

Efectos acústicos de las variaciones fonopragmáticas y ambientales*

Acoustic effects of phonopragmatic and environmental variations

Claudia Rosas¹, Jorge Sommerhoff²

¹Universidad Austral de Chile, Facultad de Filosofía y Humanidades,
Instituto de Lingüística y Literatura, Valdivia, Chile,
e-mail: claudiarosas@uach.cl

²Universidad Austral de Chile, Facultad de Ciencias de la Ingeniería,
Instituto de Acústica, Valdivia, Chile,
e-mail: jsommerh@uach.cl

En este trabajo se analiza el comportamiento del F0 y de los formantes en situaciones fonopragmáticas y ambientales diversas. El corpus analizado está integrado por 160 segmentos vocálicos correspondientes a una vocal “a” tónica inserta en ocho contextos fonéticamente balanceados emitidos por una actriz en canal anecoico en situación de lectura, evidencia, asombro, atenuación y no-conclusión, y reproducidos luego todos ellos en canal telefónico, celular y reverberante. Los resultados muestran que las situaciones fonopragmáticas y ambientales causan variaciones en las propiedades acústicas de la señal. Los efectos producidos por estos factores no son homogéneos. El F0 y los formantes son afectados de diferente manera y grado.

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

This work analyzed the behavior of F0 and formants under diverse phonopragmatic and environmental situations. The analyzed corpus was integrated by 160 vocalic segments corresponding to a tonic vowel “a” inserted in 8 phonetically balanced contexts and emitted by an actress in an anechoic channel during reading, evidence, amazement, attenuation and non-conclusion situations, and then all of them reproduced by a phone, mobile phone and reverberant channels. The results showed that the phonopragmatic and environmental situations cause variations in the acoustic properties of the signal. The effects produced by these factors were not homogeneous. The F0 and formants were affected in different ways and degrees.

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

* Los datos de este trabajo proceden del proyecto de investigación Fondecyt N° 1070210 (Conicyt).

1. INTRODUCCIÓN

En Chile, la Reforma Procesal Penal y su correspondiente implementación atrajo la atención de las voces grabadas como posibles pruebas en la determinación de inocencia o culpabilidad de los sujetos llevados a juicio oral y, con ello, el desarrollo de los análisis periciales. Sin embargo, dicho desarrollo no ha contado con el concurso necesario de criterios que permitan validar el tipo de pruebas utilizadas y, en consecuencia, asegurar la fiabilidad de las conclusiones arribadas en cada análisis. Lo anterior, no obstante la indiscutida relevancia del software utilizado (Homayounpour y Chollet 1995), trasciende el aspecto puramente instrumental del análisis y se relaciona directamente con la valoración que se debe otorgar a los hechos comparados, es decir, los indicadores y rangos que se deben conocer y manejar por los expertos para interpretar los resultados (Rosas y Sommerhoff 2005). El conocimiento de estos indicadores y criterios permanece todavía sin determinar, debido a causas relacionadas con la complejidad del fenómeno considerado, como asimismo con el tardío interés por las diferencias fonéticas idiolectales (Battaner *et al.* 2005).

En efecto, motivado quizá por el hecho más fácilmente perceptible de las diferencias fonéticas entre hablantes, la cantidad de estudios al respecto es comparativamente más numerosa que la referida a la variación individual. Entre las investigaciones del primer tipo se encuentran estudios que, con un enfoque sociolingüístico, intentan descubrir una relación entre variables físicas (sexo, edad, constitución física) y características acústicas (Van Dommelen 1995), incluso con hábitos sociales como los de un fumador (Braun y Rietveld 1995), además de otras psicosomáticas (Van Rie y Bezooijen 1995). El balance de todos ellos permite señalar que existe una relación entre las variables sociales, fisiológicas y psicológicas consideradas y las características fonocústicas de los sujetos comparados; por ejemplo, se establece que la voz de un fumador se percibe más grave que la de otro que no lo es; también se percibe más grave la voz perteneciente a un individuo de contextura gruesa frente a otro delgado, y que la diferencia sexual, con notorias implicancias acústicas entre adultos, entre menores se neutraliza.

