

<https://doi.org/10.32735/S0718-22012022000551099>

169-191

COMPARACIÓN FORENSE DE LA VOZ EN CHILE: BALANCE, PROYECCIONES Y DESAFÍOS

Forensic voice comparison in Chile: Balance sheet, projections and challenges

CLAUDIA ROSAS AGUILAR
Universidad Austral de Chile
 claudiarosas@uach.cl

JORGE SOMMERHOFF HYDE
Universidad Austral de Chile
 jsommerh@uach.cl

JAIME PACHECO QUEZADA
Universidad de Concepción (Chile)
 jaipacheco@udec.cl

CÉSAR SÁEZ ELGUETA
Policía de Investigaciones de Chile
 cesarsaez.e@gmail.com

Resumen

El objetivo de este trabajo es describir la situación actual y las proyecciones de la comparación forense de la voz en Chile, a la luz de los estándares internacionales que se enfocan en la producción de evidencia confiable, que pueda ser utilizada en los tribunales de justicia. Para ello, en principio, se dan a conocer los fundamentos científicos que caracterizan esta rama de las ciencias forenses y, luego, sobre esa base se analiza y comenta la situación de Chile y sus proyecciones, tanto en la práctica como en la investigación forense.

Palabras clave: Comparación forense de la voz; estándares internacionales; evidencia confiable; tribunales de justicia.

Abstract

The objective of this work is to describe the current situation and the projections of forensic voice comparison in Chile, in the light of international standards that focus on the production of reliable evidence that can be used in the courts of justice. To do this, first, the scientific foundations that characterize this branch of forensic sciences are disclosed and, then, on that basis, the situation in Chile and its projections are analyzed and discussed, both in practice and in the forensic research.

Key words: Forensic voice comparison; international standards; reliable evidence; courts of justice.

1. INTRODUCCIÓN

La comparación forense de la voz es una rama de las ciencias forenses en la que el profesional compara las propiedades del habla de las grabaciones de un delincuente y

Recibido: 10 marzo 2021

Aceptado: 23 junio 2021

de un sospechoso con el fin de proporcionar una opinión experta que, por último, ayude al juez o a un tribunal a determinar la identidad del primero. Es una actividad no exenta de dificultades. Dos circunstancias así lo determinan: la variabilidad de la voz humana y el importante despliegue de recursos requeridos para realizar una comparación de voz en el contexto de un caso de trabajo forense.

En efecto, por una parte, las propiedades de las voces que se comparan son, a diferencia de las que se encuentran en otras ramas de las ciencias forenses –como el perfil de ADN– de naturaleza continua y variable. Por ejemplo, la frecuencia fundamental de un hablante en particular no tiene por qué coincidir con un número entero específico, como, p.e., 124 Hz, por el contrario, puede exhibir cualquier valor, como 125, 126, 124,5, 124,9, 124,999999, 125,0000001 Hz, etc. Asimismo, el valor de la medición extraída en una oportunidad, con toda seguridad, no será el mismo, obtenido en otra ocasión, por razones de diverso orden: emocionales, de salud, de estilo, o solo por azar. Y entre hablantes, las voces variarán, además, por el grupo social o profesional de pertenencia y por las cualidades de sus órganos de fonación y articulación. A esta genuina variabilidad intra e interhablante se suma aquella otra introducida por efectos de las condiciones ambientales y técnicas en que se obtienen las grabaciones; sobre todo, por la presencia de más o menos ruido de fondo o reverberación y por el tipo de canal ocupado (microfónico o telefónico), hecho que conviene recordar debido a que en un escenario forense típico, las condiciones técnicas de la grabación del delincuente y del sospechoso, por norma, difieren; por ejemplo, en el caso del delincuente, una interceptación telefónica de una conversación, con ruido de fondo y escasa duración y, en el caso del sospechoso, una grabación microfónica de un interrogatorio policial, en una habitación tranquila y con mayor duración. Estos desajustes podrían provocar que algunas diferencias en la grabación del delincuente respecto de la grabación del sospechoso podrían verse aumentadas o enmascaradas, haciendo que la tarea de distinguir la variabilidad propia de una voz, respecto de la variabilidad entre hablantes diferentes, sea una cuestión difícil de determinar, y, por lo mismo, fuente potencial de interpretaciones erróneas (Morrison *et al.*, 2018).

La comparación forense de la voz es, contrario a lo que pudiera inferirse por su nombre, mucho más que solo observar de manera atenta semejanzas y diferencias entre las propiedades de los registros de voz de identidad desconocida y conocida, consignar cada hecho y presentar una opinión al juez acerca de la probabilidad de que procedan o no del mismo hablante. Una forma de abordar en forma correcta la naturaleza compleja de la comparación forense de la voz es asumirla como parte de un sistema que contiene ciertas decisiones que, en última instancia, deberán ser puestas a prueba para que el juez o tribunal, junto con el resultado informado, pueda asignarle un cierto grado de fiabilidad y admitirla o no como evidencia de un determinado hecho en disputa. En un escenario típico de comparación forense de la voz existen dos hipótesis contrarias, la del fiscal –o hipótesis acusatoria– y la de la defensa –o hipótesis alternativa–, donde la tarea del perito será, en

efecto, evaluar el peso de cada hipótesis e informar al juez de un resultado; sin embargo, previo a todo eso, también deberá realizar diversas acciones que incluirán, por ejemplo, los procedimientos para medir las propiedades de las grabaciones de voz y los modelos estadísticos para calcular los valores que se informarán, así como los marcos de interpretación de los resultados obtenidos, todo ello, en su conjunto, define la noción de *sistema* de comparación, la que representa de manera más fiel los conocimientos, capacidades, experiencia y recursos de diversa índole que han de estar involucrados en el proceso de comparación forense de la voz (Morrison y Thompson, 2017)¹.

Pensamos que aspectos generales como los referidos, respuestas sólidas por parte de científicos connotados en el ámbito de la inferencia forense, junto con el conocimiento del *status quo* de la comparación forense de la voz en nuestro país –de sus logros, dificultades y desafíos– pueden ayudar a comprender, de mejor manera, las necesidades de quienes trabajan en el ámbito de la comparación forense de la voz, y a valorar las ventajas de una sinergia colaborativa entre organismos gubernamentales de orden y seguridad, de investigación en centros académicos, y de profesionales públicos o privados en pro de elevar los estándares para el desarrollo de una práctica e investigación en la materia, que cumpla con los requisitos de admisibilidad internacionales.

2. LOS FUNDAMENTOS DE LA COMPARACIÓN FORENSE DE LA VOZ

Un científico forense, perito o investigador, necesita, en primer lugar, extraer datos –las propiedades de la evidencia– para, después, analizarlos, interpretarlos y, en último lugar, presentarlos a un juez o tribunal. Estamos hablando sobre todo de enfoques, marcos y procedimientos de validación. En este apartado nos referiremos de manera amplia a la forma en que se realiza una comparación forense de la voz y a los principios, conocimientos y recursos en general que dicha actividad supone.

2.1. ENFOQUES

Los enfoques o métodos generales ahora más usados para la extracción de información de la señal con fines forenses, son el *enfoque estadístico acústico-fonético* y el *enfoque automático*², debido a la naturaleza objetiva que los hace más compatibles con el

¹ Morrison y Thompson (2017) lo definen así: [...] *We use the term system to designate the whole of the data and the processes used to evaluate the strength of evidence after the forensic scientist has stated what competing hypotheses they intend to evaluate. A forensic voice comparison system can include the sampling and selection of relevant voice recordings, procedures used to measure properties of the voice recordings, and statistical models used to calculate values which will be reported. A system also includes any actions taken by a human* (p. 363).

