



Estudios Filológicos

ISSN: 0071-1713

efil@uach.cl

Universidad Austral de Chile  
Chile

Rosas, Claudia; Sommerhoff, Jorge  
Implicancias y proyecciones forenses en el análisis de la voz  
Estudios Filológicos, núm. 46, 2010, pp. 101-118  
Universidad Austral de Chile  
Valdivia, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=173418807006>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica  
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

## Implicancias y proyecciones forenses en el análisis de la voz\*

Forensic implications and projections of voice analysis

*Claudia Rosas<sup>1</sup>, Jorge Sommerhoff<sup>2</sup>*

<sup>1</sup> Universidad Austral de Chile, Facultad de Filosofía y Humanidades, Instituto de Lingüística y Literatura, Valdivia, Chile. Correo electrónico: claudiarosas@uach.cl

<sup>2</sup> Universidad Austral de Chile, Facultad de Ciencias de la Ingeniería, Instituto de Acústica, Valdivia, Chile. Correo electrónico: jsommerh@uach.cl

Este artículo describe los efectos de las condiciones fonopragmáticas y ambientales sobre el F0 y los formantes y señala la falta de datos empíricos y/o respaldo teórico de apoyo a lo individual en la identificación forense del locutor. A continuación se describe una estrategia general de investigación para situar la labor pericial en un marco de actuación confiable.

*Palabras clave:* análisis de F0, análisis de formantes, diferentes canales de transmisión, diferentes situaciones fonopragmáticas, identificación forense del hablante.

This article describes the effects of the phonopragmatic and environmental conditions on the F0 and formants, and notes the lack of empirical data and/or theoretical support to hold up particular aspects in forensic speaker identification. The following describes a general strategy research to position the expert work within reliable performance framework.

*Key words:* F0 analysis, formant analysis, different transmission channels, different phonopragmatic situations, forensic speaker identification.

1. Hace un par de años, las dificultades surgidas en el ejercicio de algunos peritajes forenses de la voz nos llevaron a buscar un respaldo teórico y metodológico para emitir informes con seguridad: descubrimos un panorama desprovisto de estudios sistemáticos que indicaran los pasos a seguir en el análisis del habla: a qué indicadores se debía prestar atención y en qué medida. Nos preocupaba saber ¿hasta qué punto los análisis de la voz en Chile con fines forenses habían sido practicados con rigurosidad metodológica? Nuestras inquietudes nos introducían en un tema complejo: la identificación del hablante.

Así pues, conscientes de la posición social del hecho, de la falta de criterios para el análisis pericial y de la asumida importancia del trabajo interdisciplinario fonético-acústico se nos impuso la necesidad de comenzar una investigación, con

---

\* Los datos de este trabajo proceden del proyecto de investigación FONDECYT Regular N° 1070210.

el fin de contribuir a la dotación de las pautas necesarias para el análisis de voces grabadas presentadas como pruebas en los juicios, cuyos resultados forman parte de este artículo (Rosas y Sommerhoff 2009a).

Para comenzar el estudio concebimos la siguiente hipótesis: en el español de Chile existe variación acústica idiolectal producida por factores ambientales y expresivos, posible de reconocer, representar y jerarquizar. Bajo esas premisas se definieron tres objetivos generales: (a) caracterizar la variación acústica idiolectal en el español de Chile producida por factores expresivos y ambientales, (b) contribuir al conocimiento de los especialistas de fonética y acústica con el fin de mejorar las aplicaciones en el ámbito forense, y (c) contribuir al desarrollo de la labor investigadora interdisciplinaria fonético-acústica; y dos objetivos específicos: (a) describir las propiedades acústicas segmentales y suprasegmentales idiolectales de carácter expresivo y ambiental en el español de Chile y (b) jerarquizar la variación acústica idiolectal expresiva y ambiental de los parámetros segmentales y suprasegmentales seleccionados en la variante chilena del español.

Por su parte, el corpus analizado estuvo integrado por 160 vocales extraídas de frases de estructura simple de tipo artículo + sustantivo, procedentes de la base de datos del proyecto de investigación (Ref.: N° 1070210 Fondecyt Regular 2007), subvencionado por la Comisión Nacional de Investigación en Ciencia y Tecnología, Gobierno de Chile (CONICYT). Los segmentos emitidos corresponden a una vocal "a" tónica inserta en ocho contextos fonéticamente balanceados, como se detalla a continuación: precedida de /p/, /t/, /k/ y seguida de /s/ (el *paso*, la *taza*, la *casa*); precedida de /s/ y seguida de /p/, /k/ (el *sapo*, el *saco*); precedida y seguida de nasal /n/ (la *nana*); y finalmente, precedida y seguida de /p/, /t/ (las *papas*, el *tata*). Primeramente, las frases fueron producidas y registradas en canal anecoico por cuatro actrices, quienes simulaban estar inmersas en distintas situaciones fonopragmáticas, de lectura, evidencia, asombro, atenuación y no-conclusión. Terminada esta etapa, se procedió a su reproducción y registro en canal telefónico, celular y reverberante.

Para el registro electroacústico se utilizó un micrófono Shure SM 48, la interfaz analógica-digital KayPENTAX, modelo CSL 4500 (Computerized Speech Lab), un PC Pentium con disco duro y grabador de CD y un sistema reproductor NTI modelo Talk Box.

La cadena electroacústica comienza en una sala anecoica con cada locutora que emite frente al micrófono las frases del corpus, que son registradas por el CSL 4500 y almacenadas en el disco duro del PC; cada una, junto a una señal digital de sincronización audible de 1000 Hz y de 0,01 segundos de longitud. A continuación, se generan las señales de tipo reverberante, telefónico y celular. El procedimiento seguido para la obtención de las distintas señales es el siguiente:

*Señal reverberante.* Se reproduce la señal grabada en el CD, utilizando el Talk Box; la señal es captada, primero, por el micrófono ubicado en el campo reverberante, registrada luego por el CSL y finalmente almacenada en el disco duro del PC.

