

**YO LO RECONOCERÍA POR SU VOZ...  
EL CASO DE EMILIO BERKHOFF**

*I would recognize him by his voice... The case of Emilio Berkhoff*

CLAUDIA ROSAS

*Universidad Austral de Chile (Chile)*  
*claudiarosas@uach.cl*

JORGE SOMMERHOFF

*Universidad Austral de Chile (Chile)*  
*jsommerh@uach.cl*

JAIME PACHECO

*Universidad de Concepción (Chile)*  
*jaipacheco@udec.cl*

CÉSAR SÁEZ

*Laboratorio de Criminalística*  
*Policía de Investigaciones (Chile)*  
*csaeze@investigaciones.cl*

Resumen

Durante la comisión de un crimen, testigos o víctimas, expuestos al habla del agresor pueden ser llamados, en forma posterior, a reconocer la voz de quien se sospecha es el autor del delito, con el fin de establecer si ellos la reconocerían como la voz del delincuente. Un ejemplo de esta situación, donde se vio involucrado el testimonio de testigos auditivos, fue durante el proceso del caso de Emilio Berkhoff (Chile, 2015), el joven universitario que fue declarado culpable de los delitos de porte ilegal de armas y municiones y condenado a cinco años y un día de cárcel. El caso no estuvo exento de cierta controversia debido a las declaraciones públicas del imputado sobre supuestas violaciones al debido proceso. En este trabajo se revisa el procedimiento ocupado para la recolección de la evidencia auditiva a la luz de los criterios más ampliamente compartidos sobre la materia.

Conceptos clave: Identificación de voces en contextos forenses; reconocimiento del hablante por testigos auditivos no profesionales.

Abstract

During the commission of a crime, witnesses or victims, exposed to the speech of the aggressor can be called, later, to recognize the voice of the person suspected of being the perpetrator of the crime, in order to establish whether they would recognize it as the voice of the offender. An example of this situation, where the testimony of hearing witnesses was involved, was during the case of Emilio Berkhoff (Chile, 2015), the young university student who was found guilty of the crimes of illegal bearing of arms and ammunition and sentenced to five years and one day in jail. The case was not exempt from some controversy due to the public statements of the accused about

alleged violations of due process. In this paper we review the procedure used to collect the auditory evidence in light of the most widely shared criteria on the subject.

Key words: Voices identification in forensic contexts; speaker recognition by naïve earwitness.

## INTRODUCCIÓN

Una identificación por voz en casos criminales mediante víctimas o testigos puede aparecer en aquellas situaciones en las que no ha quedado un registro de habla del delincuente, pero la víctima o testigo piensa que, sin conocer a la persona que habló en esa oportunidad podría reconocerla si la volviera a oír, en cuyo caso la identidad debe establecerse.

Con miras a identificar la voz del agresor, la víctima o testigo puede escuchar una muestra de habla y señalar si es o no la del delincuente (identificación de a uno); o bien, oír una serie de voces y seleccionar una de ellas como la del perpetrador del crimen (rueda de reconocimiento). Su uso en la Corte –en cualquiera de las dos modalidades– se basa en la creencia de que la mayoría de las personas, con o sin conocimientos lingüísticos, pueden identificar personas solo por su voz; por ejemplo, voces de familiares, amigos, conocidos, etc. (Nolan, 2001). Y, aunque en muchos casos, reconocemos de manera fiable quien habla, sin verle, no siempre ocurre así; porque no conocemos de manera previa a la persona con quien interactuamos, por ejemplo, cuando nos llama un vendedor de seguros, el ejecutivo de un banco, etc., o solo, porque, a pesar de conocerla, no la reconocemos por diversas razones; como, por ejemplo, cuando quien nos habla es un amigo resfriado, un conocido que se hace pasar por otro para hacernos una broma, el hermano de un amigo quien responde la llamada, etc. Si comprobamos errores como estos en nuestra vida diaria, ¿por qué no pensar que también puedan presentarse desaciertos en otras circunstancias como, por ejemplo, en una situación criminal? Porque existe en la población, en general, el mito de que toda identificación es precisa y confiable, pero, por desgracia, cuando este alcanza la esfera de la justicia, puede resultar muy perjudicial, como en el controvertido juicio de Bruno Hauptmann, por el secuestro del hijo del héroe nacional (inglés) Charles Lindberg, en 1932<sup>1</sup> (Kreiman y Sidtis, 2011). En efecto, el exceso de confianza en el reconocimiento auditivo puede ser la causa de muchas condenas injustas, como informa en su estudio Laub (2010).

---

<sup>1</sup> Lindberg oyó la voz del hombre a quien él pagó el rescate (quien pudo haber sido el perpetrador del crimen) pronunciar una frase como “*Hey, Doctor, over here, over here*” o solo “*Hey, Doctor!*” en la noche y a una distancia de 70 yardas (65 metros Aprox). Al menos, tres años después, en el juicio de Hauptmann, Lindeberg testificó, diciendo que la voz escuchada aquel día era la de Bruno Hauptmann y que luego la reconoció en la oficina del fiscal at District Attorney Foley, en nueva York, en el Bronx). Aunque fueron presentadas varias evidencias circunstanciales que apoyaban la presunción de culpabilidad, los jurados reportaron más tarde que el testimonio de Lindbergh fue un factor crítico en sus deliberaciones.

Hay muchos ejemplos de identificaciones erróneas en la literatura, desde los más emblemáticos como el ocurrido en Inglaterra en 1660 donde un soldado llamado William Hulet fue enjuiciado y condenado, de manera injusta, de llevar a la decapitación al Rey Inglés Carlos I en 1649 (Kreiman y Sidtis, 2011)<sup>2</sup> hasta los más mundanos como aquellos relacionados, por ejemplo con ataques de ladrones enmascarados y violadores, crímenes cometidos en la oscuridad, llamadas telefónicas obscenas o de amenazas, etc. En estas situaciones, una identificación del sospechoso como el hablante oído durante la comisión del crimen es a menudo una importante pieza de evidencia, y algunas veces, la única pieza de evidencia contra el acusado. En dichos casos, el testimonio es confiable en la medida en que los auditores pueden oír una voz una vez y, más tarde, distinguirla con precisión de otras voces, posiblemente similares. Al respecto, es necesario tener en cuenta que el reconocimiento de voces no oídas con anterioridad, después de una exposición limitada, es una tarea antinatural, que demanda, con seguridad, estrategias perceptivas y cognitivas distintas a las empleadas en el reconocimiento de voces familiares, y que, como cualquier tarea de identificación, no puede ostentar una certeza absoluta. Kirk y Kingston, citados por Aitken y Taroni (2010, p. 39) se han referido a esta condición de incertidumbre del siguiente modo: “Los científicos y los juristas tienen que abandonar la idea de que pueda existir certeza absoluta en un proceso de identificación de forma completamente objetiva. Si se acepta que nada es ciento por ciento cierto, entonces se ve lógico que se determine el grado de confianza que puede tenerse sobre una particular creencia”. Y esto es así, porque una identificación auditiva está sujeta a una enorme variabilidad que depende, básicamente, de la familiaridad y naturaleza de la voz oída, de la habilidad intrínseca del auditor para procesar, registrar y recuperar o describir las voces, y de una multiplicidad de eventos específicos, medioambientales y de procedimiento (Clifford, 1980).

Un caso del contexto local donde se vio involucrado el testimonio de testigos auditivos fue durante el proceso del caso de Emilio Berkhoff (2015), el joven universitario que fue declarado culpable de los delitos de porte ilegal de armas y municiones y condenado a cinco años y un día de cárcel. Berkhoff fue vinculado a las acciones de la Coordinadora Arauco Malleco (CAM) y decidió declararse en rebeldía luego de que la Corte Suprema ratificara su condena. Mientras era buscado intensamente por la Policía de Investigaciones y Carabineros, a través de un comunicado público escrito desde la clandestinidad, Berkhoff denuncia irregularidades durante el proceso judicial: “Mi

---

<sup>2</sup> El testimonio consistió en gran parte en informes de rumores de tercera mano con testigos que afirmaban que otros les habían dicho que Hulet les había comentado que él había sido el verdugo, además del testimonio de un testigo, Richard Gittens, quien testificó que había servido en el ejército con Hulet durante más de 12 años, y que reconoció la voz de Hulet en el momento de la ejecución cuando Hulet pidió el indulto al rey antes de su decapitación. Aunque Hulet negó haber prestado servicios en la misma compañía que Gittens, y pese al hecho de que varios testigos adicionales declararon que el verdugo Gregory Byon, había confesado el asesinato, Hulet fue condenado por regicidio.

proceso judicial siempre estuvo plagado de irregularidades y vicios, desde testigos protegidos hasta testigos de oída, el debido proceso jamás existió y menos aún un juicio justo” (ver comunicado en: Anexo 1).