² El método auditivo y el espectrográfico, que también se han utilizado en la comparación de voces con fines forenses, se basan en mediciones subjetivas y, por tanto, son incompatibles con un marco científico de evaluación de la evidencia. En el primero el analista –normalmente un fonetista– utiliza su habilidad natural de oír junto con la formación y experiencia que tiene en la producción y percepción de una lengua a nivel laríngeo

marco objetivo de la razón de probabilidad. El enfoque *estadístico acústico-fonético* implica realizar mediciones cuantitativas de las propiedades acústicas de las muestras de voz y analizar estadísticamente los datos numéricos resultantes. Las propiedades acústicas medidas por los profesionales del enfoque acústico-fonético son típicamente las que se utilizan en estudios empíricos de producción y percepción del habla, por ejemplo, frecuencias de formantes, frecuencia fundamental y VOT. Las mediciones acústicas en general se realizan utilizando implementaciones de *software* de algoritmos de procesamiento de señales con supervisión humana. Por lo general, se identifican unidades fonéticas comparables en grabaciones de locutores conocidos y cuestionados y luego se miden las propiedades acústicas de dichas unidades³. El enfoque *automático*, al igual que el enfoque acústico-fonético, se basa casi siempre en mediciones cuantitativas de las propiedades de los coeficientes cepstrales de frecuencia de Mel (MFCC), en lugar de las propiedades acústicas tradicionales (fonemas, F0, formantes, VOT). Los resultados numéricos se usan como entrada para los modelos estadísticos, como el modelo de fondo universal (GMM-UBM) o i-vectores más análisis discriminante lineal probabilístico (PLDA). Este enfoque tiene sobre el enfoque anterior la ventaja de que debido a que las mediciones y los modelos estadísticos operan de manera automática, permite procesar en forma rápida grandes cantidades de archivos y, por tanto, procesar cientos o miles de pares de pruebas.⁴

2.2. MARCOS

Se ha establecido que el marco conceptual correcto para evaluar la evidencia forense es el *marco de la razón de probabilidad*. De hecho, es el marco que se ocupa en las principales ramas de las ciencias forenses como lo es la comparación de perfiles de ADN⁵

y articulatorio para consignar las similitudes y diferencias entre las muestras comparadas. Detalles de este tipo de enfoque se pueden leer, p.e., en Hollien (2002, 2016). Por su parte, en el método espectrográfico –ya en desuso– el profesional forense emitía un juicio subjetivo a partir de la representación gráfica de una señal analógica. Una descripción del enfoque espectrográfico o auditivo-espectrográfico se puede encontrar, p.e., en Kersta (1962). En cualquier caso, Morrison (2014) señala que el uso de espectrogramas o de cualquier otra medición no debe ser el aspecto clave de la crítica sino el de si se ha probado o no la validez y fiabilidad de dicho enfoque bajo las condiciones del caso de trabajo particular (pp. 252-254).

³ Descripciones más detalladas del enfoque estadístico acústico-fonético se pueden encontrar, p.e., en Rose (2006, 2017).

⁴ Conviene hacer notar que estas ventajas no anulan el papel del científico forense, quien estará a cargo de la selección de los datos adecuados a las características del caso (población relevante). Una descripción más detallada del enfoque automático para la comparación de voz forense se pueden encontrar en Morrison y Enzinger (2018).

⁵ El *marco de razón de probabilidad* es un marco conceptual que puede aplicarse a todo tipo de creencias subjetivas basadas en la experiencia en cuanto a las probabilidades de la evidencia dadas las hipótesis en competencia, solo cambia la forma en que se presentan los resultados de la comparación. Para objetos cuyas

(Morrison 2009). Una razón de probabilidad es un componente del Teorema de Bayes, el que se expresa mediante tres ecuaciones: la apuesta a priori, la razón de probabilidad y las apuestas a *posteriori*. El Teorema de Bayes queda expresado en la ecuación 1:

$$p(H_{mo}|E)/p(H_{do}|E) = p(E|H_{mo})/p(E|H_{do}) \times P(H_{mo})/P(H_{do}) \quad (1)$$

apuesta a posteriori razón de probabilidad apuesta a priori

Donde la primera ecuación de izquierda a derecha representa la *apuesta a posteriori*; la segunda, la *razón de probabilidad* y, la última, la *apuesta a priori*. En el contexto de la comparación forense de la voz, estas ecuaciones representan, en sentido inverso, de derecha a izquierda, las creencias previas del juzgador de los hechos acerca de la culpabilidad o inocencia de un individuo, la evidencia –los audios de origen desconocido y conocido, en comparación– y la decisión del juez acerca de la culpabilidad o inocencia del imputado. El rol del científico forense se enmarca de manera estricta en la evaluación de la evidencia, es decir, la ecuación de la razón de probabilidad (Ec. 2):

$$RV = p(E|H_{mo})/p(E|H_{do}) \quad (2)$$

Donde RV es la razón de probabilidad (LR en inglés), E es la evidencia, H_{mo} es la hipótesis del mismo origen, y H_{do} es la hipótesis de diferente origen.

El *marco de la razón de probabilidad*, aplicado a la comparación forense de la voz, cuantifica la probabilidad de obtener las propiedades de la voz en la grabación del delincuente si hubieran sido producidas por el sospechoso, versus la probabilidad de obtener las propiedades de la voz en la grabación del delincuente si no hubieran sido producidas por el sospechoso, sino por cualquier otro hablante de la *población*. Este marco exige, así, considerar no solo la *similitud*, representada en el numerador de la ecuación, sino también la *tipicidad* de dichas características en la *población relevante*, definida esta última como la comunidad de donde procede el delincuente, y representada en el denominador de la ecuación. Si la razón de probabilidad tiene un valor mayor que 1, entonces la evidencia es más probable bajo la hipótesis del origen idéntico que bajo la hipótesis del origen diferente, y cuanto mayor sea el valor de la razón de probabilidad, mayor será el apoyo relativo para la hipótesis del mismo origen sobre la hipótesis del origen diferente. *Mutatis mutandis*, si la razón de probabilidad tiene un valor menor que 1, entonces la evidencia es más probable bajo la hipótesis de origen diferente que bajo la hipótesis del mismo origen, y cuanto menor sea el valor de la razón de probabilidad, mayor será el apoyo relativo para la hipótesis de diferente origen en relación con la hipótesis del mismo origen. Si el valor de la razón de probabilidad es cercano a 1, entonces la evidencia es también probable bajo cualquiera de las dos hipótesis.

propiedades son discretas, como el ADN, se ocupan gráficos de barras, mientras que para datos continuos como la voz, se ocupan histogramas o funciones de densidad de probabilidad (Morrison *et al.*, 2018).

En definitiva, el marco de la razón de probabilidad constituye un marco de interpretación objetiva de los resultados, se ajusta a las propiedades continuas del habla, y al rol específico del analista forense. Estos rasgos no han sido considerados en otros tipos de testificaciones como el *marco de dos etapas*, el *marco de solo similitud* y el *marco de probabilidad posterior*.⁶ En ocasiones, además, se suelen encontrar conclusiones como “con un grado razonable de certeza científica, la voz en la grabación del locutor cuestionado fue producida por el locutor conocido”. La expresión “un grado razonable de certeza científica” no se utiliza en la ciencia en general y no tiene un significado claramente definido⁷.

2.3. EL MARCO DE LA RAZÓN DE PROBABILIDAD Y LOS REQUISITOS DE ADMISIBILIDAD

La producción de evidencia confiable es consustancial al rol del perito forense, porque este debe proporcionar una opinión experta que ayude al juez de los hechos a determinar la identidad de un locutor. Y la única manera de hacer aquello es proporcionando al juez un testimonio que sea válido y confiable como para que pueda admitirlo, por lo que, enfoques, marcos y procedimientos en general están vinculados de manera lógica por ese propósito.

⁶ El *marco de dos etapas*, si bien es un marco de la razón de probabilidad, considera los resultados de las mediciones de las muestras de voces comparadas, como si se trataran de propiedades discretas, en lugar de propiedades continuas. A partir de umbrales pre especificados se pretende que un valor que quepa dentro de dicho rango sea prueba infalible de identidad, por ejemplo, si el umbral se estableciera en 10 Hz, un valor de 9,9 Hz se declararía “coincidente”, pero 10,1 (casi idéntico) se declararía “no coincidente”. Esto funciona para el ADN donde el perfil de un organismo no cambia de una ocasión a otra y es posible, en consecuencia, obtener una coincidencia, pero no en el caso de la voz humana que es continua y variable. El *marco de solo similitud* solo evalúa el grado de similitud, entre las muestras comparadas, mediante expresiones como “la frecuencia fundamental de la voz en la grabación del locutor cuestionado coincide con la de la voz en la grabación del locutor conocido”, lo que puede conducir a conclusiones engañosas, sin considerar la probabilidad de encontrar esa característica en la población relevante. El *marco de probabilidad posterior* presenta una conjetura, a partir de una probabilidad de la razón de probabilidad. Esto es incorrecto, ya que la única forma lógica de derivar una probabilidad posterior—alcanzar una conclusión respecto de la culpabilidad versus la inocencia del sospechoso—es combinar la razón de probabilidad con una probabilidad previa que, en el contexto judicial, es la creencia que tiene el evaluador de los hechos acerca de la probabilidad de que el locutor cuestionado sea el locutor conocido, antes de que se realice el análisis de la comparación. Debido a que el científico forense no conoce la probabilidad previa (solo la conoce el juez), no puede calcular la probabilidad posterior apropiada. De modo que, cualquier forma de “identificación” o “exclusión” informada por el profesional forense es una forma conceptual del marco de probabilidad posterior que no es teóricamente posible, ni legalmente adecuada.