*Señal telefónica.* Se requiere dos salas para su reproducción: la sala 1 que es la de emisión de la señal, y la sala 2, que es la de su recepción. En la sala 1, la señal es grabada en el CD, reproducida, primero, por el Talk Box, captada luego por el micrófono del teléfono y, por último, transmitida por línea telefónica a la sala 2. En la sala 2 se instala el micrófono pegado al auricular del teléfono donde se procede a su registro y almacenamiento por medio del CSL.

*Señal celular.* Ocurre de forma semejante al caso anterior, con la excepción de que el elemento receptor no es un teléfono normal, sino que uno celular.

Para el análisis, los datos de las grabaciones fueron procesados en un CSL modelo 4500 Kay-Pentax con una frecuencia de muestreo de 22050 Hz y se aplicó un análisis de FFT de Cepstrum de 512 puntos para medir el F0 y los formantes, previa localización del valor central de cada uno de los segmentos vocálicos analizados. Los resultados que se obtienen con la aplicación de esta técnica no determinan de manera automática los valores de F0 y formantes: es necesario observar directamente dentro de los datos numéricos de la frecuencia de respuesta los máximos de energía, consignando los centros de frecuencia. Para ello se consideraron los máximos de intensidad que precedían un valor negativo y cuya intensidad coincidiera o superara el valor de la media. Posteriormente, para excluir valores aberrantes se aplicó a la totalidad de datos extraídos la regla empírica del límite superior de la expresión  $\text{media} + - 3\text{desv}$ .

Se aplicaron 1.280 análisis acústicos sin considerar en este cálculo los análisis auditivos que acompañaron todo el proceso con diferentes objetivos (validación de procedimientos automáticos y semiautomáticos, resolución de puntos críticos de mediciones no automáticas, entre otros), ni los análisis relativos a los sondeos preliminares, anteriores a la definición del corpus definitivo (Rosas y Sommerhoff 2008a), como tampoco los análisis auxiliares, practicados con el propósito de complementar el perfil de cada locutora. Conviene destacar, además, que si bien el programa utilizado (Main Program del CSL Kay-Pentax Computer Speech Lab CSL Model 4500 versión 3.1.1.) permite realizar automatizaciones, todas ellas requieren, como es sabido, indefectiblemente, de un proceso visual-auditivo y manual previo. Así, se efectuaron –solo para la presentación de los análisis de donde derivan los resultados que ahora se presentan– más de 4.000 aplicaciones –no automáticas– para medir los periodos de voz que interesaba segmentar a fin de establecer, consecuentemente, las marcas de sincronización acústica que exigía el análisis. También se aplicó un tratamiento estadístico a los datos obtenidos para determinar las relaciones de significancia mediante el SPSS para Windows en su versión 11.5.

Los análisis efectuados nos permitieron obtener dos tipos de información, a saber: central, referida a los indicadores que rigen la investigación, el F0 y los formantes y auxiliar, de interés complementario a esas pesquisas.

2. A continuación se describen de manera sintética los resultados obtenidos, referidos, en primer lugar, a la *información central*; los cambios que introducen los distintos canales y modalidades fonopragmáticas de lectura, evidencia, asombro, atenuación y no conclusión (en adelante, L, E, As, At y NC, respectivamente) sobre el F0 (frecuencia del fundamental) y los formantes (F1, F2, F3, F4) en las 4 locutoras (en adelante, L1, L2, L3 y L4); y luego, más resumidamente, algunos datos sobre el comportamiento del LTF0 (frecuencia del fundamental a largo plazo), extraídos de la *información auxiliar*, complementaria.

*Información central.* Las tendencias generales –que expresan la orientación general de las muestras versus las manifestaciones particulares–, cuando ocurren, se acompañan del grado de significancia, si corresponde. Para obtener una visión más detallada se pueden consultar las relaciones de significancia que aparecen junto a las Tablas de resultados de F0 y Formantes (Anexo).

*F0.* El canal reverberante aumenta y disminuye: L1 disminuye en L, E, As y At y se mantiene en NC; L2 disminuye en todas las modalidades; L3 aumenta en E, As y NC y disminuye en L y At; y L4 aumenta en L, E, As y At y disminuye en NC.

El canal telefónico aumenta en todas las modalidades. Es significativa parcialmente (solo en L1 y L2).

El canal celular aumenta en todas las modalidades, excepto en el caso de L4 que disminuye en NC y aumenta en L, E, As y At.

Se comprobó, además, que *todas las modalidades introducen variaciones en el F0*. Al compararlas con la modalidad de lectura, se observó un aumento de ellas, aunque no es regular ni significativo para todas: en L1 todas las modalidades disminuyen; y en L4 disminuye la modalidad de NC.

*F1.* El canal reverberante disminuye, excepto en L1 que aumenta en dos modalidades: E y NC.

El canal telefónico aumenta y disminuye: L1 aumenta en E y NC y disminuye en L, As y At; L2 aumenta en E, As y NC y disminuye en L y At; L3 aumenta en E, As y NC y disminuye en L y At; y L4 aumenta en todas las modalidades.

El canal celular aumenta y disminuye: L1 disminuye en L, E, As, At y NC; L2 aumenta en E y NC y disminuye en L, As y At; L3 aumenta en As y NC y disminuye en L, E y At; y L4 aumenta en At y NC y disminuye en L, E y As.

*F2.* El canal reverberante presenta una tendencia a la disminución, aunque la orientación del cambio no es sistemática. Todas las locutoras presentan aumento en algunas de las modalidades: L1 en E, L2 en NC, L3 en NC, y L4 en NC.

El canal telefónico presenta una tendencia a la disminución, aunque la orientación del cambio no es sistemática. L1, L3 y L4 presentan aumentos en algunas de las modalidades: L1 en E, L3 en As y NC, y L4 en E y NC.

El canal celular disminuye en todas las modalidades. No es significativo para todas las locutoras en su conjunto, ni para todas las modalidades de una locutora en particular.

*F3.* El canal reverberante presenta una tendencia a la disminución, aunque la orientación del cambio no es sistemática. L1, L2 y L4 presentan aumentos en algunas de las modalidades: L1 en E, A y NC, L2 en At y L4 en NC.

El canal telefónico disminuye en todas las modalidades. Es significativo para todas las locutoras, exceptuando el caso de L4 en su modalidad de NC.

El canal celular disminuye en todas las modalidades, excepto en L1 que presenta aumento en la modalidad de At.