Fue debido al tenor de estas declaraciones que nos pareció oportuno revisar el procedimiento ocupado en la identificación auditiva, como una forma de contribuir a enfatizar la importancia de observar con cautela estos testimonios, teniendo en cuenta los errores que se han cometido en temas de identificación visual y auditiva (Broeders y Rietveld, 1995; Laub, 2010), como las dudas que se ciernen sobre la fiabilidad del reconocimiento que pueden realizar víctimas o testigos, ya sea por razones inherentes a sus propias habilidades auditivas, como por cuestiones procedimentales, a lo que se suma, en el ámbito nacional, la falta de una regulación específica en materia de identificación auditiva y los problemas que de esta situación se derivan en materia de los resguardos para las personas involucradas en delitos asociados a la voz <sup>3</sup>.

En los apartados que siguen se presentan los factores potenciales que ponen en riesgo la fiabilidad del testimonio auditivo para abordar de manera fundamentada el análisis específico de los procedimientos ocupados en la identificación auditiva en el caso de Emilio Berkhoff (2015).

## 1. FACTORES QUE IMPACTAN HIPOTÉTICAMENTE LA FIABILIDAD DEL TESTIMONIO AUDITIVO

La fiabilidad del testimonio auditivo puede verse amenazada de manera eventual por cualquier tipo de elemento, ente o circunstancia que forme parte del acto de habla. Y en ese sentido, podemos examinar ese potencial impacto en relación con el hablante, la señal, el oyente, las circunstancias acústicas y ambientales y, por supuesto, con los procedimientos que se ocupan en la recolección de la evidencia.

En relación con *el hablante*, es decir, con quien produjo la voz cuya identidad se intenta averiguar, normalmente interesará saber cómo era la voz: ¿sonaba peculiar?, ¿familiar?, ¿qué acento tenía? Incluso, podríamos preguntarnos si no se trataría acaso de

---

<sup>3</sup> El Código Procesal Penal de la República de Chile establece mediante diversas disposiciones la obligación del Ministerio Público de lograr la identificación de los intervinientes en el proceso penal, aunque sin hacer referencia a los medios por los cuales se llevaría a cabo la misión. La falta de una regulación amenaza la confiabilidad de la evidencia, dado que al no existir un protocolo que, de manera razonada, recoja los factores que pueden impactar en el testimonio de un auditor –o en la comparación de las propiedades auditivas cuando no hay registro de voz– ampliamente conocido por los jueces y abogados, es difícil que no se produzcan problemas en alguna parte del proceso, ya que los jueces no manejan toda la información y sería muy ingenuo pensar que pueden estar al día en todos los temas científicos. Al respecto, un estudio de Morales y Welsch (2011) sobre el procedimiento de reconocimiento de imputados en Chile y en derecho comparado señalan que la jurisprudencia analizada concuerda en que el reconocimiento de personas constituye una diligencia, en la cual pueden presentarse serias vulneraciones al debido proceso, al existir un espacio para la sugestividad inherente a su práctica, ya sea atribuible a la memoria del testigo como a la influencia de terceros.

una voz imitada, donde el hablante intentaba hacerse pasar por alguien conocido de la víctima. En cuanto a *la señal*, parece lógico suponer que escuchar un grito de un par de segundos coopera mucho menos en la tarea de identificación que oír una conversación telefónica de varios minutos. En otras palabras, podríamos suponer que es más confiable un testimonio que tuviera como base una conversación de varios minutos frente a otra que solo durara un par de segundos. La cuestión no es tan simple de resolver, porque depende de la interacción de diversos factores; sin embargo, es razonable, plantearse la pregunta acerca de cuál podría ser la duración adecuada para una exposición confiable. Por otra parte, con respecto a *los auditores -víctimas o testigos-* podríamos preguntarnos qué tanto podemos confiar en su capacidad auditiva, amparados en la constatación cotidiana de que hay mejores auditores que otros. Lo cierto es que los auditores pueden diferir de forma amplia en cuanto a su capacidad para reconocer voces y puede que esa variabilidad no solo dependa de sus habilidades naturales en sí, sino que se vea interferida por factores externos, como, por ejemplo, la experiencia en oír voces de un determinado tipo o el contar con algún entrenamiento especial en reconocimiento. Finalmente, *los procedimientos de recogida de la evidencia* deben tener en cuenta ante todo la frágil condición del testimonio al tipo de formato de prueba, y la construcción de la misma, como a los diversos elementos contextuales que rodean la identificación. En lo que sigue, ordenaremos la discusión de esos factores según el lugar que ocupan en la cadena comunicativa.

### 1.1. ¿CÓMO SONABA EL HABLANTE?

Podría darse el caso de que el testigo o víctima reconociera, en el acto, la voz del agresor, como la voz de una persona que ya conoce o que ha oído antes (“fue Pedro quien habló” o “fue la misma voz que escuché el otro día”), en cuyo caso la identificación ya se habría realizado y una prueba de identificación posterior no tendría sentido. Pero los casos de reconocimiento de voces familiares en casos forenses son poco frecuentes. En general, las víctimas o testigos no conocen quien habló, pero declaran que si volvieran a oír la voz del agresor desconocido podrían reconocerla. Estas dos situaciones encierran los conceptos de *familiaridad* y *tipicidad* de la voz escuchada, dos factores cuyo impacto en el reconocimiento del hablante interesa conocer.

1.1.1. LA FAMILIARIDAD. La familiaridad de las voces se puede entender de dos maneras: como la relación de parentesco o asociación previa al suceso entre el hablante y el oyente (p.e., en un caso de trabajo donde un mujer identifique al agresor como un antiguo novio); y también como la exposición previa *ad hoc* para estudiar la capacidad de reconocimiento del auditor, lo que se conoce como entrenamiento (p.e., el experimento de someter a un auditor a una muestra de determinadas características para probar si puede reconocer al hablante después de dos semanas). Este último concepto es el que está implícito en la mayoría de los trabajos sobre el tema.

Los estudios para probar la capacidad del oyente en relación con este factor son bastante variados, desde los más auspiciosos que informan porcentajes de precisión de voces familiares cercanas al 100% (p.e., LaRiviere, 1972; Hollien, Majewski, y Doherty, 1982) hasta aquellos que, como el de Ladefoged y Ladefoged (1980), en el extremo opuesto, reportan el fracaso de un auditor experto al no reconocer a su madre, en una tarea de audición de conjunto abierto, pasando por situaciones intermedias como las que informan Yarmey *et al.* (2001), quienes encontraron que voces con alta familiaridad fueron reconocidas con mayor precisión que voces con familiaridad baja, pero también, estudios como el de Read y Craik (1995) que reportaron que la precisión no fue mejor para voces medianamente familiares frente a voces no familiares.

1.1.2. LA TIPICIDAD. La tipicidad es un indicador que resulta, desde la perspectiva de quien oye -muchas veces de forma inconsciente- en una impresión general de rasgos comunes entre las voces oídas de que los hablantes forman parte de una misma comunidad. El concepto de tipicidad es relativo y depende del contexto de la población que es en la práctica, el conjunto de las voces que está habituado a escuchar un hablante particular. Sin embargo, dentro de una misma comunidad pueden existir voces distintivas; en esos casos, los patrones de habla o las propiedades acústicas de las voces son inusuales, atípicos en la población en general: no es lo que el oyente escucha de manera más frecuente. En este sentido, familiaridad y tipicidad son conceptos que están relacionados. Las voces atípicas pueden tener un origen más o menos inherente a la propia individualidad (como cuando algunos amigos o colegas de trabajo se escapan de la norma, porque tienen, p.e., un tono más alto para lo que se considera habitual en un hombre o más bajo para lo esperado, en el caso de una mujer, etc.); o bien, dialectal, por derivar de hablantes de comunidades diferentes (por ejemplo, norteos vs. sureños). Estos hablantes suelen ser más fáciles de identificar que los que tienen voces típicas. A su vez, si se seleccionan dos hablantes al azar dentro de la población, es más probable que ambos tengan -en términos relativos- voces típicas más que voces atípicas y éstas sonarán similares entre sí, no porque las voces sean producidas por el mismo hablante, sino, simplemente, porque los dos hablantes tienen voces comunes (Morrison, 2018, pp. 92-93).