⁷ El informe de la Comisión Nacional de Ciencias Forenses (NCFCS) y el Consejo de Asesores de Ciencia y Tecnología del Presidente (PCAST) han recomendado que no se utilicen expresiones de este tipo, porque ninguna expresión verbal puede competir con la objetividad de una expresión numérica (Morrison y Thompson, 2017, p. 358).

2.3.1. ESTÁNDARES DE ADMISIBILIDAD

Morrison y Thompson (2017) han enfatizado la importancia de la *demostración empírica de la validez y fiabilidad*⁸ como criterio imprescindible tendiente a garantizar que una evidencia sea confiable y pueda ser admitida por un juez o tribunal, en lugar de otros tales como *la revisión por pares y publicación, la existencia de estándares, la aceptación general, la vinculación entre el testimonio y el caso, y la prueba empírica de la validez y fiabilidad*. En efecto, ningún otro criterio de manera independiente o combinada puede dar garantías de confiabilidad, por ejemplo con relación a si la teoría o técnica ha sido sometida a *revisión y publicación por pares*, los autores argumentan que una proporción sustancial de los artículos publicados en ciencias forenses, en general, y la comparación de voces forenses, en particular, sufren de defectos metodológicos importantes, incluido el uso de bases de datos que son muy pequeñas y que no representan condiciones forenses realistas y que incluso la calidad de la revisión en revistas respetadas arbitradas por pares puede fallar debido a la variable dificultad de encontrar revisores que estén calificados, que tengan tiempo y que estén dispuestos a revisar documentos de forma voluntaria. La *existencia y mantenimiento de estándares* que controlan el funcionamiento de una técnica, por su parte, solo indica que se ha seguido un estándar y que existe documentación que lo demuestra, pero no garantiza que los resultados de un análisis particular sean necesariamente correctos, en especial si el estándar o los procedimientos de validación no son en realidad adecuados para su propósito, porque, p.e., no se han seguido los procedimientos para elaborar un estándar nacional o internacional⁹, o porque han sido producidos por los propios partidarios de una técnica en particular, etc. La *aceptación de una determinada técnica por parte de una comunidad científica relevante* tampoco es una garantía de fiabilidad, ya que depende de la definición de la comunidad; una comunidad restringida otorgará un grado de aceptación diferente a una determinada técnica dependiendo de si es compatible con su práctica o no, en contraste con una comunidad científica relativamente amplia que considere diferentes enfoques y disciplinas. La *vinculación entre el testimonio y los hechos del caso*, que sugiere que la validez general de una teoría o método no es suficiente para garantizar que

⁸ Este criterio ha sido recogido por regulaciones de justicia de Estados Unidos y Reino Unido (la Regla Federal 702, los fallos de la Corte Suprema de Estados Unidos y las Instrucciones de Práctica Criminal 19 A de Inglaterra y Gales). Ver: Federal Rule of Evidence 702 [Regla Federal de la Evidencia 702]. (1975). Pub. [Consolidado con enmiendas: 2011]. Legal Information Institute LII. Disponible en: https://www.law.com/ell.edu/rules/fre/rule_702 [Consulta: 20/8/20]. Criminal Practice Directions. [Instrucciones de Práctica Criminal] (2015). Pub.19A: Expert Evidence. [Consolidado con enmiendas:2019]. Disponible en: <https://www.judiciary.uk/wp-content/uploads/2019/03/crim-pd-amendment-no-8-consolidated-mar2019.pdf> [Consulta: 20/8/20].

⁹ Para más detalles relativas a este aspecto, ver, p.e., ISO/IEC 17025 *Requisitos generales para la competencia de los laboratorios de ensayo y calibración*. Disponible en: <https://www.iso.org/standard/39883.html> [Consulta: 22-11-20]. Resumen disponible en: <https://www.iso.org/obp/ui/#iso:std:iso-iec:17025:ed-3:v2:es> [Consulta: 22/11/20]

la evidencia producida sea válida y confiable, sino que requiere que dicha metodología haya sido aplicada de manera confiable al caso en particular, es decir, que haya tenido en cuenta las circunstancias en que este se ha producido se observa con pertinencia¹⁰. Sin embargo, por sí misma no es suficiente, se requiere probar el sistema bajo las condiciones del caso bajo investigación. Solo la *prueba empírica de la validez de la técnica* resulta ser el criterio que los autores destacan y justifican como decisivo para evaluar el funcionamiento del sistema utilizado y lo interpretan como *la comprobación empírica del grado de validez y fiabilidad del sistema, bajo condiciones que reflejen las del caso bajo investigación*, y que el investigador forense debe presentar al juez para que este pueda tenerlo en cuenta al decidir acerca de la admisibilidad del testimonio, debido a que en última instancia es el juez en una audiencia de admisibilidad, o bien en el proceso de valoración de la pericia, quien debe estar satisfecho con el sistema, conforme con lo que el experto ha hecho, a base de la simulación de las condiciones del caso, como para admitir su testimonio basado en dicho sistema, y considerar, en función de los resultados de la prueba, el grado en que confiará en la salida del mismo.

2.3.2. PROCEDIMIENTOS PARA LA VALIDACIÓN EMPÍRICA DEL DESEMPEÑO DE LOS SISTEMAS DE COMPARACIÓN FORENSE DE LA VOZ

Ha habido varios llamados para que la *validez* y la *fiabilidad* de la comparación forense de la voz sea probada empíricamente¹¹, como los contenidos en el Informe del *Consejo Nacional de Investigación* de Estados Unidos (NRC) de 2009 y el Regulador de Ciencia Forense de Reino Unido (FSR-2014) de 2014 que incluye, además, como requisito previo a la validación del método, la acreditación de los proveedores de ciencias forenses. Su importancia se sustenta en el hecho de que constituyen la única forma de demostrar qué tan bien funciona en realidad un sistema de análisis forense. Debido a lo anterior, las pruebas empíricas de validez y fiabilidad deben ser ciegas al funcionamiento interno del sistema.¹²

¹⁰ Para la comparación forense de la voz, lo anterior incluiría preguntas pertinente a si el experto eligió una población relevante apropiada para el caso en cuestión, si obtuvo una muestra de datos suficientemente representativa de esa población relevante, y si los datos o las técnicas analíticas se aplicaron de manera adecuada al estilo de discurso y condiciones de grabación en las grabaciones de hablantes conocidos y cuestionados. Los autores argumentan que el juez también debería considerar si el experto tomó las medidas adecuadas para evitar sesgos cognitivos y contextuales, ya que esto también afecta la validez de los métodos en la forma en que se aplican y, por tanto, la fiabilidad de las conclusiones del experto en el caso en cuestión.

¹¹ Para una revisión de los llamados desde la década de 1960, consultar: Morrison (2014).

¹² Para obtener descripciones más detalladas de los procedimientos para probar empíricamente la validez y fiabilidad de los sistemas de análisis forense (incluidos los sistemas de comparación de voz forense), y las métricas y gráficos para comunicar los resultados, ver Morrison (2011), Haraksim *et al.* (2016) y Morrison y Enzinger (2016).

En estadística y literatura científica, la *validez* es sinónimo de exactitud y la *fiabilidad*, de precisión, y son conceptos graduales, por lo que, aplicados a un sistema de comparación forense de la voz en particular, harán referencia a grados de mayor o menor *validez* y grados de mayor o menor *fiabilidad* encontrados en dicho sistema (Morrison *et al.* 2018).