*F4.* El canal reverberante presenta una tendencia a la disminución, aunque la orientación del cambio no es sistemática. L1 y L2 presentan aumentos en algunas modalidades: L1 en E y At y L2 en E y NC.

El canal telefónico disminuye en todas las modalidades. Es totalmente significativo solo en L1.

El canal celular presenta una tendencia a la disminución, aunque la orientación del cambio no es sistemática. L1 y L3 presentan aumentos en algunas de sus modalidades: L1 en E, At y L3 en NC.

En cuanto a la relación entre la modalidad de lectura y el resto de las modalidades, se constató lo siguiente:

*F1.* Todas las modalidades disminuyen, excepto As y At en L1, y E en L4, que se encuentran por encima.

F2. E disminuye en todos los casos, excepto en L3; As aumenta en todos los casos, siendo significativa en L1, L2 y L4; At aumenta en todos los casos, siendo significativa solo en L4; NC disminuye en todos los casos, excepto en L1.

F3. E disminuye en todos los casos, excepto en L4. As aumenta en L1 y L3 y disminuye en L2 y L4. At aumenta en todos los casos, excepto en L2. NC aumenta en L2 y L3 y disminuye en L1 y L4.

F4. E disminuye en todas las modalidades. Es parcialmente significativa (solo en L2 y en L4). As disminuye en todas las modalidades, excepto en L1. At aumenta en L1 y L4 y disminuye en L2 y L3. NC disminuye en todas las modalidades, excepto en L1.

*Información auxiliar.* En cuanto a los datos derivados de los análisis auxiliares, el *LTF0* demostró ser el parámetro más rentable para complementar de manera objetiva los datos del análisis individual de cada locutora; es decir, para colaborar en la determinación de lo sistemático e individual y rechazo de lo casual e individual. En ese sentido, se observó que el aumento del *LTF0* aumentó, sistemáticamente, en todos los casos (entre un 17,84 y un 8,15 por ciento), y que, exceptuando el caso de L2, ese cambio estaba en una relación directamente proporcional con el aumento del F0 temporal relativo. Así, por ejemplo, la locutora con el promedio de F0 temporal más alto también registró el promedio de *LTF0* superior (193 y 230 Hz, respectivamente).

3. El examen de los datos presentados hasta aquí nos permite extraer una serie de características que resumimos en los siguientes párrafos.

La exposición de *los resultados relacionados con la información central* muestra que las situaciones fonopragmáticas y ambientales (canales de transmisión) causan variaciones en las propiedades acústicas de la señal (Rosas y Sommerhoff 2008b, 2009a, 2009b y 2009c). Los efectos que estos factores producen son complejos, ya que se relacionan con distintos aspectos: parámetro estudiado sobre el que actúan (F0, F1, F2, F3 y F4) y orientación del cambio (aumento versus disminución). También son variables, porque no se desarrollan de la misma manera en cada uno de los diferentes canales (reverberante, telefónico y celular) y modalidades fonopragmáticas estudiadas (lectura, evidencia, asombro, atenuación y no conclusión).

Una conclusión destacada derivada de *los resultados auxiliares con fines complementarios* reveló la importancia de agregar a los parámetros acústicos estudiados (F0 y formantes) la frecuencia fundamental de largo plazo (*LTF0*). Este hecho valida la decisión metodológica inicial de considerar un corpus diferenciado que contuviera, además de los segmentos de laboratorio –de características fonéticas y ambientales controladas–, emisiones espontáneas de características no controladas y de mayor duración, a fin de permitir la observación de este tipo de parámetro con fines comparativos, cuya efectividad en la determinación de lo individual y sistemático versus lo individual y casual deberá evaluarse a partir de un número mayor de datos.

Por otra parte, una comparación con estudios previos, similares al nuestro, aunque con diferencias metodológicas, sugiere algunas coincidencias. Principalmente, los efectos del teléfono coinciden con los datos observados por Künzel (2001) para el alemán, aunque no de manera significativa (Rosas y Sommerhoff 2009c).

Con respecto a los objetivos e hipótesis que sostiene esta investigación, se pudo concluir que, si bien los objetivos generales y específicos fueron satisfechos, dentro de los límites propuestos, es imprescindible una etapa de consolidación estadística de

los resultados obtenidos. Efectivamente, el comportamiento del F0 y de los formantes permitió caracterizar la forma en que esos parámetros cambiaban frente a diversas condiciones fonopragmáticas y ambientales, mediante la configuración de tendencias –aumentos o disminuciones generalizadas–; sin embargo, ninguna de las tendencias observadas llegó a ser significativa para la totalidad de las locutoras. En este sentido, se descubrió el surgimiento de un nuevo objetivo de carácter más puntual, pero cuantitativamente mayor, que sobrepasa los límites del presente estudio, pero que habrá que atender, y sobre el que ya se ha empezado a trabajar (Rosas y Sommerhoff 2010): crear un corpus con el propósito de implementar una base estadística para usos alternativos clásicos y automáticos.

Por su parte, con respecto a las premisas que conforman la hipótesis se puede afirmar que cada una de ellas ha sido confirmada, dentro de los límites impuestos, de manera que es posible continuar nuestra investigación u otra de carácter similar, manteniendo esa base.

4. Al finalizar este apartado, deseamos subrayar la importancia de considerar la variabilidad comentada en este estudio para el análisis pericial de voces del español de Chile, donde el contexto de los estudios forenses del habla no es comparable aún al de otras regiones de Norteamérica y Europa (Boë 2000), e introducir una breve reflexión sobre las proyecciones del análisis forense de la voz, a la luz de la variabilidad comentada y de nuestra propia experiencia científica como investigadores y peritos. Para ello, nos centraremos en tres aspectos: los métodos de identificación, el objetivo del perito y la función del fonetista.

