neighbouring criminal justice systems», *The International Journal of Evidence & Proof*, 22, 3, pp. 218-238, <https://doi.org/10.1177/1365712718782>.
- Salmón Torralbo, Beatriz (2019).** *Un estudio de lingüística aplicada: El reconocimiento de voces en el ámbito forense*, TFG dirigido por María García Antuña, Comillas, CIESE-Comillas/Universidad de Cantabria, <https://repositorio.unican.es/xmlui/bitstream/handle/10902/17999/TFG.BST.pdf?sequence=1&isAllowed=y>.
- Sarwar, Farhan; Allwood, Carl Martin, y Zetterholm, Elisabeth (2014).** «Earwitnesses: the type of voice lineup affects the proportion of correct identifications and the realism in confidence judgments», *The International Journal of Speech, Language and the Law*, 39, 1, pp. 139-155.
- Saslove, H.; y Yarmey, A. D. (1980).** «Long term auditory memory: Speaker identification», *Journal of Applied Psychology*, 165, pp. 111-116, <https://doi.org/10.1037/0021-9010.65.1.111>.
- Smith, Harriel M. J.; 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>.
- Smith, Harriet M. J.; Bird, Katherine; Roeser, Jens; Robson, Jeremy; Braber, Natalie; Wright, David; y Stacey, Paula C. (2020).** «Voice parade procedures: optimising witness performance», *Memory*, 28, 1, pp. 2-17. <https://doi.org/10.1080/09658211.2019.1673427>.
- Solan, Lawrence; y Tiersma, Peter. (2003).** «Hearing voices: Speaker identification in court», *Hastings Law Journal*, 54, pp. 373-436, <https://brooklynworks.brooklaw.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1284&context=faculty>.
- Valiente Martín, Ana Belén (2014).** «La influencia del número de distractores en el comportamiento de los testigos de una rueda de reconocimiento de voces», en Yolanda Congosto Martín, María Lui-



- sa Montero Curiel y Antonio Salvador Plans (eds.), *Fonética experimental, educación superior e investigación. I. Fonética y Fonología*, Madrid, Arco/Libros, pp. 669-686.
- Verdú Pascual, Fernando Alejo (2012).** «Una frase desafortunada en el lenguaje peyorativo: “Sin ningún género de duda”», *Gaceta Internacional de ciencias forenses*, 5, 1, [https://www.uv.es/gicf/1Edi\\_Verdu\\_GICF\\_05.pdf](https://www.uv.es/gicf/1Edi_Verdu_GICF_05.pdf).
- Yarmey, A. Daniel (1991).** «Description of distinctive and nondistinctive voices over time», *Journal of the Forensic Science Society*, 31, pp. 421-428, [https://doi.org/10.1016/S0015-7368\(91\)73183-6](https://doi.org/10.1016/S0015-7368(91)73183-6).
- (1995). «Earwitness speaker identification», *Psychology, Public Policy and Law*, 1, 4, pp. 792-816, <https://doi.org/10.1037/1076-8971.1.4.792>.
- (2007). «The psychology of speaker identification and earwitness memory», en R. C. L. Lindsay, David F. Ross, J. D. Read y Michael P. Toglia (eds.), *The Handbook of Eyewitness Psychology. II. Memory for People*, Mahwah, NJ, Lawrence Erlbaum Associates Publishers, pp. 101-136.
- Zäske, Romi; Hasan, Bashard Awwad Shiekh; y Belin, Pascal (2017).** «It doesn't matter what you say: fMRI correlates of voice learning and recognition independent of speech content», *Cortex*, 94, pp. 100-112, <https://doi.org/10.1016/j.cortex.2017.06.005>.

## LEGISLACIÓN

- Advice on the use of voice identification paradigms. Home Office circular 057/2003.** <https://webarchive.nationalarchives.gov.uk/ukgwa/20130125153221/http://www.homeoffice.gov.uk/about-us/corporate-publications-strategy/home-office-circulars/circulars-2003/057-2003/>.
- Real Decreto de 14 de septiembre de 1882 por el que se aprueba la Ley de Enjuiciamiento Criminal.** *Gaceta de Madrid*, Ministerio de Gracia y Justicia, núm. 260, de 17 de septiembre de 1882, pp. 803-806 (referencia: BOE-A-1882-6036), [https://www.boe.es/eli/es/rd/1882/09/14/\(1\)](https://www.boe.es/eli/es/rd/1882/09/14/(1)).
- Ley 1/2000, de 7 de enero, de Enjuiciamiento Civil.** Jefatura del Estado, BOE, núm. 7, de 08 de enero de 2000 (referencia: BOE-A-2000-323), <https://www.boe.es/buscar/pdf/2000/BOE-A-2000-323-consolidado.pdf>.
- Poder Judicial España.** Jurisprudencia, <https://www.poderjudicial.es/cgpj/es/Servicios/Jurisprudencia/>.