Para probar de manera empírica la *validez*, o *exactitud*, de un sistema de comparación de voz forense, se debe disponer de un conjunto de datos de prueba que incluyan pares de grabaciones de voz de los mismos y diferentes locutores para los cuales el evaluador conoce el origen en cada situación. Los datos de la prueba deben ser representativos de la población relevante para el caso, y las condiciones de un miembro de cada par de prueba debe reflejar el estilo de habla y las condiciones de registro de la grabación del locutor conocido, y las condiciones del otro miembro de cada prueba deben reflejar el estilo de habla y las condiciones de registro del locutor cuestionado. Una medida de *validez* que es consistente con el marco de la razón de probabilidad es el *coste del logaritmo de la razón de probabilidad* (log-likelihood-ratio cost, *Cllr*).¹³ Un sistema será bueno si para los pares de grabaciones de mismo origen da valores de salida mayores que 1 y para valores de salida de pares de origen diferente valores menores que 1. El *Cllr* penaliza los valores de salida de mismo origen menores que 1 y valores de salida de pares de origen diferente mayores que 1¹⁴. Cuanto más bajo sea el *Cllr*, mejor será el rendimiento del sistema. Si se prueban varios sistemas utilizando el mismo conjunto de datos de prueba, entonces el sistema más exacto es el sistema que produce el valor *Cllr* más bajo. Para ser significativo en el caso de trabajo, los datos de la prueba deben ser muestras representativas de la población relevante, y las condiciones de cada miembro de cada par de pruebas deben reflejar lo más fiel posible las condiciones de los locutores

¹³ Morrison (2009) recomienda convertir las razones de probabilidad en razones de probabilidad logarítmicas, ya que estas últimas son simétricas en torno a cero, por ejemplo, razones de probabilidad de 1000/1 y 1/1000 se convierten en razones de probabilidad log-base-diez de +3 y -3, respectivamente, y una razón de probabilidad de 1/0 corresponde a una razón de probabilidad logarítmica de 0.

¹⁴ Para calcular *Cllr*, primero se debe calcular un valor de penalización para la razón de probabilidad de cada par de prueba (el valor de penalización aumenta a medida que los valores contrarios a los hechos son más grandes). Luego se encuentra la media de todos los valores de penalización de pares de prueba del mismo origen y la media de todos los valores de penalización de pares de prueba de origen diferente, y por último se toma la media de las últimas dos medias.

La función para calcular *Cllr* se proporciona en la siguiente fórmula:

$$C_{llr} = \frac{1}{2} \left(\frac{1}{N_s} \sum_i^{N_s} \log_2 \left(1 + \frac{1}{LR_{si}} \right) + \frac{1}{N_d} \sum_j^{N_d} \log_2 (1 + LR_{dj}) \right) \quad (3)$$

Donde N_s y N_d son el número de pares de prueba del mismo hablante y de diferentes hablantes, y LR_s y LR_d son las razones de probabilidad derivadas de los pares de prueba del mismo hablante y de diferentes hablantes. Un valor de penalización del mismo origen es $\log_2 (1 + 1 / LR_s)$, y un valor de penalización de origen diferente es $\log_2 (1 + LR_d)$ (Morrison *et al.*, 2018).

conocidos y cuestionados de las grabaciones. De esta manera, los resultados de las pruebas reflejarán el rendimiento esperado en el sistema en las condiciones del caso.

En cuanto a la medición de la *fiabilidad o precisión*, existe un desacuerdo entre los estadísticos forenses acerca de la mejor manera de manejar la imprecisión¹⁵. Por ello, Morrison *et al.* (2018) recomiendan que los abogados que contratan los servicios de un profesional forense le pregunten cómo el método ocupado para evaluar la probabilidad tiene en cuenta la precisión.

Otro aspecto importante y complementario en la evaluación empírica del sistema de comparación forense de la voz es la *representación de los resultados*. Un método gráfico para presentar los resultados de las pruebas de un sistema de análisis forense es un diagrama de Tippett. Las gráficas de Tippett se introdujeron en Meuwly (2001) –inspiradas en el trabajo de C.F. Tippett–, y ahora son un método estándar para presentar resultados en la investigación forense de comparación de voces, porque proporcionan información más detallada de los resultados de la que está disponible en una medida de resumen como el *coste del logaritmo de la razón de probabilidad*, *Cllr*.

Junto con todo lo anterior, Morrison y Thompson (2017) alertan de un factor que puede afectar los resultados del analista, y por esta, la admisibilidad de su testimonio: la influencia potencial del *sesgo contextual*. El *sesgo contextual* se refiere a toda aquella información que forma parte del caso de trabajo, pero que no es relevante para la comparación conducida por el analista y que puede influir en su percepción y, por tanto, en la forma en que construye su sistema de comparación forense de la voz. El tema es sensible, porque en ocasiones el analista requiere información clave para tomar decisiones, p.e., acerca de la definición de la población relevante del caso bajo investigación, pero cuya obtención está asociada a aspectos que podrían introducir sesgo por contaminación con información contextual. Para sortear este riesgo se han diseñado diversas fórmulas, que incluyen, por ejemplo: (a) mediadores de información en los casos de trabajo para impedir que los implicados tomen contacto, (b) información suministrada al analista según las necesidades del caso de trabajo, (c) evitar que se conozcan los detalles de la fuente de la muestra, (d) uso de enfoques de análisis objetivos (mediciones cuantitativas).

¹⁵ Por ejemplo, señalan Morrison *et al.* (2018, p. 39), que algunos proporcionan una mejor estimación para el valor de la razón de probabilidad y un rango de valores dentro del cual creen que es probable que el valor se encuentre, por ejemplo, una mejor estimación de 1000 con una probabilidad del 90% de estar en el rango de 900 a 1100. Una alternativa a esto es informar solo el límite más cercano a una razón de probabilidad de 1, por ej., una probabilidad del 95% de que la razón de probabilidad sea de al menos 900. Otros tienen objeciones filosóficas a este enfoque y, en cambio, calculan un valor único, pero ese valor único en general estaría más cerca de una razón de probabilidad de 1 que la mejor estimación del primer grupo.

3. LA SITUACIÓN DE LA COMPARACIÓN FORENSE DE LA VOZ EN CHILE

Basado en fundamentos científicos, en breve presentados, se expone un análisis y comentario de la situación de la comparación forense de la voz en Chile tanto en la práctica como en la investigación, previa descripción del marco legal en que se inscribe.

3.1. MARCO LEGAL

En esta sección nos referiremos, sobre todo, a los requisitos que establece la ley en Chile sobre *la procedencia de la comparación forense de la voz como medio de prueba, la testificación del perito, la admisibilidad del informe y su compatibilidad con los estándares internacionales observados.*

Respecto de *la procedencia de la comparación forense de voz*, el Código Procesal Penal (2020) establece la libertad probatoria (art. 295; 323), de modo que el *testimonio de comparación de voz* puede ser ocupado como medio de prueba. En cuanto a *la testificación del perito*, se establece que un perito acreditado profesionalmente podrá testificar “en los casos determinados por la ley y siempre que para apreciar algún hecho o circunstancia relevante para la causa fueren necesarios o convenientes conocimientos especiales de una ciencia, arte u oficio” (art. 314), “Los informes deberán emitirse con imparcialidad, ateniéndose a los principios de la ciencia o reglas del arte u oficio que profesare el perito” (art. 314) y “Sin perjuicio del deber de los peritos de concurrir a declarar ante el tribunal acerca de su informe, este deberá entregarse por escrito y contener: a) La descripción de la persona o cosa que fuere objeto de él, del estado y modo en que se hallare; b) La relación circunstanciada de todas las operaciones practicadas y su resultado, y c) Las conclusiones que, en vista de tales datos, formularen los peritos conforme a los principios de su ciencia o reglas de su arte u oficio” (art. 315). Interpretamos este requisito como *la descripción de las condiciones del caso, de los métodos y protocolos aplicados y del marco de razonamiento ocupado para evaluar los resultados y presentar la evidencia de acuerdo con principios científicos.*

En cuanto a *la admisibilidad del informe*, el Código establece que “El juez de garantía admitirá los informes y citará a los peritos cuando, además de los requisitos generales para la admisibilidad de las solicitudes de prueba, considerare que los peritos y sus informes otorgan suficientes garantías de seriedad y profesionalismo” (art. 316). Como consecuencia de la anterior formulación, resulta que será necesario demostrar la “seriedad” y “profesionalismo”, hecho que implica, en principio, conocer qué significan ambos conceptos y, luego, definir la forma mediante la cual se pueden demostrar los contenidos implicados. El término “seriedad” se define en la RAE (2014) como la “cualidad de serio” y esta, a su vez, en una de sus acepciones como, “real, verdadero y sincero, sin engaño o burla, doblez o disimulo”. Por su parte, el término “profesionalismo” se define como

“cultivo o utilización de ciertas disciplinas, artes o deportes, como medio de lucro”.¹⁶ Así, a base de una consideración combinada de los conceptos definidos, podemos interpretar que el perito es un profesional que realiza un trabajo verdadero o –en la terminología que se ha venido comentando– presenta un testimonio de manera *válida y confiable* para ello habrá de *explicar al juez que el sistema se puede aplicar convenientemente al caso (mediante la descripción de los enfoques y procedimientos generales aplicados) y demostrando en forma empírica su grado de validez y fiabilidad*, de forma que el juez pueda asignarle algún crédito a los resultados obtenidos.