Por el contrario, la variación interna del hablante, de escaso tratamiento –condición determinada, en gran medida, además de la complejidad de la propia naturaleza que su estudio implica por la creencia no probada de que la variabilidad intrahablante es siempre menor que la existente entre hablantes– ha permitido constatar en contra de esta idea que el grado con que se presenta la variación en un mismo sujeto puede llegar a ser incluso mayor que la existente entre extraños, como evidencian algunos estudios al respecto. Por ejemplo, se ha comprobado, en este sentido, que en condiciones expresivas con distintos grados de seguridad los espectrogramas que puede ofrecer la emisión fónica de una misma persona son manifiestamente diferentes (Baldwin y French 1990); e inversamente, como se recoge en otro estudio, en contextos similares, los resultados espectrográficos provenientes de fuentes distintas arrojan resultados semejantes (Braun 1995).

Hechos como los descritos ponen en evidencia la necesidad de considerar el estudio de la variación interna del hablante; sin embargo, a las dificultades propias de dicho estudio se agregan en la práctica las de aplicación de los conocimientos allegados.

Ello se deriva de la diferente naturaleza de los materiales con los que se trabaja en el laboratorio y los que normalmente se obtienen en situaciones reales donde el tamaño de la muestra y el tipo de registro distan mucho de lo deseable, contextos en los cuales, por definición, las muestras son breves, el estilo, las condiciones ambientales y el canal, generalmente telefónico o celular, junto con el ruido de fondo que suele estar presente en las grabaciones, impiden la comparación directa con la voz del sujeto conocido cuyas circunstancias de emisión son diferentes. Estas características repercuten en el grado de seguridad de las conclusiones que los expertos deben informar a los jueces. De hecho, hay ocasiones en que definitivamente no es posible afirmar nada, al menos desde la perspectiva acústica manejada.

Lo anterior explica una posición difundida entre los especialistas que ha llevado a señalar a más alguien que no es posible reconocer la voz de una persona con un 100% de certeza (Ladefoged 2003). Algunos autores como Broeders (1995) y Nolan (1995) que comparten la mencionada posición, aunque desde una perspectiva menos radical, han visto una salida al problema mediante el tratamiento de las conjeturas en términos probabilísticos. Para ello han propuesto categorías que distribuyen en una escala que va desde “altamente probable”, “probable”, “posible” hasta “improbable”.

Sin duda, una escala como la mencionada resulta claramente más adecuada para rotular la certeza de las afirmaciones. Con todo, el escollo de la determinación de los criterios intrínsecos adscritos a la naturaleza del fenómeno acústico continúa existiendo. Llegados a este punto, convendría señalar que no se escapa el hecho de que la naturaleza comunicativa del fenómeno oral considerado comporta otras propiedades derivadas de su carácter comunicativo que pueden y deben ser aprovechadas, aunque la investigación pueda parcelar los objetivos y niveles de análisis propuestos. En esta misma línea de análisis paralelos o complementarios se encuentran algunos estudios de carácter perceptivo (Pollack *et al.* 1954; Compton 1963; Stevens 1971; Hollien 1995; Kuwabara y Takagi, 1991; Kreiman y Papçun 1991; Pisoni 1993), como también otros de corte perceptivo y acústico (De Figuereido y Olivier 1995).

Dentro de la amplia gama de factores que pueden contribuir a la intravariabilidad del hablante (expresivos, psicosomáticos, cognitivos, conscientes, inconscientes e involuntarios, etc.), y a las perspectivas de análisis (comunicativa, perceptiva y acústica, básicamente) en el ámbito forense se ha privilegiado el estudio acústico de la señal, porque los resultados se ajustan mejor al perfil de prueba deseada por la representación objetiva que pueden ofrecer los gráficos. Lo anterior introduce el tema de los indicadores utilizados en dichas mediciones: cuáles y qué grado de responsabilidad en la variación del sujeto se atribuye a cada uno de ellos. Este problema todavía se discute; sin embargo, se ha logrado reconocer la incidencia de los siguientes indicadores: frecuencia del fundamental (F0), pausas, velocidad de la elocución, intensidad, formantes vocálicos, sonidos sordos, sonidos nasales, fluidez y precisión articulatoria (Stevens 1971; Battaner *et al.* 2004; Van Rie y Bezooijen 1995). Siguiendo con el mismo enfoque de análisis, hay un aspecto que ha sido menos investigado y que tiene que ver con cuánto y cómo pueden variar las características acústicas de la voz en diferentes condiciones, especialmente ligadas a los contextos forenses de interés constatados. Se ha establecido que el tipo de transmisión telefónica filtra un rango de frecuencias (entre 300 y 3.400 Hz) que repercute en la información acústica de