Una situación que permite ilustrar mejor el concepto de tipicidad puede ser la siguiente: entre la población de periodistas de televisión chilenos, Amaro Gómez-Pablos, tiene una voz muy atípica, ya que cecea<sup>4</sup>, en contra de la norma general hispanoamericana, no distinguidora. Por lo tanto, la voz de Amaro Gómez-Pablos sería

---

<sup>4</sup> Amaro Gómez-Pablos Benavides, conocido periodista chileno-español, es ceceante, pero no lleísta. Un ejemplo de su emisión oral puede oírse en *Chile Profundo*, programa de Megavisión (2018). Disponible en: [https://www.youtube.com/watch?v=kD4eUnG\\_v40](https://www.youtube.com/watch?v=kD4eUnG_v40)

muy fácil de identificar en una selección de grabaciones de locutores chilenos. En contraste, las voces de los periodistas chilenos con voces típicas, que hablan español de Chile y no son ceceantes, serían más difíciles de identificar. Ahora, si cambiamos la población de TV chilena por la de TV española, entonces la voz de Amaro Gómez Pablos se vuelve más típica y será más difícil identificarlo en una selección de grabaciones de audio con ceceo. El punto que interesa destacar es el de que por pertenecer a una determinada comunidad de hablantes estamos más familiarizados con determinados acentos frente a otros y *podremos distinguir sutiles cambios dentro de nuestra comunidad, pero no de otras comunidades más alejadas*. Por ejemplo, para un chilote puede resultar inequívoco distinguir un hablante de la isla frente a otro del sur (p.e., Valdivia y Osorno). En cambio, probablemente, tendrá que esforzarse para diferenciar un hablante del centro o norte de Chile.

La situación anterior enlaza con el multilingüismo que es una característica común en la vida urbana, y, en muchos lugares, oír un hablante que tiene un acento distinto o que no habla la lengua local no llama hoy día la atención. En situaciones forenses, es razonable preguntarse en qué medida las diferencias de acento o lengua entre víctima y agresor afectan la fiabilidad de las identificaciones auditivas. Un auditor no familiarizado con el inventario fonético de una lengua particular podría no ser sensible a pequeños detalles idiosincráticos de esa lengua presentes en el hablante objetivo y no ser capaz de diferenciar entre los hablantes que hablan dicha lengua, pudiendo incluso percibir que todos hablan de la misma manera. Al respecto, Rosas, Sommerhoff y Sáez (2018) realizaron una prueba en la que auditores chilenos no entrenados y sin manejo del alemán fueron expuestos a la voz de una mujer alemana que habló en español y luego de dos horas la tuvieron que reconocer dentro de una serie de voces de mujeres no alemanas que hablaron en alemán. Los auditores identificaron la voz original, pero con poca certeza.

Otra cuestión que deriva de la tipicidad es qué tan robusto puede resultar el recuerdo de una voz peculiar al transcurso del tiempo. La asociación entre lo que es distintivo y lo reconocible no es sencilla. Orchard y Yarmey (1995) reportaron que las voces que eran consideradas “distintivas” fueron recordadas mejor para dos días después que las voces “no distintivas” y Clifford, Rathborn y Bull (1981) encontraron que voces no familiares tasadas como “altamente reconocibles” no fueron mejor reconocidas que aquellas que fueron tasadas como “difíciles de reconocer” (para un retraso entre aprendizaje y prueba de 10 minutos a 14 días). De manera similar, para retrasos de una a cuatro semanas, Papcun, Kreiman, y Davis (1989) no encontraron una diferencia importante entre las voces calificadas como “difíciles de recordar” versus las voces “fáciles de recordar”. Así, las voces tachadas, a priori, de distintivas no son necesariamente fáciles de recordar.

1.1.3. DISFRACES E IMITACIONES. No es ilógico suponer que el victimario intentara ocultar su voz, para evitar que una potencial familiaridad con la víctima lo pudiera delatar. Pero,

a pesar de que la experiencia diaria y el sentido común apoyan su efectividad, las voces disfrazadas no son, por lo general, comunes en casos criminales que incluyen reconocimiento de voz. Al respecto, Kunzel (1994) estimó que solo sobre un 15% de casos en las Cortes de Alemania involucraron voces disfrazadas. En cuanto a los disfraces, los más utilizados que se han encontrado son la voz de falsete junto con la voz rota, los falsos acentos extranjeros y las hiponasalizaciones (producidas por pinzamiento de la nariz); mientras que el habla sintetizada es el menos frecuente (Reich y Duke, 1979; Kunzel, 2000). En cuanto al reconocimiento, los estudios muestran que los auditores tienen pocas dificultades para detectar una voz disfrazada, pero, en cambio, tienen dificultades para reconocer el patrón vocal subyacente (Reich, 1981).

No todos los cambios de patrones son intencionales; de hecho, muchos cambios de la voz son totalmente incontrolados e inconscientes. Es lo que ocurre, por ejemplo, con las variaciones de la voz que introducen los diferentes estados de ánimo o las alternancias de registros que pueden igualar, y hasta exceder, los efectos de la voz disfrazada. Bahr y Pass (1996) grabaron las voces de hombres afroamericanos en estilo coloquial y formal y preguntaron a los auditores que juzgaran si los pares de voces eran de la misma o distintas personas. Cuando las muestras de un mismo locutor compartían el estilo, las tasas de correctas fueron altas (87% de aciertos y 13% de falsa alarma). En cambio, cuando los estilos se cambiaron, las tasas disminuyeron (39% de aciertos y con 21% de falsa alarma).

1.1.4. LA IMITACIÓN. Podríamos imaginar también una situación en la que quizá le interesara al agresor involucrar a otro, intentando imitar su patrón de voz. Se trata de una tarea que pone en acción otras estrategias vocales. Mientras que una voz disfrazada requiere que los hablantes oscurezcan sus patrones de voz subyacentes, los imitadores reproducen, o incluso, exageran las características vocálicas más importantes de un hablante con el detalle suficiente como para que los oyentes de forma inmediata las asocien a la voz que se intenta reproducir; se trata, en verdad, de una caricatura verbal (Blomberg, Elenius, y Zetterholm, 2004). Si bien, tanto las voces disfrazadas como las imitaciones de voces requieren de la plasticidad del hablante, un punto que interesa resaltar es el hecho de que la imitación a diferencia del disfraz requiere de una audiencia que espera escuchar la voz objetivo. El éxito de muchos doblajes radica en ese hecho. Es bien sabido el caso de Rich Little, quien dobló el diálogo de David Niven cuando su habla se hizo ininteligible, por la enfermedad que padecía, poco antes de que fuera realizada *The Curse of the Pink Panther* (Edwards, 1983) y, por supuesto, nadie percató de ello hasta que no fue dado a conocer públicamente (Kreiman y Sidtis, 2011).

## 1.2. ¿CÓMO ERA LA VOZ OÍDA?

Al menos, hay tres cuestiones básicas que parece relevante atender: la duración del material (la cantidad de segmentos articulados por el agresor), el registro o estilo de la emisión (diferentes grados de formalidad o informalidad que pueden darse en la interacción) y la calidad de la señal (grado de inteligibilidad de la palabra).

1.2.1. LA DURACIÓN. Parece razonable esperar un mejor desempeño a medida que se incrementa la cantidad de material disponible. Por ejemplo, Ladefoged y Ladefoged (1980) y Rose y Duncan (1995) informaron tasas de identificación correcta para voces familiares que oscilaron entre el 31% para palabras sueltas y el 95% para segmentos de oraciones (véase también: Solan y Tiersma, 2003, pp. 397–399). Sin embargo, es posible que esa expectativa requiera algún matiz. De hecho, en un experimento de identificación de locutores, Bull y Clifford (1984) encontraron que el número de expresiones escuchadas provenientes del locutor objetivo no tenían casi ningún efecto en el rendimiento de los oyentes y, de acuerdo con Bricker y Pruzansky (1966), no es tanto la longitud del material lo que importa, sino el grado en que el material escuchado por el oyente es representativo del repertorio del locutor.