3.2. PRÁCTICA

Sobre la base de la fundamentación científica, y mediante una revisión de la documentación existente disponible, se ofrece un análisis general acerca de cómo los investigadores y peritos forenses han analizado, evaluado y presentado la evidencia.¹⁷

En 2009, Celis en su *Descripción de los métodos utilizados en reconocimiento forense de locutores y su implementación en Chile* describe una situación bastante discreta en términos de profesionales e investigadores abocados a la práctica –e investigación– en nuestro país. Una encuesta¹⁸ aplicada recientemente ha demostrado que tras una década de dicha constatación, el número de profesionales dedicados a la práctica forense de la voz sigue siendo reducido, y como veremos más adelante, la Policía de Investigaciones de Chile se mantiene no solo como la entidad que más informes de análisis de la voz emite, sino como la única institución que ha aplicado de manera sistemática la metodología bayesiana a la comparación forense de la voz.

3.2.1. LAS CONDICIONES DE LOS CASOS DE TRABAJO EN CHILE

Una evaluación de las condiciones de los casos de trabajo en Chile, desarrollada en 2019 por la Universidad Austral de Chile y la Policía de Investigaciones de Chile, en el marco de un proyecto internacional,¹⁹ reveló que –en promedio– las características demográficas de los casos recibidos entre 2009 y 2019 por la Policía de Investigaciones de Chile, y de acuerdo con la experiencia de sus funcionarios, la mayoría de los hablantes de origen desconocido (delincuentes) eran hombres adultos de entre 18 y 50 años (95%). El estudio –que permite orientar una potencial recolección de bases de datos– también provee información específica importante acerca de las condiciones específicas de

¹⁶ Real Academia Española: Diccionario de la lengua española, 23ª ed. [versión 23.3 en línea]. Disponible en: <https://dle.rae.es>. [Consulta: 2/10/20].

¹⁷ Para una evaluación del desarrollo de las pericias en Chile, a nivel general, ver: Duce (2018), pp. 243-250.

¹⁸ Se envió vía electrónica una encuesta confidencial a profesionales de quienes los investigadores han tenido noticia de su trabajo e interés en el ámbito forense de la voz.

¹⁹ Ver detalles, más adelante, en sección 3.4.2.1.

registro, canal, estilo, variedad dialectal, duración y condiciones ambientales para los audios de origen desconocido y conocido, como veremos enseguida. En forma particular, informa que las grabaciones de origen desconocido corresponden frecuentemente a interceptaciones de conversaciones (94%) en español de Chile, con acento norteño (I y IV regiones) (70%), realizadas desde teléfono celular a celular (95%), con una duración aproximada de menos de 30 s brutos y de entre 30 a 120 s brutos (45% y 45%, en cada caso), y que incluyen ruido de fondo de la calle (tráfico de vehículos) (70%) y de habitación (sistemas de calefacción, ventilación o refrigeración, impresoras, restaurante, televisión o radio) (30%). Por su parte, las grabaciones de hablantes conocidos (sospechosos) corresponden a entrevistas cara a cara realizadas en unidades policiales bajo condiciones controladas (90%), con una duración aproximada de 5 minutos puros de habla, o más, sin ruido de fondo y sin reverberación o reverberación moderada.

3.2.1.1. LAS INTERCEPTACIONES TELEFÓNICAS

La interceptación telefónica, o vulgarmente llamada “pinchazo”, que es el tipo de registro más frecuente constatado por la PDI, es una medida intrusiva que está regulada en el artículo 222 del Código Procesal Penal y que procede en casos muy acotados²⁰, siendo utilizada por el Ministerio Público, el que solicita la autorización correspondiente al tribunal para delitos que merezcan pena de crimen (sanción privativa de libertad que parte en los cinco años y un día, hasta perpetuo calificado)²¹. Una vez que el requerimiento es aceptado por el juez de garantía correspondiente, las policías, o el organismo que esté participando de la investigación, ejecuta la diligencia.

En un artículo publicado por el diario *La Tercera* el 18 de febrero de 2018 se señala que, con información obtenida mediante la Ley de Transparencia²², la Fiscalía Nacional informó que entre 2014 y 2016 se realizaron 59.380 de estas diligencias, de las estas el 86% de las intervenciones las realizó la PDI. En tanto, Carabineros llevó a la práctica el 12% de estas diligencias, y entre Gendarmería y la Dirección General del Territorio Marítimo y Marina Mercante de Chile (Directemar) se reparten el 2% restante.

²⁰ El Art. 222 señala que dicha medida procede cuando “existieren fundadas sospechas, basadas en hechos determinados, de que una persona hubiere cometido o participado en la preparación o comisión, o que ella prepare la comisión o participación en un hecho punible que mereciere pena de crimen, y la investigación lo hiciera imprescindible, el juez de garantía, a petición del ministerio público, podrá ordenar la interceptación y grabación de sus comunicaciones telefónicas o de otras formas de telecomunicación.

²¹ Además del Ministerio Público, basado en la Ley de Inteligencia, un jefe de inteligencia puede solicitar la autorización de esta diligencia a un juez de la Corte de Apelaciones. Ver: Ley N° 19.974 de 2004, Biblioteca del Congreso Nacional de Chile/ BCN. Disponible en: <https://www.bcn.cl/leychile/navegar?idNorma=230999>. [Consulta: 2/2/21].

²² Ley N° 20.285 de 2008, Biblioteca Nacional de Chile/ BCN. <https://www.bcn.cl/leychile/navegar?idNorma=276363>. [Consulta: 2/2/2021].

3.2.1.2. LA POLICÍA DE INVESTIGACIONES DE CHILE Y EL BATVOX

Siendo la Policía de Investigaciones de Chile la Institución que provee del mayor número de peritajes de voz, conviene saber qué metodología ocupa y cuáles son los problemas que ha debido sortear. Esta institución fue la primera en Latinoamérica en identificar científicamente a los delincuentes mediante un sistema automático basado en razones de probabilidad²³. Sin embargo, si bien es cierto que la Sección Sonido y Audiovisual del Laboratorio de Criminalística viene trabajando en sus análisis de voz con relaciones de probabilidad desde el 2006 con la utilización del *software* Batvox –cuyos resultados han sido valorados y comprendidos por una gran cantidad de fiscales y jueces– ha existido siempre de parte de los profesionales a cargo de las pericias una inquietud respecto de ciertas limitaciones metodológicas del sistema.²⁴ Efectivamente, a pesar de sus bondades, uno de los problemas que tiene este sistema, a la luz de los fundamentos de esta rama científica, descritos en la secciones anteriores, tiene que ver sobre todo con no considerar las características de la población relevante específica al caso para establecer la tipicidad, y usar en cambio una población genérica, y de no abordar como algo necesario el problema de los desajustes de las grabaciones, sino más bien como algo accesorio (Morrison y Thompson, 2017). La importancia de medir la tipicidad, a base de las características del caso, es algo que reviste la mayor importancia para el correcto funcionamiento del sistema, lo mismo que la consideración de los desajustes de las grabaciones, previo a cualquier comparación, y así ha sido constatado (Enzinger 2016; Morrison 2018).

3.2.2. CONCLUSIONES PRELIMINARES

Vista la situación general, en términos de profesionales peritos e investigadores en comparación forense de voz, condiciones de casos de trabajo y procedimientos ocupados, podemos concluir al terminar este apartado que el país enfrenta necesidades teóricas y metodológicas, sobre todo asociadas con el uso y difusión de un marco adecuado para la interpretación de resultados –como el marco de la razón de probabilidad presentado–, la incorporación de pruebas empíricas de validez del sistema, bases de datos que sean representativas de los casos de trabajo –a base del conocimiento allegado por los usuarios–, y herramientas de *software* que permitan implementar los procedimientos requeridos. Quizá también sea necesario incorporar a esta lista capacitación en el manejo práctico de herramientas y en la fundamentación de su uso (base para la toma de decisiones). Agregaríamos que nos parece importante y necesario que lo anterior no solo

²³ “La biometría de la voz”. Ciencia y Tecnología. *Revista Detective*. Policía de Investigaciones de Chile. Abril de 2007 / Edición 127, p. 5.