la señal que escapa a esa selección (Cf. Künzel 1995; Künzel 2001; Rose 2003 y Nolan 1983). En este sentido, se han realizado estudios referidos a los efectos de la transmisión telefónica y celular sobre las propiedades acústicas de las vocales (Künzel 2001; Byrne y Foulkes 2004). También se han comprobado cambios en la señal que introduce el propio hablante al hablar por teléfono (motivados por el uso del aparato en sí): pronunciación, volumen y tono (Cf. Wells (1982: 28); Summers *et al.* 1988; Summers *et al.* 1989; French (1998); Hirson, French y Howard (1995)).

Siguiendo con el punto anterior, en el ámbito forense nacional se ha comprobado que los factores de tipo expresivo y ambiental (tipo de transmisión –canal– y condiciones de ruido que rodean la emisión grabada) tienen un alto impacto en el análisis de casos (Sommerhoff y Rosas 2003a, 2003b, 2004a, 2004b, 2006).

En la actualidad se desarrolla una investigación (Rosas y Sommerhoff 2008: 25) que persigue caracterizar la variabilidad intrahablante producida por factores expresivos y ambientales en el español de Chile con el objeto de apoyar la labor de los expertos forenses en los casos donde interviene la voz, mediante la disposición de pautas que contengan una disposición jerarquizada de la variabilidad encontrada.

El objetivo del presente trabajo es comparar las señales de la voz producidas por una hablante en diferentes situaciones fonopragmáticas¹ (lectura, evidencia, asombro, atenuación y no conclusión) y que se han transmitido a través de distintos canales de comunicación (anecoico, reverberante, telefónico y celular) y captadas por la cadena electroacústica de registro sonoro.

2. MATERIALES

El corpus analizado está integrado por 160 segmentos vocálicos extraídos de segmentos más largos que coinciden con una frase de estructura simple de tipo artículo + sustantivo, emitidos por una hablante. Los segmentos corresponden a una vocal “a” tónica inserta en ocho contextos fonéticamente balanceados emitidos: precedida de /p/, /t/, /k/, y seguida de /s/ (paso, taza, casa); precedida de /s/ y seguidas de /p/, /k/ (sapo, saco); precedida de nasal /n/ y seguida de consonante nasal /n/ (nana); precedida y seguida de /p/, /t/ (papas, tata) (ver Cuadro 1). Estos segmentos fueron emitidos por una actriz en canal anecoico, quien simuló estar inmersa cada vez en cinco situaciones fonopragmáticas: lectura, evidencia, asombro, atenuación y no-conclusión, y reproducidos, luego, todos ellos en tres canales: telefónico, celular y reverberante.

Por su parte, la cadena electroacústica de registro está conformada por un micrófono Shure SM 48, la interfase analógica-digital KayPENTAX, modelo CSL 4500 (Computerized Speech Lab) y un computador con procesador Pentium con disco duro y grabador de CD. La cadena de emisión de la voz está conformada por un CD conectado a un sistema reproductor NTI modelo Talk Box.

¹ Usamos este término para referirnos, como actualmente se hace, a los efectos sonoros de los valores comunicativos (Cf. Hidalgo 2002: 29).

El proceso de registro comienza con la locutora leyendo las frases del corpus. Esta se ubica frente al micrófono, el cual está instalado en una sala anecoica (ver Figura 1.1.). Las frases son registradas por el Computerized Speech Lab y almacenadas en el disco duro del PC. Posteriormente, se agrega digitalmente, al inicio de estas frases, una señal de sincronización consistente en una señal audible de 1000 Hz y de 0,01 segundos de longitud. La señal con el pito de sincronización es finalmente grabada en un CD, información que corresponde al canal de comunicación A.