1.2.2. EL REGISTRO. Las diferencias inherentes a la situación del crimen y del interrogatorio dan lugar a la posibilidad de que el estilo de discurso también cambie. Bahr y Pass (1996) evaluaron la influencia de la alternancia de estilos en la identificación del hablante, mediante un experimento donde los locutores aparecían en pares de estilos similares y diferentes y evidenciaron que los locutores en estilos distintos estuvieron más propensos a ser etiquetados como personas diferentes que cuando se compararon dentro de la misma condición, por lo que concluyen que el cambio de estilo, inducido por los cambios en la situación y el entorno social de la grabación, puede dar lugar a la identificación errónea de un hablante en particular. Öhman, Eriksson, Granhagm (2013), en un experimento en el que la exposición original era una voz enojada, mientras que los locutores de la rueda de reconocimiento realizada, a continuación, hablaban con voces normales, encontraron que los oyentes adultos nunca identificaron correctamente al hablante objetivo.

1.2.3. LA CALIDAD DE LA MUESTRA. La pérdida de información acústica debida a factores como el ruido de fondo, la baja calidad de la grabación y/o de la reproducción, las distorsiones debidas a la transmisión de la voz a través de un sistema telefónico y la discrepancia entre la calidad de las grabaciones de los hablantes conocidos y cuestionados, también tienden a reducir las tasas de identificación correcta (Rathborn, Bull, Clifford, 1981; Bull y Clifford, 1984; Rose, 2002, p 102). Nolan, McDougall, Hudson (2013) encontraron que cuando se presentaban grabaciones de pares de hablantes con calidad telefónica, se percibían más similares que cuando se presentaban grabaciones de buena calidad (no telefónica) de los mismos pares de locutores. Y cuando un miembro

de la pareja era de buena calidad y el otro miembro era de calidad telefónica, se percibía que eran más diferentes.

### 1.3. ¿CUÁNTO PODEMOS CREER EN EL TESTIMONIO DEL AUDITOR?

En relación con el auditor, hay dos aspectos de interés: la experticia en temas de identificación auditiva por parte del oyente y la confianza que manifiesta el auditor al realizar la tarea de identificación. Interesa saber si estos indicadores tienen algún peso en la confiabilidad del resultado de selección o no; es decir, si la experticia y la certeza pueden servir de referencia para evaluar la validez del testimonio auditivo en algún grado.

1.3.1. LA EXPERTICIA. En general, víctimas o testigos son personas, sin conocimientos especiales en reconocimiento de voces con fines forenses y eso plantea saber si su testimonio puede ser o no confiable. Hollien (1990) muestra la diversidad de posturas sobre el valor que se le puede otorgar al testimonio de un experto frente a otro no lo es. Clifford (1980) sostiene que el sistema de justicia criminal debería aplicar la mayor reserva posible cuando utiliza la identificación de voz en cualquier tipo de acusación. Para Deffenbacher *et al.* (1989), la identificación de testigos auditivos nunca debiera considerarse como la única prueba sobre la cual se determine la identidad. Más recientemente, Morrison, Enzinger y Zhang (2018, p. 86) destacan que, en el reconocimiento por no expertos, con frecuencia, los auditores, solo seleccionan la voz que ellos piensan pertenece al delincuente, pero tienen dificultades al identificar y/o describir las propiedades de la voz particular seleccionada que les permiten distinguirla de las otras voces. Por el contrario, la comparación forense de la voz es idealmente realizada por profesionales forenses –con amplio entrenamiento en fonética o procesamiento del habla y entrenamiento en evaluación de la evidencia forense, a menudo con un doctorado en un campo relevante– que trabajan dentro de un paradigma científico que hace uso de una base de datos relevante, mediciones cuantitativas, modelos estadísticos para cuantificar la fuerza de la evidencia, y que prueban de forma empírica la validez y confiabilidad de su sistema bajo condiciones que reflejan las condiciones del caso bajo investigación (Morrison, 2011).

1.3.2. LA CONFIANZA. Uno podría sentirse inclinado a dar más crédito a la elección de un oyente que dice estar “100% seguro” frente a otro que manifiesta dudas respecto de su selección. Por desgracia, no es una cuestión simple predecir qué voces serán realmente reconocidas con precisión (Bull y Clifford, 1999, pp. 217-218). Los estudios demuestran que los auditores no son buenos jueces prediciendo su reconocimiento. Por ejemplo, Van Lancker, Kreiman, y Emmorey (1985) encontraron que los auditores solo identificaron correctamente un 26,6% de aquellas voces de famosos que habían dicho que podrían

reconocer. También, Saslove y Yarmey (1980) encontraron una correlación bastante baja entre el nivel de confianza y la precisión de la identificación. En forma más reciente, Sarwar, Allwood y Zetterholm (2014) encontraron que los oyentes estaban demasiado confiados en su identificación, es decir, sus niveles de certeza informados en la exactitud de sus identificaciones eran mayores que su tasa de identificación correcta real. Clifford, Bull y Rathbom (1980, citado en Bull y Clifford, 1999, p. 218) y Bull y Clifford (1984, pp. 121-123) encontraron una correlación positiva entre la certeza de los oyentes individuales y sus tasas de identificación correcta, pero solo en los casos en que la voz objetivo se incluyó en la alineación. En un estudio de Van Wallendael, Surace, Hall-Parsons y Brown (1994) se encontró que cuando el locutor objetivo no estaba incluido en la alineación, los oyentes siempre identificaron incorrectamente a uno de los locutores en la serie como el objetivo. Que los oyentes elijan con confianza uno de los hablantes en la alineación cuando el objetivo no está presente es bastante preocupante. Esto extrapolado a la situación de un caso real equivale a seleccionar al sospechoso cuando en realidad no es el delincuente (Zetterholm, Sarwar, Thorvaldsson, Allwood, 2012). También se ha observado una falta de correlación en el sentido inverso, por ejemplo, Zetterholm *et al.* (2012) encontraron que los oyentes no tenían confianza suficiente en su afirmación de que el locutor objetivo no estaba en la rueda de reconocimiento cuando, de hecho, no estaba presente. El relativo peso que puede tener la certeza del oyente en su identificación de la voz cuestionada obliga a pensar que para que su testimonio sea en realidad valioso para el tribunal, el oyente tendría que ser examinado para encontrar la relación entre su grado de certeza y su tasa de identificación correcta. (Morrison *et al.*, 2018, p. 412)

#### 1.4. ¿CÓMO SERÍA UN BUEN PROCEDIMIENTO?

Básicamente, nos interesan cuestiones como: ¿es posible a la vista de los estudios revisados usar como medio de prueba válido un reconocimiento auditivo por víctimas o testigos?, si lo es, ¿bajo qué condiciones?, y ¿cómo debiera llevarse a cabo este objetivo?

Si la única pieza de evidencia es el testimonio de oído –y no existe una muestra tangible de la voz cuestionada como para realizar una comparación forense por expertos– esta podría ser útil, pero los jueces deberían estar informados de sus limitaciones (como las descritas en los apartados anteriores). Así, cuando no fuere posible realizar una comparación forense de la voz, se le puede pedir a un oyente que escuche una alineación de voces.

1.4.1. TIPOS DE FORMATO. La identificación de personas mediante testigos se trata de un procedimiento que puede adoptar dos tipos de formatos, *la confrontación simple*, donde el testigo y/o víctima es expuesto simplemente a una sola voz y la *confrontación múltiple*, más conocida como *rueda de reconocimiento de voz*, *desfile de voces* o *alineación de voces*, donde el oyente es expuesto a una secuencia de varias voces. El primer formato ha sido ampliamente criticado por considerarse una práctica intrínsecamente sugestiva

(Solan y Tiersma, 2003, pp. 381-82, 390-393, 427-428). Aquí es donde se escucha una grabación de una sola persona, o se escucha a esa persona en vivo, y esa persona es el sospechoso. Se considera sugestivo, porque es razonable que el testigo deduzca que la razón por la que se le expone una sola voz es que la policía cree que esa voz particular pertenece al malhechor, y lo que la policía requiere es que el testigo haga una confirmación.

En cuanto al formato aceptado, de manera más extendida, *la rueda de reconocimiento*, en él, por lo general, el testigo describe primero la voz del agresor al policía investigador o a un experto forense, en términos tales como: tenía un acento del sur, voz masculina, voz profunda, etc. A continuación, un experto forense prepara uno o más conjuntos de grabaciones de audio. Cada conjunto consta de muestras de hablantes que coinciden con la descripción dada por el testigo y pueden incluir, o no, una grabación del sospechoso. El contenido lingüístico, el estilo de habla o las condiciones de grabación no deben hacer que el sospechoso se destaque en comparación con las demás muestras de hablantes.