*Los métodos de identificación.* Más allá del carácter complementario de todos los métodos, técnicas y herramientas que se puedan utilizar en las tareas de identificación del hablante (Nolan 2001), y que nadie discute, conviene tener presente a la luz de los hechos la variabilidad intrínseca asociada a todos ellos, esto es, el margen de error a que puede inducir cada procedimiento. Aunque autores de la talla de Künzel (2001) lo han observado, el fenómeno no ha recibido la atención que requiere. Esta variabilidad inherente algunas veces desorienta al investigador, más preocupado en la descripción del objeto –como sería deseable que ocurriera– que en el instrumento que ha elegido para su medición. En nuestras mediciones, empeñados en descubrir la intravariabilidad subyacente a los fenómenos analizados, observábamos con cierta desazón y angustia cómo los valores cambiaban según las diferentes técnicas de análisis utilizadas (Rosas y Sommerhoff 2009c), y ese hecho, que no podíamos ocultar, se convertía en una traba adicional, de la que no habíamos sido avisados.

Relacionado con lo anterior se encuentra el sesgo que puede introducir el método acústico o auditivo utilizado para medir el grado de inteligibilidad y, en consecuencia, determinar su carácter de analizable. A menudo, quienes defienden los métodos más tradicionales y exageran las ventajas del método auditivo –cuando las condiciones de la señal han sido degradadas e impiden un análisis acústico– olvidan que estas ventajas son de igual manera, perceptivamente, limitadas (Nolan 2001). Al respecto, una experiencia previa, donde aplicábamos un test de articulación a estudiantes universitarios para medir la inteligibilidad de la palabra, nos permitió comprobar cómo logatomos aislados de toda interferencia ambiental (ruido de fondo) eran percibidos de distinta manera por los auditores, a pesar de que habían sido entrenados para ello

y conocían las palabras del test (Sommerhoff y Rosas 2007; Rosas y Sommerhoff 2008c). Experiencias como éstas deben hacernos tomar más en serio el componente subjetivo que está presente en el investigador, de cómo algunos fenómenos cognitivos, grado de experticia, conocimiento previo (Schiller y Köster 1998) y emocionales, entre otros, pueden alterar el juicio de la evidencia; es decir, la percepción de los hechos y su valoración (Saks 2005).

*El objetivo del perito.* En cuanto al objetivo del perito forense, normalmente su tarea se focaliza en la comparación de características entre una voz conocida o dubitada (la del criminal) y otra que no lo es o indubitada (la del sospechoso); básicamente, en hallar similitudes o diferencias. El fenómeno encierra una gran complejidad. Primeramente, suponiendo que se encuentran rasgos parecidos, segmentos o suprasegmentos, válidamente comparables –lo que ya es hasta cierto punto irregular (Kredens *et al.* 1998), dadas la consabidas condiciones forenses de la muestras que se obtienen (Sommerhoff y Rosas 2003a, 2003b, 2004a y 2004b)–, es necesario plantearse la pregunta acerca del valor que esos rasgos tienen. Como señalan Romito y Galatà 2004, es preciso conocer el código al que se adscribe cada muestra. Esos datos no se pueden presuponer. El investigador tiene en sus manos una sustancia y debe indagar el sistema funcional de pertenencia (el registro, dialecto y sociolecto). Ello implica conocer en profundidad los varios subsistemas de cada lengua común en cuestión: las formas que los diferentes valores comunicativos adquieren en los distintos contextos, situaciones y registros, lo que entraña obviamente una gran dificultad. En este punto los estudios dialectales y sociolingüísticos son imprescindibles, porque pueden proporcionar un antecedente más a la valoración de la evidencia. En Chile, los estudios dialectales –especialmente éstos– y geolingüísticos pueden realizar un interesante aporte, alertando al investigador acerca de la peculiaridad de ciertos fenómenos y de su distribución geográfica y –presumiblemente– social (Wagner *et al.* 1994; Wagner y Rosas 2003).

Una vez aceptada la compatibilidad formal y funcional de las muestras, el segundo escollo que surge es el de sopesar qué tan significativas son esas semejanzas. Por una parte, un enfoque bayesiano, adoptado por los sistemas automáticos de reconocimiento del locutor (SRAL) (Robertson y Vignaux 1995; González-Rodríguez *et al.* 2003), ha hecho ver cómo esas diferencias no pueden ser vistas de manera tan simple y la conveniencia de que sean enfocadas desde una perspectiva estadística (Broeders 1999). Aunque, como señala Nolan (2001), tal enfoque no esté todavía en condiciones de ser aplicado a la identificación del hablante ofrece un marco conceptual adecuado para tratar el problema. En efecto, el grado de certeza alcanzado por los sistemas biométricos del habla está muy lejos de alcanzar la precisión que se logra en otras áreas, y que por cierto tampoco es total (Moore 2007). De cualquier forma, el progreso de los sistemas biométricos de la voz requiere un indiscutido compromiso con los aspectos lingüísticos, encaminado a elaborar bases de datos más ricas, aunque su incorporación desafíe, de hecho, la ingeniería de los sistemas de modelado de la voz (Benzeghiba *et al.* 2007).

*La función del fonetista.* De lo anterior se desprende la importancia del fonetista como interfaz entre los conocimientos lingüísticos, acústicos y auditivos. Efectivamente, el fonetista forense se convierte en el primordial y último eslabón frente a la sustancia fónica al tener que aplicar en ella todo lo que sabe de acústica, de percepción auditiva, de fonopragmática, etc.



En síntesis, frente a lo reseñado, nuestra postura es hacer acopio de las experiencias obtenidas en otras sociedades vanguardistas –propiciadas quizá por el mayor peso que la evidencia de voz allí tiene– y profundizar en el conocimiento de las herramientas objetivas y subjetivas, en sus ventajas y desventajas, haciendo uso de ellas con cautela, en orden a describir la muestra desde una perspectiva que considere su relación con el sistema al que pertenece, en un marco contextual amplio, donde se conecten, fundamentalmente, lo lingüístico con lo acústico y lo pragmático bajo la mirada rectora del fonetista.