Cuadro 1
Corpus fonético

Contexto fonético	Palabra en frase portadora
Tónica /a/ precedida de /p/, /t/, /k/, y seguida de /s/	El paso La taza La casa
Tónica /a/ precedida de /s/ y seguida de /p/, /k/	El sapo El saco
Tónica /a/ precedida de nasal /n/ y seguida de consonante nasal /n/	La nana
Tónica /a/ precedida y seguida de /p/, /t/	Las papas El tata

Para generar la señal del canal de comunicación B en una sala reverberante se reproduce la señal grabada en el CD utilizando el Talk Box. Esta señal es captada por el micrófono ubicado en el campo reverberante, registrada por el Computerized Speech Lab y almacenada en el disco duro del PC (ver Figura 1.2.).

Para generar la señal del canal de comunicación C se cuenta con dos salas: 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 grabada en el CD es reproducida por el Talk Box, es captada por el micrófono del teléfono y 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 Computerized Speech Lab y disco duro (ver Figura 1.3.).

Para generar la señal del canal de comunicación D se procede en forma semejante al caso C, con la excepción de que el elemento receptor no es un teléfono normal, sino que un teléfono celular (ver Figura 1.4.).

Como herramienta de análisis de la voz registrada se utilizó el software correspondiente al Computerized Speech Lab de KayPENTAX, modelo CSL 4500.

Figura 1
Montaje de las grabaciones en los distintos ambientes:
anechoico, reverberante, telefónico y celular

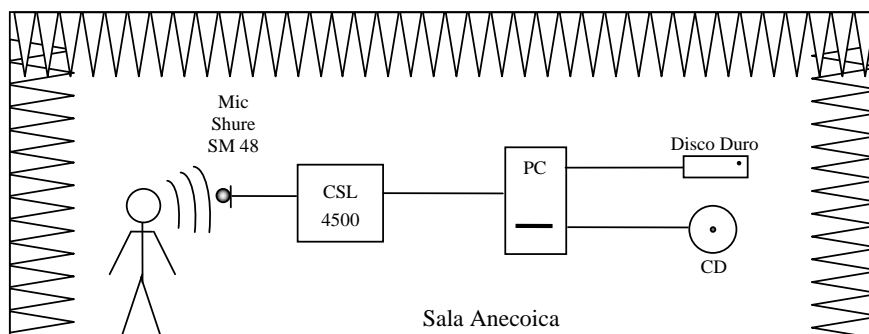


Figura 1.1. Grabación anecoica. Sonido de la voz emitido en una sala anecoica, que corresponde al caso del sonido transmitido a través de un campo sonoro directo, es decir, el caso en que el canal de transmisión no afecta al sonido emitido por el locutor.

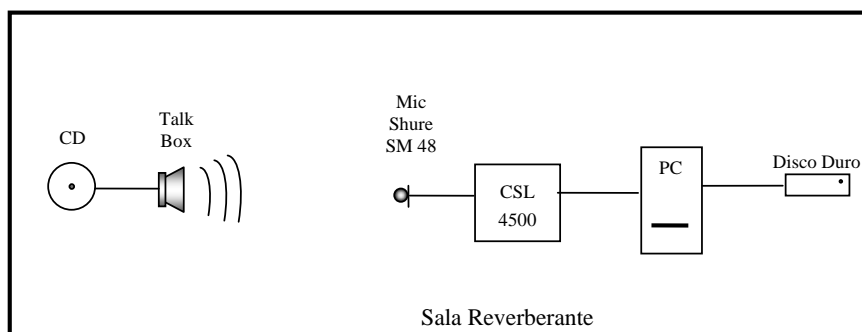


Figura 1.2. Grabación reverberante. Sonido de la voz emitido en una sala reverberante, donde el canal de transmisión le agrega sonido reverberante al sonido directo.

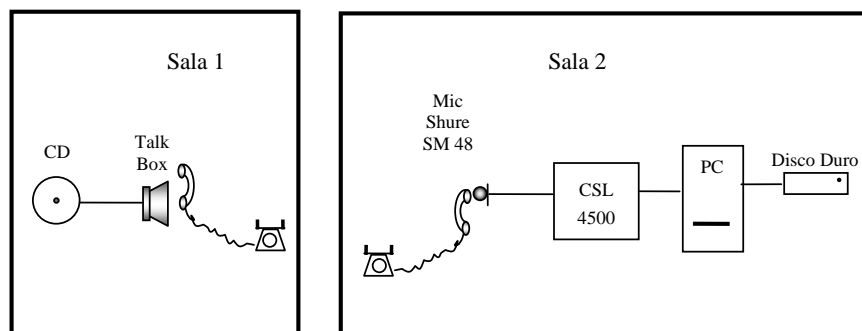


Figura 1.3. Grabación telefónica. Sonido de la voz transmitido a través de dos teléfonos, uno ubicado en la sala 1 y el otro ubicado en sala 2.