²⁴ “Tratamiento de la evidencia de voz”. *Revista Detective*. Policía de Inestigaciones de Chile. Edición especial 159, aniversario 80, junio, 2013, pp. 68-70.

sea comprendido –y exigido– por quienes tienen una relación directa con los análisis en casos de trabajo, profesionales e investigadores, sino también –y muy en especial– por quienes están en el rol de encargar informes periciales y evaluar la admisibilidad de dichos testimonios, como fiscales, abogados y jueces.

3.4. INVESTIGACIÓN

En este apartado se presenta un breve resumen de la investigación en comparación forense de la voz en nuestro territorio.

En la investigación muchos actores –a veces, sin sospecharlo– realizan esfuerzos de diverso orden en pro de un propósito general que van en forma gradual desbrozando el camino y acercando el objetivo. Podemos distribuir dichos empeños individuales y colectivos, temporal y nocionalmente, en dos etapas: la primera, de estudios generales respecto de evidencia de voz, y la segunda –ahora en desarrollo–, de estudios insertos en el marco de la razón de probabilidad.

3.4.1. ETAPA DE ESTUDIOS GENERALES

Como se constató para el caso de la práctica forense, la comunidad científica nacional abocada al tema del análisis de voz con fines forenses es bastante reducida. De hecho, aparte del grupo interdisciplinar e interinstitucional UACH-PDI²⁵, no hemos logrado encontrar otros grupos trabajando en el tema forense de la voz.

El interés por las ciencias forenses del habla²⁶, y más aún por la rama de comparación forense de la voz, es de data reciente en nuestro país, con un primer estudio correspondiente a la “Variabilidad idiofónica en español como herramienta forense” publicado por Rosas y Sommerhoff en 2006, al que siguieron “Efectos acústicos de las variaciones fonopragmáticas y ambientales” (Rosas y Sommerhoff, 2009), “Implicancias y proyecciones forenses en el análisis de la voz” (Rosas y Sommerhoff, 2010), “Comparación de voz bajo el cociente de probabilidad en el caso de Luis Tralcal (2011)” (Rosas, Sommerhoff, Sáez y Saavedra, 2014), y recientemente “Un método para calcular la fuerza de la evidencia asociada con el supuesto reconocimiento por parte de un testigo auditivo de un hablante conocido”²⁷ (Rosas, Sommerhoff y Morrison, 2019). A ellos se

²⁵ El equipo lo conforman la Dra. Claudia Rosas y el Dr. Jorge Sommerhoff, por la Universidad Austral de Chile, y el Ing. Acústico, César Sáez, por la Policía de Investigaciones de Chile.

²⁶ Algunos estudios sobre otras ramas de las ciencias forenses del habla en nuestro país, se pueden encontrar en los artículos de Rosas, Sommerhoff y Sáez (2018) “Identificación de voces por auditores profanos no nativos”, Rosas, Sommerhoff, Pachecho y Sáez (2020) “Yo lo reconocería por su voz ... El caso de Emilio Berkhoff”, y en la tesis de magíster de Barria, Cristián (2018) “Influencia del tiempo y la experticia en rondas de reconocimiento de voz en acústica forense”.

²⁷ Título original en inglés: “A method for calculating the strength of evidence associated with an earwitness’s claimed recognition of a familiar speaker” (*Science & Justice*, Volume 59, Issue 6, November 2019, pp. 585-596).

suma una interesante aunque no larga lista de tesis de grado o titulación realizados en el país, todas relativas a la evidencia forense con miras a la identificación del hablante, que pueden ser considerados como una primera etapa de acercamiento al tema que nos ocupa, como veremos en las líneas que siguen.

Un número acotado de tesis nacionales de pre y postgrado que han incursionado en el ámbito forense de la voz desde disciplinas diversas, lingüísticas, acústicas y fonaudiológicas confirman la falta de investigación respectiva en dichas áreas. Así, en acústica, Bello (2008) demanda a la academia involucrarse en el trabajo de la Policía de Investigaciones de Chile, Celis (2009)²⁸ suma la falta de una metodología bayesiana, bases de datos que describan la realidad fonética-acústica de la población de referencia (español de Chile) de uso público y evaluación de las metodologías. Estamos hablando de llamados a crear sinergia entre instituciones de investigación y de orden de seguridad, de marcos de interpretación de los resultados objetivos, de bases de datos representativas de la población relevante y de pruebas empíricas de validez de los sistemas de comparación forense de la voz. Por su parte, Castillo *et al.* (2011), desde una perspectiva fonaudiológica, reparan en la falta de consenso respecto de los términos conceptuales y métodos para la identificación vocal. Barría (2009)²⁹ realiza un recuento de los análisis periciales de la voz en Chile entre el 2003-2008 al que añade el sistema judicial y administrativo que regula esta práctica a partir de la reforma procesal penal. Una aproximación a las características de la población relevante en el contexto de casos de trabajos de Chile se encuentra en Farías (2013) y Novoa (2013). Ambos trabajos ocupan para la extracción de características una base de datos de casos cerrados de la Policía de Investigaciones de Chile³⁰.

El origen de los trabajos más sistemáticos en evidencia de voz se inicia formalmente en 2005 con un proyecto interno de la Universidad Austral de Chile, “Estudio fonético-acústico de la variabilidad ambiental idiofónica en español como herramienta forense para validar análisis periciales de voces presentadas en los juicios” (Ref.: DID S-2005-72), en cuya realización intervino un equipo integrado por Claudia Rosas, Jorge Sommerhoff y Víctor Poblete, todos investigadores de la UACh. Consecuencia de la experiencia adquirida en ese primer emprendimiento, entre 2007 y 2014 se desarrollan dos proyectos nacionales: “Pautas para el análisis pericial de voces presentadas en los juicios” (Ref.: Fondecyt Regular N° 1070210), con Claudia Rosas y Jorge Sommerhoff y “Voces en contextos periciales para el reconocimiento automático” (Ref.: Fondecyt Regular N° 1110742)³¹, al que se une

²⁸ En el marco de un proyecto Fondecyt Regular (Ref.: 1070210) con Rosas como Investigadora Responsable.

²⁹ *Ibid.*

³⁰ En el marco de un proyecto Fondecyt Regular (Ref.: N° 1110742) con Rosas como Investigadora Responsable.

³¹ De manera preliminar, se diseñó un proyecto interno entre la PDI-UACH 2010-2011 “Elaboración de un corpus de hablantes para su estudio como población de referencia en la comparación de voz – Mejoramiento 184 | ALPHA N° 55 (DICIEMBRE 2022) PÁGS. 169-191. ISSN 07 16-4254

César Sáez, funcionario del Laboratorio de Criminalística de la Policía de Investigaciones de Chile, como parte de una alianza estratégica de investigación y fuerte vinculación con el medio³².

3.4.2. ETAPA DE APROXIMACIÓN AL NUEVO PARADIGMA CIENTÍFICO

En 2019 el equipo UACH-PDI fue invitado a participar en un proyecto de ciencia forense de la voz de alcance mundial³³. Se trata del proyecto de ciencia forense del habla, *Desarrollo de un sistema de comparación de voz forense, Development of a Forensic Voice Comparison System*³⁴ anclado en el Laboratorio de Ciencias Forenses del Habla (*Forensic Speech Science Laboratory FSSL*) del Instituto Aston de Lingüística Forense (*Aston Institute for Forensic Linguistics AIFL*) de la Universidad de Aston (Reino Unido). Los usuarios finales se han definido como: (1) investigadores y profesionales que utilizarán (potencialmente) el sistema para realizar investigaciones y realizar trabajos de casos, y (2) usuarios del servicio, es decir, organizaciones que encargan a los profesionales que realicen análisis forenses de comparación de voz. Los usuarios potenciales del servicio incluyen abogados defensores y agencias de aplicación de la ley. Si bien las necesidades de los usuarios variarán dependiendo de la procedencia, estarán en la base de los requerimientos contemplados herramientas de *software* que sean de bajo costo, flexibles, y fáciles de usar (controlables por medio de una interfaz gráfica de usuario o que solo requieran habilidades de programación muy limitadas), muy bien documentados, diseñados para facilitar la validación y con código abierto a la inspección (el proyecto prevé código abierto, pero no

de la población de referencia para la optimización del BATVOX en Chile” con el objetivo de conocer cómo funcionaba el Batvox con muestras recolectadas *ad hoc*.