#### OBRAS CITADAS

- Benzeghiba, M. *et al.* 2007. "Automatic speech recognition and speech variability: a review". *Speech Communication* 49, 10-11: 763-786.
- Boë, Louis-Jean. 2000. "Forensic voice identification in France". *Speech Communication* 31, 2-3: 205-224.
- Broeders, A. 1999. "Some observations on the use of probability scales in forensic identification". *Forensic Linguistics* 6, 2: 228-241.
- González-Rodríguez, J. *et al.* 2003. "Forensic identification reporting using automatic speaker recognition systems". *Proceedings of ICASSP'03*, 2, Hong Kong (China). 93-96.
- Künzel, H. J. 2001. "Effects of voice disguise on speaking fundamental frequency". *International Journal of Speech, Language and the law (formerly Forensic Linguistics)*, VII, 2: 149-179.
- Kredens, K. y Goralewska-Lach, G. 1998. "Language as Sole Incriminating Evidence: The Augustynek Case". *Forensic Linguistics*, 1998, 5, 2: 193-202.
- Moore, Ariana-Michele. 2007. "Biometric technologies – an introduction". *Biometric Technology Today* 15, 1: 6-7.
- Nolan, F. 2001. "Speaker identification evidence: its forms, limitations, and roles". *Proceedings of the conference "Law and Language: Prospect and Retrospect"*, diciembre 12-15, Levi (Finnish Lapland).
- Robertson, B. y G.A. Vignaux. 1995. *Interpreting Evidence. Evaluating Forensic Science in the Courtroom*. Cheicester, Reino Unido: John Wiley & Sons.
- Romito, L. y V. Galatà. 2004. "Towards a protocol in speaker recognition analysis". *Forensic Science International* 146S: S107-S111.
- Rosas, C. y J. Sommerhoff. 2008a. "La variación idiolectal fonopragmática y ambiental: ejemplos del español de Chile", *IV Congreso de Fonética Experimental Granada CFE*, febrero 11-14, Granada (España).
- \_\_\_\_\_. 2008b. "La variación acústica del hablante: aplicaciones forenses". *Actas del VI Congreso Iberoamericano de Acústica FIA*, noviembre 5-7, Buenos Aires (Argentina). 25.
- \_\_\_\_\_. 2008c. "Inteligibilidad acústica en español: una propuesta para su medición". *Estudios Filológicos* 43: 179-190.
- \_\_\_\_\_. 2009a. "The variations of F0, F1 and F2: examples in the Spanish language of Chile". *9th Biennial Conference on forensic Linguistics/Language and Law*, Julio 6-9, Amsterdam (Holanda).
- \_\_\_\_\_. 2009b. "Efectos acústicos de las variaciones fonopragmáticas y ambientales". *Estudios Filológicos* 44:195-210.
- \_\_\_\_\_. 2009c. "La variación del F0, F1 y F2". *XVIII Congreso de la Sociedad Chilena de Lingüística. Los estudios de lingüística hacia el bicentenario*, noviembre 4-6, Santiago (Chile).

- . 2010. “Diseño y evaluación de poblaciones de referencia con fines forenses”, *Segunda Reunión Panamericana e Ibérica de Acústica* 2010, noviembre 15-19, Cancún (México).
- Saks, M. y J. J. Koehler. 2005. “The coming paradigm shift in forensic identification science”. *Science* 309, 5736: 892-895.
- Schiller, N.O. y O. Köster. 1998. “The ability of expert witnesses to identify voices: A comparison between trained and untrained Listeners”. *Forensic Linguistics* 5: 1-9.
- Sommerhoff, J. y C. Rosas. 2003a. Informe pericial Defensoría Penal Pública de Caldera y Chañaral.
- . 2003b. Informe pericial Defensoría Penal Pública La Serena.
- . 2004a. Informe pericial Defensoría Penal Pública La Serena.
- . 2004b. Informe pericial Defensoría Penal Pública Coronel-Lota.
- . 2007. “Evaluación de la inteligibilidad del habla en español”. *Estudios Filológicos* 42: 215-225.
- Wagner, C. *et al.* 1994. “Chile en el Atlas Lingüístico de Hispanoamérica”. *Estudios Filológicos* 29: 15:24.
- Wagner, C. y C. Rosas. 2003. “Geografía de la “ll” en Chile”. *Estudios Filológicos* 2003 38: 189-200.

ANEXO  
MUESTRA DE RESULTADOS. INFORMACIÓN CENTRAL. LOCUTORA I (L1)

Tabla 1  
F0 por Modalidad y Canal. Valores expresados en Hz

| F0               |          |              |            |         |
|------------------|----------|--------------|------------|---------|
| Modalidad/ Canal | anecoico | reverberante | telefónico | celular |
| lectura          | 206      | 202          | 220        | 215     |
| evidencia        | 192      | 183          | 218        | 202     |
| asombro          | 189      | 183          | 210        | 208     |
| atenuación       | 203      | 193          | 210        | 218     |
| no conclusión    | 193      | 193          | 197        | 196     |
| Max              | 206      | 202          | 220        | 218     |
| Min              | 189      | 183          | 197        | 196     |
| Promedio         | 197      | 191          | 211        | 208     |
| Desv St          | 7        | 8            | 9          | 9       |

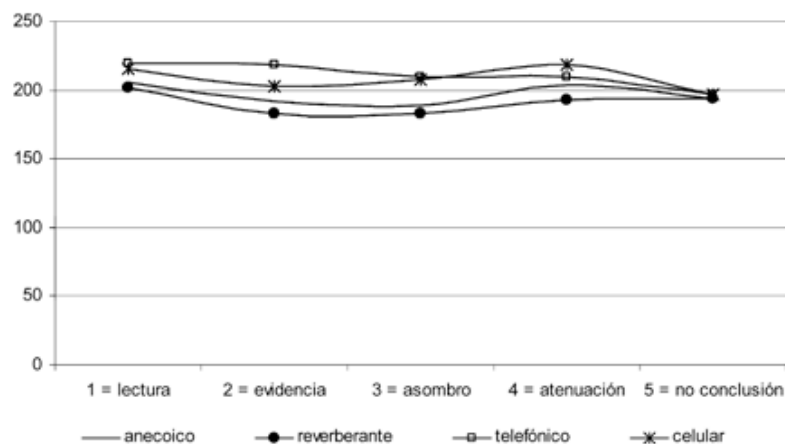


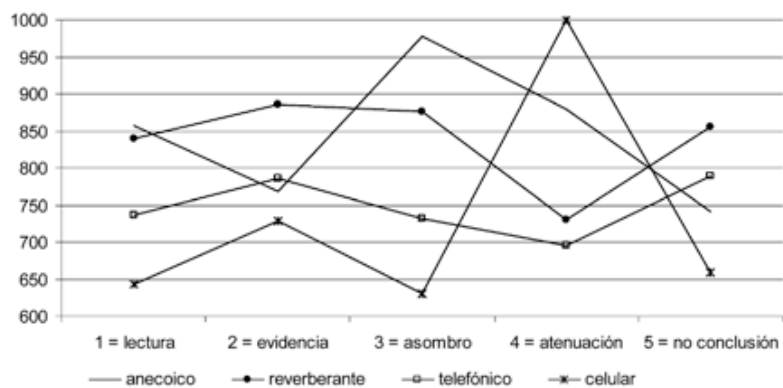
Tabla 2  
Diferencias Estadísticamente Significativas entre Modalidad y Canal<sup>1</sup>