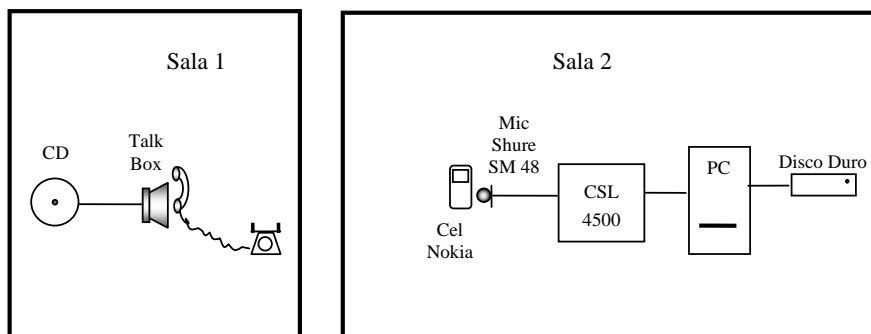


Figura 1.4. Grabación celular. Sonido de la voz transmitido a través de un teléfono y celular, donde el teléfono se instala en la sala 1 que corresponde a la del locutor y el celular en la sala 2 que corresponde a la recepción.

3. 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 $\text{media} + 3\text{desv}$.

4. RESULTADOS

Las Tablas que se presentan a continuación muestran los valores resumidos de las mediciones de F0, F1, F2, F3 y F4. Estos datos corresponden a los valores de la vocal “a” tónica en los ocho contextos fonéticos seleccionados emitidos por una actriz en cinco modalidades fonopragmáticas: lectura, evidencia, asombro, atenuación y no conclusión en canal anecoico y reproducidas todas ellas en tres canales ambientales: reverberante, telefónico y celular.

La primera columna “Todas” representa los ocho tipos de segmentos (la vocal tónica a en ocho contextos fonéticos), las dos columnas siguientes muestran los valores en Hz obtenidos en cada canal para dichos segmentos. Las tres columnas siguientes muestran el valor porcentual, la desviación estándar y la significación, aplicados a la diferencia obtenida entre los valores de los canales comparados. En los casos en que no se pudo aplicar la prueba de significación por falta de datos para conformar pares se especifica con el rótulo: No aplica. En la última columna se señala el valor límite superior de la expresión $\text{media} + 3\text{desv}$.

Tabla 1
Mediciones de F0, F1, F2, F3 y F4 por canal y modalidad fonopragmática

LECTURA						
Anecoico/ Reverberante						
	F0 ANECOICO	F0 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	247,63	323	30,43	53,29	No aplica	414,36
D. E. REVERB.		30,46				
	F1 ANECOICO	F1 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	753,66	818,26	8,57	45,68	n.s. (p=0,058)	1240,22
D. E. REVERB.		140,65				
	F2 ANECOICO	F2 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1238,16	1213,04	-2,03	17,76	n.s. (p=0,5)	1941,23
D. E. REVERB.		242,73				
	F3 ANECOICO	F3 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2616,29	1837,5	-29,77	550,68	No aplica	3705,61
D. E. REVERB.		622,7				
	F4 ANECOICO	F4 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2576,81	2222,23	-13,76	250,73	No aplica	3120,13
D. E. REVERB.		299,3				
EVIDENCIA						
	F0 ANECOICO	F0 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	274,55	263,78	-3,92	7,61	n.s. (p=0,498)	391,82
D. E. REVERB.		42,68				
	F1 ANECOICO	F1 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	780,58	705,21	-9,65	53,29	n.s. (p=0,398)	1266,00
D. E. REVERB.		186,93				
	F2 ANECOICO	F2 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1286,61	1135,88	-11,72	106,58	n.s. (p=0,500)	2203,38
D. E. REVERB.		355,83				
	F3 ANECOICO	F3 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2277,14	1798,02	-21,04	338,78	n.s. (p=0,063)	3881,01
D. E. REVERB.		694,33				
	F4 ANECOICO	F4 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	3007,47	2128,71	-29,22	621,38	n.s. (p=0,116)	4046,29
D. E. REVERB.		639,19				
ASOMBRO						
	F0 ANECOICO	F0 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	253,01	236,87	-6,38	11,42	n.s. (p=0,180)	391,29
D. E. REVERB.		51,47				
	F1 ANECOICO	F1 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	764,43	732,13	-4,23	22,84	n.s. (p=0,674)	1293,17
D. E. REVERB.		187,01				
	F2 ANECOICO	F2 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1356,59	1152,03	-15,08	144,65	n.s. (p=0,484)	1907,75
D. E. REVERB.		251,91				
	F3 ANECOICO	F3 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1997,2	1878,77	-5,93	83,74	n.s. (p=0,726)	4482,92
D. E. REVERB.		868,05				
	F4 ANECOICO	F4 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2734,72	2159,47	-21,03	406,76	n.s. (p=0,176)	5038,26
D. E. REVERB.		959,59				