³² La alianza UACH-PDI arranca con el “Convenio específico para el Desarrollo del Conocimiento, Discusión de Métodos, Capacitación y Coloquios sobre el Análisis Pericial de Voces Empleadas en los Juicios”, firmado por el Rector UACH, Sr. Víctor Cubillos y el Director General PDI, Sr. Arturo Herrera (8 de junio de 2009). El reconocimiento extra institucional del trabajo conjunto realizado entre la Policía de Investigaciones de Chile y la UACH ha quedado de manifiesto en diversas instancias, como p.e., en *Revista Detective* (Edición N° 159 Aniversario, Jun 19, 2013, pp. 70-73). Asimismo, la valoración y alcance de la investigación que en la actualidad se desarrolla queda de manifiesto en la propia autorización del jefe Nacional de Criminalística, Sr. prefecto Víctor Arriagada Mauna (comunicación personal, Santiago, 27 de agosto de 2018), que permite tratar información confidencial acerca de sus necesidades, herramientas y prácticas.

³³ Este proyecto se lleva a cabo en colaboración con varias organizaciones socias, que incluyen: Chile, Universidad Austral de Chile, Policía de Investigaciones, colaboradora principal: Dra. Claudia Rosas; China, Universidad del Suroeste de Ciencias Políticas y Derecho, colaborador principal: Prof. Cuiling Zhang; Alemania, Policía Federal Alemana, Bundeskriminalamt (BKA), colaborador principal: Dr. Michael Jessen; Países Bajos, Instituto Forense de los Países Bajos (NFI), colaborador principal: Sr. David van der Vloed; Suecia, Centro Nacional Forense de Suecia (NFC), Colaboradora principal: Sra. Fanny Carlström Plaza; Estados Unidos de América, Oficina Federal de Investigaciones (FBI), Colaborador principal: Sr. David Marks.

³⁴ Información disponible en: <https://www.aston.ac.uk/research/forensic-linguistics/data-science-laboratory>. [Actualizado en: 2020-02-09].

distribución abierta). El progreso general del proyecto incluye como avance lo siguiente: evaluaciones de las necesidades del usuario final para identificar los datos que deben recopilarse y las herramientas de *software*, los protocolos y los programas de capacitación que desarrollan prototipos para herramientas de *software* básicas; y evaluación consolidada de las necesidades del usuario final y borrador de los requisitos funcionales completados y distribuidos a los socios y colaboradores para recibir comentarios. La sección-Chile ha finalizado la primera etapa referida a las evaluaciones de las necesidades de los usuarios. Esto ha implicado, en la práctica primero gestionar la autorización de la Jefatura Nacional de Criminalística de la Policía de Investigaciones de Chile y luego reunir la información correspondiente a las características y condiciones de los casos de trabajo. La siguiente etapa, proyectada originalmente para el periodo de marzo 2020-marzo 2021 es recopilar un conjunto acotado de muestras preliminares de acuerdo con las condiciones de los casos de trabajo documentadas en la primera etapa³⁵ y, enseguida, probar cómo responde el sistema.

3.4.3. CONCLUSIONES PRELIMINARES

El cuadro general de investigación presentado revela un escenario donde han confluído fuerzas de diversa índole: individual y colectiva, nacional e internacional, disciplinar e interdisciplinar, institucional e interinstitucional. Se ha pasado de una fase inicial donde se recababa información de todo tipo de manera descriptiva, a una fase centrada en la inferencia forense bajo el marco lógicamente correcto y científico de la razón de probabilidad. Ahora, en el marco de un proyecto internacional, se trabaja para la recolección de una base de datos que permitirá (a) alimentar una base general de acuerdo con las características más corrientes de los casos de trabajo en nuestro país y (b) probar el sistema de comparación forense de la voz –en desarrollo– con datos representativos de la población relevante de nuestro territorio y, con ello, hacer que la ansiada aspiración de producir comparación forense de la voz en Chile de manera científica comience a ser una realidad.

4. CONCLUSIÓN

Basados en nuestra interpretación de la práctica e investigación de la comparación forense de la voz en Chile, a la luz de los estándares internacionales, podemos situar el desarrollo de esta ciencia en Chile en un promisorio estado de apogeo donde relevamos dos hechos: la relación entre la academia y las Fuerzas de Orden y Seguridad Pública, y su incorporación a un proyecto de comparación forense de la voz, de impacto mundial.

³⁵ Ver las características más corrientes de los casos de trabajo en Chile en sección 3.2.1.

OBRAS CITADAS

- Arriagada Mauna, V. Prefecto, jefe Nacional de Criminalística. (2018). Comunicación personal, Santiago, 27 de agosto.
- Artículos 222; 295; 314-316; 323. En *Código Procesal Penal*. (2020). Biblioteca del Congreso Nacional de Chile/ BCN. Disponible en: <https://www.bcn.cl/leychile/navegar?idNorma=176595> [Consulta: 6/12/20].
- Barría, Cristian (2018). *Influencia del tiempo y la experticia en rondas de reconocimiento de voz en acústica forense*. [Tesis] Facultad de Ciencias de la Ingeniería. Universidad Austral de Chile.
- Barría, P. (2009). *Recuento de los análisis periciales de la voz en Chile 2003-2008*. [Tesis] Facultad de Filosofía y Humanidades. Universidad Austral de Chile.
- La biometría de la voz (2007). *Revista Detective*. Policía de Investigaciones de Chile. Abril de 2007 / Edición 127, p. 5.
- Castillo, Karin; Cruz, Natalia; Escobar, Mauricio y Bustamante, Estela (2011). *Fundamentos de la identificación vocal del español de Chile: una mirada fonaudiológica*. [Tesis] Facultad de Medicina. Universidad de Chile.
- Celis, Pablo (2009). *Descripción de los métodos utilizados en reconocimiento forense de locutores y su implementación en Chile*. [Tesis] Facultad de Ciencias de la Ingeniería. Universidad Austral de Chile.
- Convenio específico para el Desarrollo del Conocimiento, Discusión de Métodos, Capacitación y Coloquios sobre el Análisis Pericial de Voces Empleadas en los Juicios. (8/1/2009). Universidad Austral de Chile UACH-Policía de Investigaciones de Chile PDI.
- Criminal Practice Directions. [Instrucciones de Práctica Criminal] (2015). Pub.19A: *Expert Evidence*. [Consolidado con enmiendas:2019]. Disponible en: <https://www.judiciary.uk/wp-content/uploads/2019/03/crim-pd-amendment-no-8-consolidated-mar2019.pdf> [Consulta: 20/8/20].
- Curso de fonética forense aplicada al español de Chile. Dirigido a funcionarios del Laboratorio de Criminalística, Sección Sonido y Audiovisual de la Policía de Investigaciones de Chile. Valdivia: 15-17 de diciembre de 2010. *Noticias UACH*. <http://noticias.uach.cl/principal.php?pag=noticia-externo&cod=27491> [Consulta: 4/2/2021].
- Dra. Claudia Rosas busca crear laboratorio de comparación forense del habla en Chile. (2015). *Diario UACH*. Disponible en: <https://diario.uach.cl/dra-claudia-rosas-busca-crear-laboratorio-de-comparacin-forense-del-habla-en-chile/>. [Consulta: 2/2/21].
- Duce J., Mauricio (2018). Prueba pericial y su impacto en los errores del sistema de justicia penal: antecedentes comparados y locales para iniciar el debate. *Ius et Praxis*, 24(2), 223-262. <https://dx.doi.org/10.4067/S0718-00122018000200223>