F0  
 Evidencia - Lectura: no significativa (n.s.,  $p = 0,197$ )  
 Asombro - Lectura: \* significativa 5% ( $p = 0,026$ )  
 Atenuación - Lectura : no significativa (n.s.,  $p = 0,708$ )  
 No Conclusión - Lectura: no significativa (n.s.,  $p = 0,067$ )  
 Anecoico-Reverberante: no significativa (n.s.,  $p = 0,270$ )  
 Anecoico-Telefónico \* significativa 5% ( $p = 0,025$ )  
 Anecoico-Celular : no significativa (n.s.,  $p = 0,065$ )

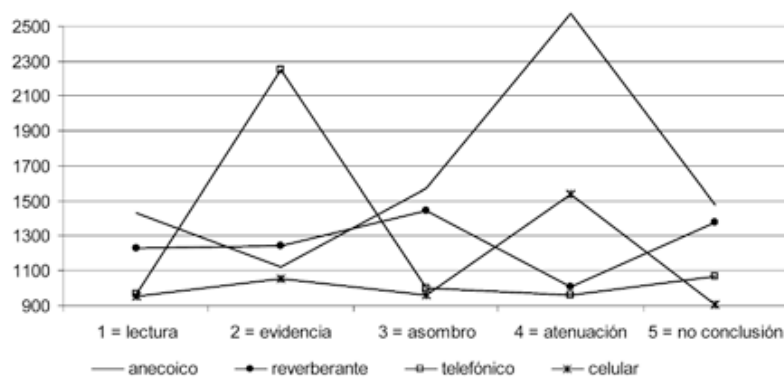
<sup>1</sup> El tratamiento estadístico de los datos fue realizado en el Laboratorio de Estadística de la Universidad Austral de Chile por el Profesor Luis Ojeda. El software utilizado fue el SPSS para windows en su versión 11.5. Los asteriscos expresan el grado de significación: \* significativa y \*\* altamente significativa.

Tabla 3  
Formantes por Modalidad y Canal. Valores expresados en Hz

| F1              |           |              |            |           |
|-----------------|-----------|--------------|------------|-----------|
| Modalidad/Canal | anecoico  | reverberante | telefónico | celular   |
| Lectura         | 857,5225  | 839,09375    | 736,31625  | 643,3     |
| evidencia       | 767,48375 | 884,93625    | 785,61875  | 728,39125 |
| asombro         | 978,48875 | 876,7375     | 732,02125  | 629,99625 |
| atenuación      | 879,415   | 730,42       | 695,13     | 999,76    |
| no conclusión   | 741,35125 | 854,8325     | 788,3275   | 658,13625 |
| Max             | 978       | 885          | 788        | 1000      |
| Min             | 741       | 730          | 695        | 630       |
| Promedio        | 845       | 837          | 747        | 732       |
| Desv St         | 95        | 62           | 39         | 154       |

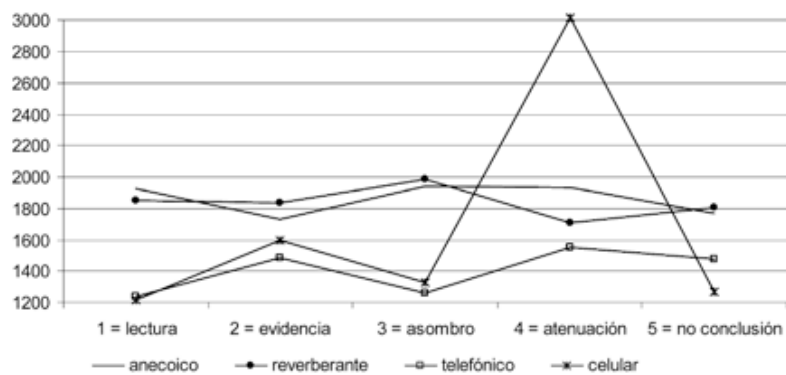


| F2               |            |              |            |            |
|------------------|------------|--------------|------------|------------|
| Modalidad/ Canal | anecoico   | reverberante | telefónico | celular    |
| Lectura          | 1433,5875  | 1232,19125   | 967,87     | 954,6325   |
| evidencia        | 1121,20875 | 1244,37625   | 2252,58125 | 1051,27125 |
| asombro          | 1570,64625 | 1443,06125   | 1002,7975  | 963,68125  |
| atenuación       | 2575,55625 | 1006,58      | 959,06     | 1538,83    |
| no conclusión    | 1479,87125 | 1376,33625   | 1066,99    | 904,89125  |
| Max              | 2576       | 1443         | 2253       | 1539       |
| Min              | 1121       | 1007         | 959        | 905        |
| Promedio         | 1636       | 1261         | 1250       | 1083       |
| Desv St          | 552        | 168          | 562        | 260        |