EFFECTOS ACUSTICOS DE LAS VARIACIONES FONOPRAGMATICAS Y AMBIENTALES

ATENUACIÓN

	F0 ANECOICO	F0 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	242,25	231,48	-4,44	7,61	n.s. (p=0,480)	298,35
D. E. REVERB.		22,29				

	F1 ANECOICO	F1 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	694,45	656,76	-5,43	26,65	n.s. (p=0,481)	1097,61
D. E. REVERB.		146,95				

	F2 ANECOICO	F2 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1103,58	1082,04	-1,95	15,23	n.s. (p=0,778)	1903,36
D. E. REVERB.		273,77				

	F3 ANECOICO	F3 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1911,07	1359,67	-28,85	389,9	* (p=0,043)	1912,15
D. E. REVERB.		184,16				

	F4 ANECOICO	F4 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	3730,63	1870,31	-49,87	1315,44	No aplica	3037,54
D. E. REVERB.		389,07				

NO CONCLUSIÓN

	F0 ANECOICO	F0 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	269,17	226,1	-16	30,45	No aplica	290,70
D. E. REVERB.		21,54				

	F1 ANECOICO	F1 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	672,91	532,95	-20,8	98,97	n.s. (p=0,063)	781,35
D. E. REVERB.		82,8				

	F2 ANECOICO	F2 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1076,66	850,56	-21	159,88	* (p=0,017)	941,92
D. E. REVERB.		30,45				

	F3 ANECOICO	F3 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1642,68	1410,43	-14,14	164,23	n.s. (p=0,463)	2027,15
D. E. REVERB.		205,57				

	F4 ANECOICO	F4 REVERBERANTE	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2467,09	1937,99	-21,45	374,13	n.s. (p=0,091)	3351,61
D. E. REVERB.		471,21				

Anecoico/ Telefónico

LECTURA

	F0 ANECOICO	F0 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	247,63	258,4	4,35	7,61	No aplica	258,40
D. E. TELEF.		0				

	F1 ANECOICO	F1 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	753,66	947,46	25,71	137,04	* (p=0,046)	1130,17
D. E. TELEF.		60,9				

	F2 ANECOICO	F2 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1238,16	1415,81	14,35	125,62	n.s. (p=0,208)	1809,13
D. E. TELEF.		131,11				

	F3 ANECOICO	F3 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1925,68	2029,51	5,39	73,41	n. s. (p>0,05)	3540,07
D. E. TELEF.		503,52				

	F4 ANECOICO	F4 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2576,81	2686,27	4,25	77,4	n.s. (p=0,753)	4039,45
D. E. TELEF.		451,06				

EVIDENCIA

	F0 ANECOICO	F0 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	274,55	301,46	9,8	19,03	No aplica	301,46
D. E. TELEF.						