Un punto muy relevante en la construcción de las ruedas de reconocimiento es la eventual incompatibilidad entre el perfil del sospechoso y el proporcionado por el oyente. Al respecto, Broeders y Rietveld (1995, p. 36) consideran que una eventual discrepancia de perfiles, debiera conducir a su cancelación, Morrison *et al.* (2018, p.88), en cambio, sostienen que las muestras deben construirse siguiendo ambos modelos; por ejemplo, si el sospechoso tiene un acento de Chiloé, pero el testigo describió un acento de Temuco, al menos una proporción sustancial de las muestras de los hablantes deben sonar como chilote). Consideramos que es más apropiada esta solución, dado que las diferencias de perfiles no solo pueden producirse por una limitación en la capacidad terminológica para describir sonidos particulares, por parte de la víctima o testigo, sino también, por la existencia de más de un testigo y las eventuales subsecuentes diferencias en la percepción y descripción de detalles sobre la voz oída.

1.4.2. EL SESGO DE LA IDENTIFICACIÓN POSITIVA PARA COOPERAR CON LA INVESTIGACIÓN. Si un oyente espera escuchar una voz en particular, entonces es más probable que identifique la voz que escucha como la voz que espera escuchar (Rose, 2002, p. 104). Ladefoged y Ladefoged (citado por Morrison *et al.* 2018, pp. 95-96) describieron el caso de una identificación errónea motivada por expectativas previas: un narcotraficante, Kalkin, arrendó una habitación de hotel. Los agentes de narcóticos llamaron por teléfono a la habitación del hotel donde esperaban que atendiera Kalkin y sostuvieron una conversación telefónica con quien contestó, pensando que se trataba de Kalkin. La defensa pudo demostrar que el Sr. Kalkin no estaba en la habitación en ese momento, y un socio suyo fue la persona que habló con los agentes de narcóticos.

1.4.3. EL INTERVALO DE TIEMPO. El intervalo de tiempo es otro factor a considerar en la aplicación de un test auditivo. En general, uno esperaría que la precisión de la identificación de voz de los auditores de oído disminuyera a medida que aumenta el intervalo de tiempo entre la exposición a la voz desconocida y la audición de la rueda. El primer trabajo que recoge este interés es el de McGehee (1937) donde se evalúa la capacidad de recordar voces no familiares con distintos intervalos de tiempo, basado en el caso real de *United States versus Hauptmann* de 1935. McGehee encontró que la precisión de la identificación fue del 83% a los dos días, del 68% a las dos semanas, del 35% a los tres meses y del 13% a los cinco meses (Clifford, 1980). Los resultados de la investigación son algo variados, por ejemplo, Öhman *et al.* (2013) encontraron que tanto los adultos como los niños de 11 a 13 años de edad tenían un 19% y un 25% de tasas de identificación correctas inmediatamente después de la exposición, pero las tasas de identificación correctas tenían una probabilidad menor después de un retraso de dos semanas (la probabilidad era de 12.5%). Otros estudios sugieren poca disminución durante largos períodos de tiempo. Por ejemplo, Kerstholt, Jansen, van Amelsvoort y Broeders (2006) encontraron tasas de identificación correctas de 24 a 32% sin diferencias significativas entre un intervalo de tres semanas y un intervalo de ocho semanas.

1.4.4. LA VALIDEZ. A modo casi de cierre de esta primera parte, dado que hemos comprobado una variación idiosincrásica sustancial entre interlocutores, además de no existir correlación objetiva entre la certeza y acierto en la identificación de un oyente, el testimonio auditivo por legos no tiene ningún peso, a menos que el oyente individual pueda ser probado en circunstancias similares a aquellas en las que identificó la voz cuestionada. Morrison *et al.* (2018, pp. 89-90) refieren un procedimiento donde se propone que el oyente tendría que identificar un gran número de muestras de voz para poder obtener su tasa de identificación correcta (Morrison *et al.* 2018, sección 99.290) y que para que una prueba de este tipo sea significativa para el tribunal, tendría que realizarse en condiciones similares a las condiciones en las que el oyente hizo la identificación de la voz interrogada. La cantidad de muestras de voz que tendrían que identificarse dependería de un compromiso entre el nivel deseado de precisión y la viabilidad, por ejemplo, 10 sería muy práctico, pero probablemente no tendría la suficiente precisión y 1000 probablemente satisfaría las preocupaciones de todos sobre la precisión, pero sería poco práctico.

## 2. LA IDENTIFICACIÓN DE EMILIO BERKHOFF POR TESTIGOS AUDITIVOS (2015)

En base a los factores que potencialmente inciden en el reconocimiento auditivo y la información proporcionada por la Defensoría Penal Pública de la Región del Biobío sobre algunos hechos que le fueron imputados a Emilio Berkhoff Jerez y las respectivas declaraciones de los testigos, se discute la fiabilidad de los testimonios de oído que

constaban en la carpeta investigativa y se proporcionan algunas sugerencias sobre lo que consideramos –a la luz de todo lo expuesto en los apartados anteriores– habría sido un procedimiento adecuado de ocupar. Previo a lo anterior, se proporcionan unas notas sobre el peso que tuvieron las declaraciones de los(las) testigos auditivos en la sentencia (ver sentencia en Anexo 2.).

## 2.1. VALORACIÓN DE LOS TESTIMONIOS AUDITIVOS PRESENTADOS

Todos los testigos que aluden a las características de voz del sospechoso eran además testigos “secretos o con reserva de identidad” (en adelante, TRI). El testimonio de estos testigos lo era en relación con múltiples delitos de robo con violencia e incendio atribuidos al imputado y por los cuales fue absuelto. El ministerio público se negó a revelar la identidad de estos testigos, a pesar de que la defensa obtuvo una decisión de los tribunales en el sentido de dar a conocer dicha identidad, en definitiva, no fueron presentados al juicio y sus testimonios se introdujeron por los policías que los habían escuchado o les habían tomado las declaraciones. Estos testimonios fueron considerados insuficientes y de ahí que el imputado fue absuelto de los delitos de robo con violencia e incendio. Sin embargo, a pesar de la absolución, los testimonios de los testigos secretos, fueron relevantes y la base para mantener en prisión preventiva al imputado durante la fase de investigación y para sustentar la acusación que presentó el ministerio público.

## 2.2. TESTIMONIOS Y TESTIGOS

2.2.1. GENERALIDADES. Hubo seis situaciones delictuales en las que se involucró a Emilio Berkhoff como autor principal de una banda entre 2012 y 2013. Los hechos suceden en sectores de las comunas de Tirúa y Cañete en la provincia de Arauco, IX Región.

Se presenta un extracto de las declaraciones de los testigos, seguido de una evaluación de los testimonios –en esencia, el perfil del hablante– en relación con las características relevantes de los testigos y las condiciones de identificación. De los 14 testigos que figuraban en la carpeta investigativa<sup>5</sup>, analizamos solo las declaraciones de los testigos que proporcionaron indicaciones sobre la voz: TRI N° 1, TRI N° 2, TRI N° 3, TRI N° 6, TRI N° 7, TRI N° 8, TRI N° 9, TRI N° 10 y TRI N° 13.

### 2.2.1.1. SITUACIÓN N° 1 (14-06-2012)

Esta situación involucró a 5 testigos: TRI N° 1, TRI N° 2, TRI N° 3, TRI N° 6 y TRI N° 7 (no mapuches).

---

<sup>5</sup> Información proporcionada por la Defensa de acuerdo a los antecedentes que constaban en la carpeta investigativa.

TRIN°1:

“el acento era distinto hablaba claro”, “Lo reconocería por su voz” (D 2) [1 mes después].

“Se notaba educado que sabía hablar [...]” “Lo reconoció a penas lo vio en la TV.” (D 3) [8 meses después].

TRIN° 2:

“por su voz creo que tenía alrededor de 25 a 30 años, en su forma de hablar se notaba que sabía o que tenía que decir” (D 2) [1 mes después].

“Hablaban en forma enérgica, al principio con groserías y en forma intimidante...Se notaba en su forma de hablar que tenía educación y sabía lo que hacía, tenía don de mando. Su voz no era aguda ni grave, diría termino medio. “[...]no era mapuche porque su forma de hablar no indicaba que fuera mapuche ya que estos hablan con un acento que les cuesta hablar en español, pero este sujeto hablaba claramente” (D 3) [ca. 10 meses después].