IMPLICANCIAS Y PROYECCIONES FORENSES EN EL ANÁLISIS DE LA VOZ

| F3               |            |              |            |            |
|------------------|------------|--------------|------------|------------|
| Modalidad/ Canal | anecoico   | reverberante | telefónico | celular    |
| Lectura          | 1926,5325  | 1848,54125   | 1239,88625 | 1212,39125 |
| evidencia        | 1731,475   | 1832,84125   | 1480,8175  | 1596,58875 |
| asombro          | 1937,75875 | 1986,8225    | 1260,71125 | 1330,71375 |
| atenuación       | 1932,52875 | 1710,61      | 1552,6     | 3016,8     |
| no conclusión    | 1770,795   | 1802,1875    | 1474,36875 | 1270,37375 |
| Max              | 1938       | 1987         | 1553       | 3017       |
| Min              | 1731       | 1711         | 1240       | 1212       |
| Promedio         | 1860       | 1836         | 1402       | 1685       |
| Desv St          | 100        | 100          | 142        | 759        |



| F4               |            |              |            |            |
|------------------|------------|--------------|------------|------------|
| Modalidad/ Canal | anecoico   | reverberante | telefónico | celular    |
| Lectura          | 2570,25375 | 2933,4275    | 1851,54    | 1849,14125 |
| evidencia        | 2268,02625 | 2804,815     | 1885,475   | 2455,30375 |
| asombro          | 2771,89    | 2899,32875   | 1863,85625 | 2022,075   |
| atenuación       | 2575,55625 | 2326,35      | 1925,31    | 5916,64    |
| no conclusión    | 2680,9025  | 2813,46875   | 1956,21625 | 1910,665   |
| Max              | 2772       | 2933         | 1956       | 5917       |
| Min              | 2268       | 2326         | 1852       | 1849       |
| Promedio         | 2573       | 2755         | 1896       | 2831       |
| Desv St          | 190        | 246          | 44         | 1741       |

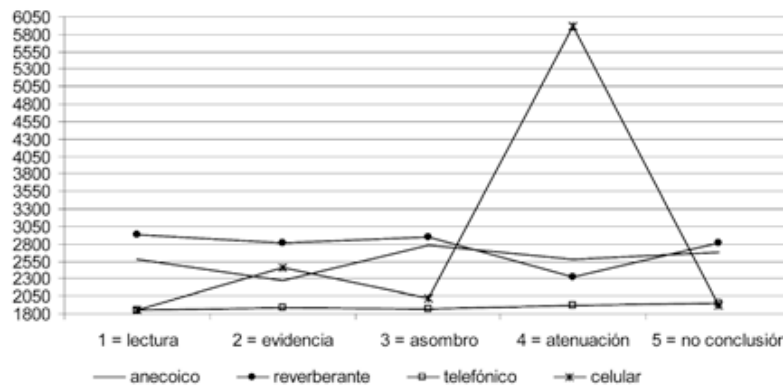


Tabla 4  
Diferencias Estadísticamente Significativas entre Modalidad y Canal

|  |
|--|
| <b>F1</b>  |
| Lectura  |
| Anecoico-Reverberante: no significativo (n.s., $p = 0,730$ ) |
| Anecoico-Telefónico: ** significativa 1% ( $p = 0,004$ )     |
| Anecoico-Celular: ** significativa 1% ( $p = 0,000$ )        |
| Evidencia  |
| Anecoico-Reverberante: no significativo (n.s., $p = 0,079$ ) |
| Anecoico-Telefónico: no significativo (n.s., $p = 0,770$ )   |
| Anecoico-Celular: no significativo (n.s., $p = 0,615$ )      |
| Asombro  |
| Anecoico-Reverberante: no significativo (n.s., $p = 0,160$ ) |
| Anecoico-Telefónico: ** significativa 1% ( $p = 0,000$ )     |
| Anecoico-Celular: ** significativa 1% ( $p = 0,000$ )        |
| Atenuación   |
| Anecoico-Reverberante: no significativo (n.s., $p = 0,623$ ) |
| Anecoico-telefónico: no significativo (n.s., $p = 0,052$ )   |
| Anecoico-celular: ** significativa 1% ( $p = 0,001$ )        |
| No Conclusión  |
| Anecoico-reverberante: no significativo (n.s., $p = 0,367$ ) |
| Anecoico-telefónico: ** significativa 1% ( $p = 0,000$ )     |
| Anecoico-celular: ** significativa 1% ( $p = 0,000$ )        |
| <b>F2</b>  |
| Lectura  |
| Anecoico-reverberante: no significativo (n.s., $p = 0,100$ ) |
| Anecoico-telefónico: ** significativa 1% ( $p = 0,001$ )     |
| Anecoico-celular: ** significativa 1% ( $p = 0,001$ )        |
| Evidencia  |
| Anecoico-reverberante: no significativo (n.s., $p = 0,178$ ) |
| Anecoico-telefónico: no significativo (n.s., $p = 0,367$ )   |
| Anecoico-celular: no significativo (n.s., $p = 0,542$ )      |
| Asombro  |
| Anecoico-reverberante: no significativo (n.s., $p = 0,428$ ) |
| Anecoico-telefónico: ** significativa 1% ( $p = 0,002$ )     |
| Anecoico-celular: ** significativa 1% ( $p = 0,001$ )        |
| Atenuación   |
| Anecoico-reverberante: no significativo (n.s., $p = 0,743$ ) |
| Anecoico-telefónico: ** significativa 1% ( $p = 0,009$ )     |
| Anecoico-celular: ** significativa 1% ( $p = 0,000$ )        |
| No Conclusión  |
| Anecoico-reverberante: no significativo (n.s., $p = 0,197$ ) |
| Anecoico-telefónico: ** significativa 1% ( $p = 0,000$ )     |
| Anecoico-celular: ** significativa 1% ( $p = 0,000$ )        |



### F3

#### Lectura

Anecoico-reverberante: no significativo (n.s.,  $p = 0,535$ )  
Anecoico-telefónico: \*\* significativa 1% ( $p = 0,000$ )  
Anecoico-celular: \*\* significativa 1% ( $p = 0,000$ )

#### Evidencia

Anecoico-reverberante: no significativo (n.s.,  $p = 0,354$ )  
Anecoico-telefónico: \* significativa 5% ( $p = 0,023$ )  
Anecoico-celular: no significativo (n.s.,  $p = 0,613$ )

#### Asombro

Anecoico-reverberante: no significativo (n.s.,  $p = 0,780$ )  
Anecoico-telefónico: \*\* significativa 1% ( $p = 0,000$ )  
Anecoico-celular: \*\* significativa 1% ( $p = 0,000$ )