- Farías Álvarez, Felipe Andrés (2013). *Análisis de los espectros nasales de un corpus de casos cerrados de la PDI*. Tesis de Magíster en Acústica y Vibraciones. Facultad de Ciencias de la Ingeniería. Universidad Austral de Chile.
- Federal Rule of Evidence 702 [Regla Federal de la Evidencia 702]. (1975). Pub. [Consolidado con enmiendas: 2011]. Legal Information Institute LII. Disponible en: https://www.law.cornell.edu/rules/fre/rule_702 [Consulta: 20/8/20].
- Forensic Science Regulator FSR [Regulador de Ciencia Forense] (2014). Versión 2.0. *Codes of Practice and Conduct for forensic science providers and practitioners in the Criminal Justice System*. Disponible en: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/351197/The_FS_R_Codes_of_Practice_and_Conduct_-_v2_August_2014.pdf [Consulta: 20/8/20].
- Forensic Speech Science Laboratory FSSL [Laboratorio de Ciencias Forenses del Habla]. Aston Institute for Forensic Linguistics AIFL. Aston University (UK). Disponible en: <https://www.aston.ac.uk/research/forensic-linguistics/data-science-laboratory> [Consulta: 10/2/2020].
- Haraksim, Rudolf; Ramos, Daniel y Meuwly, Didier (2016). Validation of likelihood ratio methods for forensic evidence evaluation handling multimodal score distributions. *IET Biometrics*, 6(2), 61-69.
- Hollien, Harry (2016). An approach to speaker identification. *Journal of Forensic Sciences*, 61, 334-344. <http://dx.doi.org/10.1111/1556-4029.13034>
- (2002). *Forensic Voice Identification*. Academic Press.
- International Standard Organization / International Electrotechnical Commission. ISO/IEC [Organización Internacional de Normalización/Comisión Electrónica Internacional] 17025 *Requisitos generales para la competencia de los laboratorios de ensayo y calibración*. Disponible en: <https://www.iso.org/standard/39883.html> [Consulta: 22-11-20].
- Kersta, L. G. (1962). Voiceprint identification. *Nature*, 196, 1253-1257. <https://doi.org/10.1038/1961253a0>
- Ley N° 19.974 de 2004. *Sobre el Sistema de Inteligencia del Estado y Crea la Agencia Nacional de Inteligencia*. Biblioteca del Congreso Nacional de Chile/ BCN. <https://www.bcn.cl/leychile/navegar?idNorma=230999> [Consulta: 2/2/21].
- Ley N° 20.285 de 2008. *Sobre Acceso a la Información Pública*. Biblioteca Nacional de Chile/ BCN. <https://www.bcn.cl/leychile/navegar?idNorma=276363> [Consulta: 2/2/2021].
- Meuwly, Didier (2001). *Reconnaissance de locuteurs en sciences forensiques: l'apport d'une approche automatique*. PhD dissertation. University of Lausanne.
- Morrison, Geoffrey Stewart (2019-2022). *Development of a forensic voice comparison system/Desarrollo de un sistema de comparación forense de la voz*. GBP 5.4 M grant from Research England's Expanding Excellence in England (E3) programme.

- Disponible en: <https://www.aston.ac.uk/research/forensic-linguistics/data-science-laboratory> [Consulta: 10/2/2020].
- (2014). Distinguishing between forensic science and forensic pseudoscience: Testing of validity and reliability, and approaches to forensic voice comparison. *Science & Justice*, 54, 245–256. <http://dx.doi.org/10.1016/j.scijus.2013.07.004>
- (2011). Measuring the validity and reliability of forensic likelihood-ratio systems. *Science & Justice*, 51, 91–98. <http://dx.doi.org/10.1016/j.scijus.2011.03.002>
- (2009). Forensic voice comparison and the paradigm shift. *Science & Justice*, 49, 298–308. <http://dx.doi.org/10.1016/j.scijus.2009.09.002>
- Morrison, Geoffrey Stewart; y Enzinger, Ewald (2018). Score based procedures for the calculation of forensic likelihood ratios—Scores should take account of both similarity and typicality. *Science & Justice*, 58, 47–58. <http://dx.doi.org/10.1016/j.scijus.2017.06.005>
- (2016). What should a forensic practitioner’s likelihood ratio be? *Science & Justice*, 56, 374–379. <http://dx.doi.org/10.1016/j.scijus.2016.05.007>
- Morrison, Geoffrey Stewart y Thompson, William C. (2017). Assessing the admissibility of a new generation of forensic voice comparison testimony. *Columbia Science and Technology Law Review*, 18, 326–434.
- Morrison, Geoffrey Stewart; Enzinger, Ewald y Zhang, Cuiling (2018). Forensic Speech Science. En I. Freckelton, & H. Selby (Eds.), *Expert Evidence* (Ch. 99). Sydney, Australia: Thomson Reuters.
- National Commission on Forensic Science [Comisión Nacional de Ciencias Forenses] (2016). Views of the Commission Use of the Term “Reasonable Scientific Certainty”. <https://www.justice.gov/archives/ncfs/file/839726/download> [Consulta: 20/8/2020].
- National Research Council NRC [Consejo Nacional de Investigación]. (2009). *Strengthening Forensic Science in the United States: A Path Forward*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/12589>. [Consulta: 20/8/20].
- Novoa, José Eduardo (2012). *Análisis espectral de los segmentos vocálicos de un recuento de voces de la PDI (Policía de Investigaciones de Chile)*. [Tesis] Facultad de Ciencias de la Ingeniería. Universidad Austral de Chile.
- La PDI realiza el 86% de las interceptaciones telefónicas en el país (2018). *La Tercera*. Disponible en: <https://www.latercera.com/nacional/noticia/la-pdi-realiza-86-las-interceptaciones-telefonicas-pais/72194/> [Consulta: 20/8/20].
- Pérez, Eduardo (2008). *Ingeniería acústica aplicada a la criminalística*. [Tesis] Facultad de Ciencias de la Ingeniería. Universidad Austral de Chile.
- President’s Council of Advisors on Sci.&Tech. PCAST [Consejo de Asesores de Ciencia y Tecnología] (2016). *Forensic Science in Criminal Courts: Ensuring Scientific*

- Validity of Feature-Comparison Methods*. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf [Consulta: 20/8/20].
- Rosas, Claudia; Sommerhoff, Jorge (2010). Implicancias y proyecciones forenses en el análisis de la voz. *Estudios filológicos*, (46), 101-118. <https://dx.doi.org/10.4067/S0071-17132010000200006>
- (2009). Efectos acústicos de las variaciones fonopragmáticas y ambientales. *Estudios filológicos*, (44), 195-210. <https://dx.doi.org/10.4067/S007117132009000100012>.
- (2007-2010). *Pautas para el análisis pericial de voces presentadas en los juicios*. Fondecyt Regular N° 1070210.
- (2006). Variabilidad idiofónica en español como herramienta forense. *Estudios filológicos*, (41), 235-249. <https://dx.doi.org/10.4067/S0071-17132006000100014>
- Rosas, Claudia; Sommerhoff, Jorge y Poblete, V. (2005-2006). *Estudio fonético-acústico de la variabilidad ambiental idiofónica en español como herramienta forense para validar análisis periciales de voces presentadas en los juicios*. Proyecto Interno DID S-2005-72. Universidad Austral de Chile.
- Rosas, Claudia; Sommerhoff, Jorge; y Sáez, César (2018). Identificación de voces por auditores profanos no nativos. *Revista Alpha* (Osorno), (46), 129-150. <https://dx.doi.org/10.4067/S0718-22012018000100129>
- (2011-2014). *Voces en contextos periciales para el reconocimiento automático*. Fondecyt Regular N° 1110742.
- (2010-2011). *Elaboración de un corpus de hablantes para su estudio como población de referencia en la comparación de voz. Mejoramiento de la población de referencia para la optimización del BATVOX en Chile*. Proyecto de Investigación PDI-UACH.
- Rosas, Claudia; Sommerhoff, Jorge; y Morrison, Geoffrey Stewart (2019). A method for calculating the strength of evidence associated with an earwitness's claimed recognition of a familiar speaker. *Science & Justice*. 59: 585-596. DOI: 10.1016/j.scijus.2019.07.001
- Rosas, Claudia; Sommerhoff, Jorge; Pacheco, Jaime; y Sáez, César (2020). Yo lo reconocería por su voz... El caso de Emilio Berkhoff. *Revista Alpha* (Osorno), (51), 137-160. <https://dx.doi.org/10.32735/s0718-2201202000051851>.
- Rosas, Claudia; Sommerhoff, Jorge; Sáez, César; y Saavedra, Sebastián (2014). Comparación de voz bajo el cociente de probabilidad en el caso de Luis Tralcal (2011). *RLA. Revista de lingüística teórica y aplicada*, 52(1), 13-33. <https://dx.doi.org/10.4067/S0718-48832014000100002>

- Rose, Phil (2017). Likelihood ratio-based forensic voice comparison with higher level features. Research and reality. *Computer Speech & Language* 45, September, 475-502. <https://doi.org/10.1016/j.csl.2017.03.003>
- (2006). Technical forensic speaker recognition. *Computer Speech and Language*, 20, 159-191. <https://doi.org/10.1016/j.csl.2005.07.003>
- Tratamiento de la evidencia de voz. (2013). *Revista Detective*. Policía de Investigaciones de Chile. Edición especial 159, aniversario 80, junio, 68-70.
- Vera Lama, Rodrigo (2017). *Sistema de Inteligencia chileno*. Metropolitana.