TRIN° 3:

“Me llamó mucho la atención su forma de hablar ya que lo hacía de manera clara y buen vocabulario atreviéndome a decir que no es mapuche ya que estos hablan distinto y su voz es más ronca” (D 2) [1 mes después].

TRIN° 6:

“su “marido” (TRIN°1) le habría comentado que el que daba las instrucciones a los encapuchados” [...] "por su forma de hablar no era mapuche" Cuando aparece en Tv (Emilio) su esposo inmediatamente le señala que era el autor” (D 1) [ca. 9 meses después].

TRIN° 7:

“su padre (TRIN° 1) le habría dicho “ la persona que lideraba el grupo era bien hablado, por su vocabulario no era mapuche". Cuando ve a Emilio por TV “mi padre me comentó inmediatamente que era el sujeto del atentado” (D 1).

2.2.1.2. SITUACIÓN N° 2 (24-07-2012 02:40 HORAS).

Esta situación involucró solo a un testigo: TRIN° 8 (mapuche).

TRIN° 8:

“no hablaban mapudungën y por su voz se notaba que no eran del sector” (D 2) [ca. 1 mes].

“le escucho decir en voz alta “apúrense” (D 3) [7 meses].

2.2.1.3. SITUACIÓN N° 3 (24-07-2012 02:55 HORAS).

Esta situación involucró dos testigos: TRI N°9 (no mapuche) y TRI N° 10 (mapuche).

TRI N° 9:

“una voz, aparentemente de una persona sexo masculino, fuerte y clara que gritaba.” (D 2) [ca. 1 mes después].

“Su hablar no era típico de una persona mapuche ni campesino, sino de una persona común y corriente” (D 3) [6 meses después].

“voz normal, ni grave ni aguda, no enredada, educado por la forma de hablar” (D 4) [7 meses después].

TRI N° 10:

“Dos o tres de ellos no eran mapuches” [...] “y forma de hablar no corresponden a las características de los mapuches” (D 2) [ca. 1 mes después].

Pero en una testificación posterior entra en contradicción:

“Había uno que hablaba”, “era el único que hablaba por eso le dije a la PDI que él daba las ordenes...” “Parecía que estas personas le hablaban... pero no pude escuchar que le decían” (D 3).

2.2.1.4. SITUACIÓN N° 4 (24-07-2012 04:15 HORAS).

Sin indicaciones sobre la voz.

2.2.1.5. SITUACIÓN N° 5 (07-09-2012 CA. 16:30 HORAS).

Esta situación involucró solo un testigo: TRI N° 13 (de procedencia dialectal no especificada).

TRI N° 13:

“Lo insultan con palabras groseras” “3 individuos” (D 1) [mismo día].

“no era mapuche del sector, debido a su forma de hablar, su acento” [...] (D 2) [ca. 2 ½ meses].

“en su forma de hablar era claro yo diría que era una voz normal, no era mapuche a mi parecer... me trataba con groserías” (D 3). “Es la misma persona que vi en la TV No tienes dudas que es él aunque solo vio la mitad de su rostro” [ca. 3 ½ meses].

2.2.1.6. SITUACIÓN N° 6 (25-01-2013).

Esta situación involucró solo un testigo: TRI N° 9 (no mapuche). Este testigo también declaró por los hechos de la SITUACIÓN N° 3 (junto a TRI N° 10).

(TRI N° 9)

“Le habrían dicho que eran los mismos del atentado a su casa el 24/7/12..” “era el líder” “voz normal, ni grave ni aguda, clara no enredada, notaba educado por la forma de hablar” (D 3) [ca. 1 mes].

“voz normal, ni grave ni aguda, no enredada, educado por la forma de hablar” (D 4) [7 meses después].

### 2.3. LA EXPERTICIA

Una condición general que cruza a todos los testigos es su falta de experticia en temas de identificación del hablante. No es un hecho anormal, casi por regla, los testigos o víctimas son personas profanas en asuntos de la voz. Tampoco invalida, a priori, esta característica la utilidad que pudieran llegar a tener los testimonios, pero conviene tenerla presente. Para dimensionar su alcance, pensemos cómo cualquier hablante nativo de español de Chile, podrá distinguir, describir e incluso, reproducir las diferencias que percibe cuando escucha hablar en la calle o en la televisión a un ciudadano extranjero, como en el caso de un español cuando dice, por ejemplo: *caza*, *zapato* o *mantequilla*, *castellano*, porque la interdental sorda y la palatal lateral, en el primer y segundo caso, respectivamente, constituyen diferencias sobresalientes entre la variedad peninsular frente a la chilena e hispanoamericana más amplia; sin embargo, por lo general, este hablante tendrá problemas con otras distinciones más finas. Es probable que pase por alto sonidos como, por ejemplo, las diferencias entre un sonido alveolar fricativo frente a otro apicoalveolar fricativo, incluso percibiéndolos no será capaz reproducir ni describir con precisión la diferencia. Este es un aspecto sensible a la hora de construir el perfil del delincuente. Sin duda, demandará de parte del experto una habilidad especial para traducir a términos técnicos los rasgos generales y ambiguos de la caracterización del oyente no especialista. Qué interpretación dar, p.e., a las expresiones “se notaba educado que sabía hablar [...]” (TRI N° 1) o “en su forma de hablar se notaba que sabía o que tenía que decir.: (sic)” (TRI N°2). Se trata de una tarea que no es simple, pero una discusión de las posibilidades y limitaciones de este aspecto iría claramente más allá del alcance de este artículo.

### 2.4. TESTIGOS INDIRECTOS

No existe ningún argumento lógico por el cual los testimonios de testigos auditivos *indirectos* pudieran ser considerados para apoyar alguna hipótesis (acusatoria o contraria). Una declaración de personas ajenas al suceso solo podría tener cabida en el supuesto caso de existir un registro de la voz desconocida (del agresor) con la cual comparar las propiedades cuantificables de la muestra de voz del imputado, y solo si la persona que declara es un experto en comparación forense de la voz, lo que se traduce en un amplio entrenamiento en fonética o procesamiento del habla y en evaluación de

evidencia forense, además del trabajo dentro del nuevo paradigma en las ciencias forenses que se caracteriza por el uso de una base de datos relevante, mediciones cuantitativas, y modelos estadísticos para cuantificar la fuerza de la evidencia (Morrison, 2011; Morrison et al., 2018:86-87). Así, un error que se hace evidente es la presentación de los testimonios de TRI N° 6 y TRI N° 7.

## 2.5. EL SESGO EN LA IDENTIFICACIÓN

Observamos, al menos, tres tipos de sesgo que tienen que ver con una (1) eventual acción de los medios de información (televisión, principalmente), (2) interacción de los testigos entre sí y (3) cooperación de los testigos.

2.5.1. ACCIÓN DE LOS MEDIOS DE INFORMACIÓN. Una identificación posterior por medio de las noticias de la TV no es lo más apropiado, porque constituye una identificación sesgada, dado que el contexto colabora en hacer suponer que el detenido es el agresor (si la Policía ha detenido a alguien hay razones lógicas para suponer que se trata del delincuente). Por otra parte, una situación en la que la víctima o testigo es expuesto a un video de TV (noticias, programa, etc.) podría compararse, en la práctica, con el formato de identificación de a uno (que no da ninguna salvaguarda al imputado) al quedar focalizada la atención solo en un sujeto. Sumado a lo anterior, una respuesta positiva, en este contexto, se puede explicar también como un intento por parte del testigo de querer colaborar con la investigación. De modo que, una declaración como la que proporcionan TRI N° 13 (en D 3) y TRI N° 3 (en D 3) no pueden aceptarse si estos reparos.

Por otra parte, con excepción de TRI N° 13, todos los testigos declaran aspectos sobre la voz del agresor con un desfase de 1 mes o de casi 1 mes, al menos. El desfase en sí no indica nada, pero unido a una eventual contaminación de información recibida de los medios y/o de interacción de los testigos entre sí, podría explicar el surgimiento de características sobre la voz desconocida.

2.5.2. INTERACCIÓN DE LOS TESTIGOS ENTRE SÍ. En situaciones donde convergen más de un testigo habría que tomar todos los resguardos necesarios para evitar que pudiera darse una contaminación de versiones por el contacto interpersonal entre los testigos. Si esto no fuera posible, al menos, habría que tener presente que una eventual coincidencia de caracterizaciones dadas en las testificaciones podría explicarse por esta condición, tal como ocurre en el caso de las declaraciones de TRI N° 1, TRI N° 2 y TRI N° 3, que participan de los mismos hechos y utilizan expresiones similares para referirse al perfil de la voz desconocida. Compare: "el acento era distinto hablaba claro" (TRI N° 1, D2), "este sujeto hablaba claramente" (TRI N° 2, D3) y "me llamó mucho la atención su forma de hablar ya que lo hacía de manera clara" (TRI N° 3, D3); o como también ocurre con

de TRIN° 9 y TRIN° 10 que también comparten una situación. Compare: “Su hablar no era típico de una persona mapuche ni campesino” TRI N° 9, D3 [y] "y forma de hablar no corresponden a las características de los mapuches" (TRIN° 10, D 2).