#### Atenuación

Anecoico-reverberante: \* significativa 5% ( $p = 0,039$ )  
Anecoico-telefónico: \*\* significativa 1% ( $p = 0,001$ )  
Anecoico-celular: \*\* significativa 1% ( $p = 0,000$ )

#### No Conclusión

Anecoico-reverberante: no significativo (n.s.,  $p = 0,796$ )  
Anecoico-telefónico: \*\* significativa 1% ( $p = 0,000$ )  
Anecoico-celular: \*\* significativa 1% ( $p = 0,000$ )

### F4

#### Lectura

Anecoico-reverberante: no significativo (n.s.,  $p = 0,071$ )  
Anecoico-telefónico: \*\* significativa 1% ( $p = 0,002$ )  
Anecoico-celular: \*\* significativa 1% ( $p = 0,002$ )

#### Evidencia

Anecoico-reverberante: \*\* significativa 1% ( $p = 0,009$ )  
Anecoico-telefónico: \* significativa 5% ( $p = 0,041$ )  
Anecoico-celular: no significativo (n.s.,  $p = 0,739$ )

#### Asombro

Anecoico-reverberante: no significativo (n.s.,  $p = 0,322$ )  
Anecoico-telefónico: \*\* significativa 1% ( $p = 0,000$ )  
Anecoico-celular: \*\* significativa 1% ( $p = 0,001$ )

#### Atenuación

Anecoico-reverberante: no significativo (n.s.,  $p = 0,183$ )  
Anecoico-telefónico: \*\* significativa 1% ( $p = 0,004$ )  
Anecoico-celular: \* significativa 5% ( $p = 0,010$ )

#### No Conclusión

Anecoico-reverberante: no significativo (n.s.,  $p = 0,969$ )  
Anecoico-telefónico: \*\* significativa 1% ( $p = 0,000$ )  
Anecoico-celular: \*\* significativa 1% ( $p = 0,000$ )

Tabla 5  
Muestra de Resultados Información Auxiliar

| Locutora/<br>Parámetros<br>Auxiliares | Velocidad<br>(Palabras/<br>Min.) | Precisión Articulatoria   | Fluidez Articulatoria                                   | Frecuencia<br>Fundamental<br>a Largo Plazo<br>(LTF0) |
|---------------------------------------|----------------------------------|---|---|--|
| L1                                    | 184                              | Adecuada (escasa presencia de debilitamientos y aspiraciones)   | Adecuada (el discurso progresa sin dificultades)        | 227,62   |
| L2                                    | 190                              | Regular (escasa presencia de aspiraciones, repeticiones, debilitamientos, elisiones, dubitaciones, pausas llenas) | Adecuada (el discurso progresa sin dificultades)        | 226,26   |
| L3                                    | 193                              | Regular (aspiraciones, repeticiones, debilitamientos, elisiones, dubitaciones, pausas llenas)                     | Regular (el discurso progresa con algunas dificultades) | 230,36   |
| L4                                    | 174                              | Regular (aspiraciones, repeticiones, debilitamientos, elisiones, dubitaciones, pausas llenas)                     | Regular (el discurso progresa con algunas dificultades) | 222,06   |

| Locutora/Parámetros Auxiliares | F0 temporal (Hz) | F0 a Largo Plazo (Hz) |
|--------------------------------|------------------|-----------------------|
| L1                             | 197              | 227,62                |
| L2                             | 192              | 226,26                |
| L3                             | 213              | 230,36                |
| L4                             | 194              | 222,06                |

Transliteraciones Textos Auxiliares (L1)

Texto

me llamaron por la mañana  
por teléfono  
Avisándome  
que había un espectáculo  
no me lo podía perder  
tú no te lo puedes perder me dijeron  
dije bueno  
si eh...  
así entonces me voy a preparar  
bueno una tenida  
más o menos ad hoc  
que no fuera ni muy calurosa ni muy fría  
tomé finalmente  
muchas cosas  
nada me quedó bueno  
opté por un jeans  
que siempre se lleva bien  
me puse el jeans  
me fui a Santiago

llegué a lah seih de la tarde el espectáculo era a lah ocho  
tenía tiempo para hacer algunoh tipoh de...  
Travesurah  
hice algunah travesurah y me fui al ehpectáculo  
máh preparada todavía  
llegué al ehpectáculo  
por supuehto madreporadanza  
se apagaron lah luceh  
todo en silencio...  
y aparecen dos bailarines  
un hombre y una mujer  
Moviéndose al compás  
como olas se movían como olas los dos  
muy juntos  
cada paso llevaba al otro compañero era muy bonito  
lah luces eran intimas  
muy minimalihta  
madreporadanza se llamaba  
madreporadanza con música flamenca  
yo solo quiero [tarearea] yo la había escuchado muchas veces muchas veces  
me dije alguna vez voy a hacer algo con esa música como me sucede con mucha otra música que yo escucho y digo ah...  
con esto voy a hacer algo pueden pasar muchoh añoh y la idea va madurando en mi cabeza  
hahta que de repente aparece el momento adecuado...  
para hacerlo  
era un hombre moreno  
de pelo em pecho  
muy ehpañol ah  
y ella...  
muy flaca  
mucho hueso  
mucho múculo mucha fibra  
se movían precioso  
él la llevaba por supuehto  
ella era un objeto  
era un objeto no  
pero él...  
tenía la pasión del movimiento mucho stacatto  
el movimiento eh firme y bien concreto  
ella mah bien se dehplazaba por el escenario  
como un atuendo  
muy simple loh doh de blanco solo de pantalón y ella con un vehtido largo  
dehcalzoh loh doh  
sin ni una gota de agotamiento  
era precioso verloh bailar  
madreporadanza decía yo  
lo voy a anotar para que alguna vez  
Yo  
en algún momento lo buhco en el computador  
por supuehto que yo ni siquiera tengo correo pero  
dije éhto sí vale la pena  
éhto sí vale la pena verlo por computador  
me voy a sentar y me voy a conectar me van a salir lah antenna y todo pero me conecto...  
madreporadanza  
terminó el ehpectáculo ...  
y salí contenta  
[no] como cuando uno ve un ehpectáculo y termina cabreao  
no saleh ni enojado ni...  
ni...  
ni...  
deprimido porque veh algo que...