2.5.3. COOPERACIÓN DE LOS TESTIGOS. Imaginemos que una víctima o testigo debe declarar por un hecho A en tres ocasiones que se suceden en el tiempo y que aun en una cuarta ocasión debe hacerlo nuevamente, junto con declarar por un hecho B, distinto. Si a las anteriores circunstancias se suma el hecho de que en todas las declaraciones previas de A no se indican detalles sobre la voz del delincuente en contraste con la última declaración donde sí aparecen, siendo coincidentes, además, con los rasgos declarados por B, que se parecen a los del imputado, existen razones para pensar que la generación del parecido entre la voz del imputado y la voz del agresor por parte del testigo pueda estar motivada por un intento de este último de cooperar con la investigación más que por la calidad de su memoria o capacidad de reconocimiento auditivo. Esto es justo lo que sucedió con un par de testificaciones de TRI N° 9, donde la coincidencia del contenido de las declaraciones del 20-02-13 al testificar por los hechos del 25-01-13 (SIT 6) y, por cuarta vez, sobre los hechos del 24-07-2012 (SITUACIÓN 3) podrían tomarse como una declaración sesgada por el intento de colaborar con la investigación.

## 2.6. LAS RUEDAS DE RECONOCIMIENTO Y LAS PRUEBAS DE VALIDEZ DE LOS TESTIGOS

Aunque no hubo una construcción de ruedas de reconocimiento, en caso de haber existido tales pruebas, consideramos necesario tener en cuenta: (1) el perfil general del agresor, derivado de los testimonios, como la condición de (2) varios locutores, presentes en algunas situaciones, y la necesidad de considerar (3) las pruebas de validez, para medir la capacidad de reconocimiento de cada testigo.

2.6.1. EL PERFIL GENERAL DEL AGRESOR. La rueda de voz puede ser ocupada cuando una comparación forense de voz (realizada por expertos forenses) no es posible. Hay varios protocolos, que, en general, con diferencias de opinión con respecto al número y tipo de muestras que la serie de voces debe contener, coinciden en el hecho de que un experto debe construir un perfil del delincuente en base a la descripción del sospechoso y que el sospechoso no debe salir de la alineación. Ante una eventual discrepancia de perfiles entre la descripción del testigo y la caracterización del sospechoso o entre diferentes descripciones de testigos, Morrison *et al.* (2018, p.88), sostienen que las muestras deben construirse siguiendo los modelos del sospechoso y del delincuente; por ejemplo, si el sospechoso tiene un acento de Chiloé pero el testigo describió un acento de Temuco, al menos una proporción sustancial de las muestras de los hablantes deben sonar como un hablante chilote. En el presente caso no hay discrepancias de perfil, sino todo lo contrario, una tendencia general a caracterizar la voz dubitada como “no mapuche”, con algunas

pocas excepciones que añaden otras características (referidas, p.e., al tono, la calidad vocal, la edad y la educación). Pero una caracterización así es demasiado amplia como para atribuirle la autoría a un hablante específico –que calza en ella– entre todos los habitantes de la población de Chile o solo de la novena región que “no son mapuches” (o incluso, no mapuches “educados”, “que hablan claro”, “que tienen entre 25 y 30 años de edad” como aparece en otras descripciones) y que podrían incluirse en ese conjunto. En otras palabras, la sola coincidencia de un rasgo dialectal como el enunciado no determina, de forma inequívoca, por sí misma la correspondencia entre el imputado y la voz dubitada.

La escasez de rasgos proporcionados por los testigos supondría, además, un problema para la selección de las muestras representativas de la voz del delincuente.

2.6.2. VARIOS LOCUTORES. Una condición particular, presumiblemente presente en todos los hechos imputados a Emilio Berkhoff, aunque solo confirmada en dos testificaciones, se refiere a la presencia de varios locutores al momento de la agresión. La existencia de más de una voz desconocida como lo señalan TRI N° 8 (en D 2), TRI N° 10 (en D 2) y TRI N° 13 (en D 1) plantea la necesidad de encontrar más de una voz objetivo (en la práctica, más de un imputado) y, por supuesto, a la construcción específica de una ronda que responda a la caracterización específica que pueda hacerse de cada una de esas voces. Sin embargo, dado que la caracterización es demasiado general y coincidente, el perfil resultará, con toda probabilidad, semejante en todos los casos y bastará seleccionar cualquier imputado que suene como no mapuche, como señalábamos arriba, con lo cual las ruedas de reconocimiento perderían su carácter diferenciado y no tendrán ningún efecto sobre la identificación de las voces objetivo. Esto se hace especialmente evidente en TRI N° 8 y TRI N°, quienes señalan rasgos dialectales semejantes: “no hablaban mapudungën y por su voz se notaba que no eran del sector” (TRI N° 8) [ y ] “Dos o tres de ellos no eran mapuches” [...] “y forma de hablar no corresponden a las características de los mapuches” (TRI N° 10).

2.6.3. LAS PRUEBAS DE VALIDEZ. Finalmente, a modo de cierre de esta discusión, conviene tener presente que, con independencia de la existencia de un reconocimiento positivo por parte de una o más de una víctima o testigo y de la certeza declarada, la identificación no tiene ningún peso y mal servicio podría proporcionar al juez en su decisión. En otras palabras, más allá de una identificación positiva en una confrontación múltiple y de la concordancia entre testigos independientes como señalan Broeders y Rietveld (1995) esta no puede tomarse como prueba de corroboración. Para que la identificación tenga alguna utilidad es necesario conocer la fiabilidad del oyente, es decir, poner a prueba su capacidad. Para evaluar la validez de la identificación del hablante de un oyente individual, ese testigo tendría que identificar un gran número de muestras de voz para poder obtener su tasa de

identificación correcta (Morrison et al., 2018:99; Morrison, 2011: 14) y realizarse en condiciones similares a las condiciones en las que el oyente hizo la identificación de la voz interrogada. Solo de esa forma una prueba de este tipo podría ser significativa para el tribunal.

### 3. CONCLUSIÓN

Dados los errores procedimentales señalados y la falta de pruebas de validez de los testigos, no es posible determinar si estos eran buenos, promedio o deficientes para identificar a los hablantes por sus voces, y la probabilidad de que las identificaciones sean correctas o incorrectas es, en consecuencia, desconocida. Por lo tanto, no fue apropiado ocuparlas para apoyar la hipótesis acusatoria que mantuvo en prisión preventiva al imputado Emilio Berkhoff durante la fase de investigación y para sustentar la acusación que presentó el ministerio público.

---

Este trabajo considera algunos aspectos abordados en el marco del proyecto Fondecyt Regular N° 1110742 (finalizado).

#### OBRAS CITADAS

- Aitken, Colin y Taroni, Franco (2010). *Estadística y evaluación de la evidencia para expertos forenses*. Madrid: Dykinson: 34-39.
- Bahr, Huntley y Pass, Kimberley (1996). "The influence of style-shifting on voice identification". *Forensic Linguistics* 3 (1): 24-38.
- Blomberg, Mats; Elenius, Daniel; Zetterholm, Elisabeth (2004). "Speaker verification scores and acoustic analysis of a profesional impersonator", *Proceedings of FONETIK 2004*, (Dept. of Linguistics, Stockholm University).
- Bricker, Peter D. y Pruzansky, Sandra (1996). "Effects of stimulus content y duration on talker identification", *Journal of the Acoustical Society of America* 40: 1441-1449.
- Broeders, A. y Rietveld, A. (1995) "Speaker identification by earwitnesses". En J.P. Köster y J.P. Braun (Eds.), *Studies in Forensic Phonetics*. Trier: University Press: 24-40.
- Bull Ray y Clifford Brian (1999). "Earwitness testimony". In Heaton-Armstrong A., Shepherd E., Wolchover D. eds. *Analysing witness testimony: A guide for legal practitioners and other professionals*. London: Blackstone Press: 194-206.
- (1984). "Earwitness Voice Recognition Accuracy". En Wells, Gary y Loftus, Elizabeth (Eds.), *Eyewitness Testimony: Psychological Perspectives*. Cambridge: Cambridge University Press: 92-123.
- Clifford, Brian (1980). "Voices Identification by Human Listeners: On Earwitness Reliability", *Law and Human Behavior* 4: 373-394.

- Clifford, Brian R.; Rathborn, Harriet; Bull, Ray (1981). "The effects of delay on voice recognition accuracy", *Law and Human Behavior* 5: 201-208.
- Deffenbacher, Kenneth, *et al.* (1989). "Relevance of Voice Identification Research to Criteria for Evaluating Reliability of an Identification". *Journal of Psychology*, 123: 109-119.
- Hollien, Harry (1990). *The acoustics of crime*. New York: Plenum.
- Hollien, Harry; Majewski, Wojciech; Doherty, Thomas E. (1982). "Perceptual identification of voice under normal, stress, and disguise speaker conditions". *Journal of Phonetics* 10: 139-148.
- Kassin, Saul M.; Dror, Itiel E.; Kukucka Jeff (2013). "The forensic confirmation bias: Problems, perspectives, and proposed solutions", *Journal of Applied Research in Memory and Cognition*, 2: 42-51. <http://dx.doi.org/10.1016/j.jarmac.2013.01.001>
- Kerstholt, José H.; Jansen, Noortje J. M.; Van Amelsvoort, Adri G., y Broeders, A. P. A. (2006). "Earwitnesses: Effects of accent, retention and telephone. *Applied Cognitive Psychology*", 20(2): 187-197. <http://dx.doi.org/10.1002/acp.1175>
- Kreiman, Jody y Sidtis, Diana (2011). *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception*. UK: Wiley-Blackwell: 237-259.
- Kunzel, Hermann J. (2000). "Effects of voice disguise on speaking fundamental frequency". *Forensic Linguistics* 7: 149-179.
- (1994). "Current approaches to forensic speaker recognition", *Proceedings of the ESCA Workshop on Automatic Speaker Recognition, Identification and Verification*: 135-142.
- Ladefoged, Jenny y Ladefoged Peter (1980). "The ability of listeners to identify voices". *UCLA Working Papers in Phonetics* 49: 43-51.
- LaRiviere, Conrad (1982). "Some acoustic and perceptual correlates of speaker identification". in *Proceedings of the 7th International Congress of Phonetic Sciences*: 558-564.
- Laub, Cindy (2010). *Can earwitness limitations be overcome by the court system? Strategies to help mock jurors appreciate the limitations of earwitness testimony*. PhD dissertation. Lincoln: University of Nebraska.
- Morales, Ana María y Welsch, Gherman (2011). *El reconocimiento de imputados en Chile y a nivel comparado [informes]*. Santiago de Chile: Fundación Paz Ciudadana, 60 p., Solicitar como: DOC/458
- Morrison, Geoffrey Steward (2009). "Forensic voice comparison and the paradigm shift", *Science & Justice* 49: 298-308. doi:10.1016/j.scijus.2009.09.002, que a su vez es una versión revisada de una presentación realizada en el 2nd International Conference on Evidence Law and Forensic Science, Beijing, China, 25-26 julio, 2009. Traducción de Carlos Curiá (Estudios Fónicos / Cuadernos de Trabajo 1 (2011), 1-38).

- Morrison, Geoffrey Stewart; Enzinger, Ewald; Zhang, Cuiling (2018). "Forensic Speech Science". En Freckelton I., Selby H. (Eds.). *Expert Evidence* (Cap. 99). Sydney, Australia: Thomson Reuters: 86-99.
- McGehee, Frances (1947). "The reliability of the identification of the human voice". *Journal of General Psychology* 17: 249-271.
- Nolan, Francis; McDougall, Kirsty; y Hudson, Toby (2013). "Effects of the telephone on perceived voice similarity-Implications for voice line-ups", *International Journal of Speech, Language and the Law*: 20, 229-246. <http://dx.doi.org/10.1558/ijssl.v20i2.229>
- Öhman, Lisa; Eriksson, Anders; y Granhag, Pär Anders (2013). "Angry voices from the past and present: Effects on adults' and children's earwitness testimony". *Journal of Investigative Psychology and Offender Profiling* 10: 57-70. <http://dx.doi.org/10.1002/jip.1381>
- Orchard, Tara L. y Yarmey, Daniel A. (1995). "The effects of whispers, voice-sample duration and voice distinctiveness on criminal speaker identification". *Applied Cognitive Psychology* 9: 249-260.
- Papcun, George; Kreiman, Jody; Davis, Anthony (1989). "Long-term memory for unfamiliar voices". *Journal of the Acoustical Society of America* 85: 913-925.
- Rathborn, Harriet; Bull, Ray; y Clifford, Brian (1981). "Voice recognition over the telephone", *Journal of Police Science and Administration* 9: 280-284.
- Read, Daniel y Craik, Fergus I. M. (1995). "Earwitness identification: Some influences on voice recognition" *Journal of Experimental Psychology: Applied* 1: 6-18.
- Reich, Alan R. (1981). "Detecting the presence of vocal disguise in the male voice". *Journal of the Acoustical Society of America* 69:1458-1461.
- Reich, Alan R. y Duke, James (1979). "Effects of selected vocal disguise upon speaker identification by listening". *Journal of the Acoustical Society of America* 66:1023-1028.
- Rosas, Claudia, Sommerhoff, Jorge y Sáez, César (2018). "Identificación de voces por auditores profanos no nativos". *Alpha* (Osorno), (46), 129-150. <https://dx.doi.org/10.4067/S0718-22012018000100129>.
- Rose, Philip (2002). *Forensic speaker identification*. London: Taylor and Francis.
- Rose, Philip y Duncan Sally (1995). "Naive auditory identification and discrimination of similar voices of familiar speakers". *Forensic Linguistics* 2: 1-17.
- Sarwar, Farhan; Allwood, Carl Martin; Zetterholm, Elisabeth (2014). "Earwitnesses: The type of voice lineup affects the proportion of correct identifications and the realism in confidence judgments". *International Journal of Speech, Language and the Law*, 21: 139-155. <http://dx.doi.org/10.1558/ijssl.v21i1.139>
- Saslove, Howard y Yarmey, Daniel (1980). "Long-Term Auditory Memory: Speaker Identification". *Journal of Applied Psychology* 65: 111-116.

- Solan, Lawrence M. y Tiersma Peter M. (2003). "Hearing voices: Speaker identification in court". *Hastings Law Journal* 54: 373–435.
- Van Lancker, Diana; Kreiman, Jody; Emmorey, Karen (1985). "Familiar voice recognition: Patterns and parameters. Part I. Recognition of backward voices". *Journal of Phonetics* 13:19-38.
- Van Wallendael Lori R.; Surace, Amy; Hall-Parsons, Deborah Hall; Brown, Melissa (1994). "Earwitness' voice recognition: Factors affecting accuracy and impact on jurors". *Applied Cognitive Psychology* 8: 661–677.
- Yarmey, A. Daniel; Yarmey, A. Linda; Yarmey, Meagan J.; Parliament, Lisa (2001). "Common sense beliefs and the identification of familiar voices". *Applied Cognitive Psychology* 15: 283-299.
- Zetterholm, Elisabeth; Sarwar, Farhan; Thorvaldsson, Valgeir; Allwood, Carl Martin (2012). "Earwitnesses: The effect of type of vocal differences on correct identification and confidence accuracy", *International Journal of Speech, Language and the Law*, 19: 219-237. <http://dx.doi.org/10.1558/ijssl.v19i2.219>

ANEXO 1: Comunicado público escrito por Emilio Berkhoff, desde la clandestinidad. Disponible en: <http://www.eldesconcierto.cl/2015/07/13/emilio-berkhoff-desde-la-clyestinidad-es-mi-deber-como-revolucionario-declararme-en-rebeldia/>

ANEXO 2: Sentencia sobre caso de Emilio Berkhoff Jerez (2015) / Poder Judicial. República de Chile. Tribunal Oral en Lo Penal. Cañete. 2015. Causa RIT número 121-2014, seguida en contra de EMILIO JUAN PABLO BERKHOFF JEREZ. Disponible en: <http://www.pjud.cl/documents/396729/0/Sentencia+Berkhoff+%28121-2014%29.pdf/84970a1d-b7e2-4602-b3d0-878da7295d69>