	F1 ANECOICO	F1 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	780,58	947,46	21,38	118	n.s. (p=0,058)	1405,56
D. E. TELEF.		152,7				

	F2 ANECOICO	F2 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1286,61	1431,96	11,3	102,78	n.s. (p=0,398)	2100,63
D. E. TELEF.		222,89				

	F3 ANECOICO	F3 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2277,14	1873,39	-17,73	285,49	n.s. (p=0,263)	2835,28
D. E. TELEF.		320,63				

	F4 ANECOICO	F4 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	3007,47	2341,74	-22,14	470,75	n.s. (p=0,173)	3748,49
D. E. TELEF.		468,92				

ASOMBRO

	F0 ANECOICO	F0 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	253,01	310,08	22,55	40,35	No aplica	451,61
D. E. TELEF.		47,18				

	F1 ANECOICO	F1 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	764,43	893,63	16,9	91,36	n.s. (p=0,204)	1413,88
D. E. TELEF.		173,42				

	F2 ANECOICO	F2 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1356,59	1318,91	-2,78	26,65	n.s. (p=0,779)	2167,35
D. E. TELEF.		282,82				

	F3 ANECOICO	F3 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1997,2	1706,51	-14,56	205,55	n.s. (p=0,327)	2596,11
D. E. TELEF.		296,54				

	F4 ANECOICO	F4 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2734,72	2417,1	-11,61	224,59	n.s. (p=0,362)	3693,19
D. E. TELEF.		425,36				

ATENUACIÓN

	F0 ANECOICO	F0 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	242,25	323	33,33	57,1	No aplica	489,80
D. E. TELEF.		55,6				

	F1 ANECOICO	F1 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	694,45	802,11	15,5	76,13	n.s. (p=0,414)	1574,03
D. E. TELEF.		257,31				

	F2 ANECOICO	F2 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1103,58	1291,99	17,07	133,23	n.s. (p=0,128)	1808,80
D. E. TELEF.		172,27				

	F3 ANECOICO	F3 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1911,07	1701,12	-10,99	148,46	n.s. (p=0,498)	2482,45
D. E. TELEF.		260,44				

	F4 ANECOICO	F4 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2540,92	2223,3	-12,5	224,59	n.s. (p=0,173)	3306,33
D. E. TELEF.		361,01				

EFFECTOS ACUSTICOS DE LAS VARIACIONES FONOPRAGMATICAS Y AMBIENTALES

NO CONCLUSIÓN

	F0 ANECOICO	F0 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	269,17	227,64	-15,43	29,36	* (p=0,041)	325,30
D. E. TELEF.		32,56				

	F1 ANECOICO	F1 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	672,91	936,69	39,2	186,52	* (p=0,027)	1087,21
D. E. TELEF.		50,17				

	F2 ANECOICO	F2 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1076,66	1318,91	22,5	171,3	n.s. (p=0,161)	1634,91
D. E. TELEF.		105,33				

	F3 ANECOICO	F3 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2352,5	1771,11	-24,71	411,11	n.s. (p=0,674)	2887,74
D. E. TELEF.		372,21				

	F4 ANECOICO	F4 TELEFÓNICO	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2467,09	2508,62	1,68	29,37	n.s. (p=0,779)	3501,62
D. E. TELEF.		331				

Anecoico/ Celular

LECTURA

	F0 ANECOICO	F0 CELULAR	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	247,63	323	30,43	53,29	No aplica	307,45
D.E. ANEC.	19,94					

	F1 ANECOICO	F1 CELULAR	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	753,66	823,64	9,29	49,49	n.s. (p=0,089)	1251,65
D.E. ANEC.	166					

	F2 ANECOICO	F2 CELULAR	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1238,16	1431,96	15,65	137,04	n.s. (p=0,735)	2271,18
D. E. ANEC.	344,34					

	F3 ANECOICO	F3 CELULAR	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1925,68	1716,5	-10,86	147,91	n.s. (p=0,249)	3415,15
D. E. ANEC.	496,49					

	F4 ANECOICO	F4 CELULAR	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2576,81	2251,76	-12,61	229,84	No aplica	3971,72
D. E. ANEC.	464,97					

EVIDENCIA

	F0 ANECOICO	F0 CELULAR	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	274,55	258,4	-5,88	11,42	No aplica	370,67
D. E. ANEC.	32,04					

	F1 ANECOICO	F1 CELULAR	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	780,58	861,33	10,34	57,1	n.s. (p=0,345)	1299,39
D. E. ANEC.	172,94					

	F2 ANECOICO	F2 CELULAR	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1286,61	1421,19	10,46	95,16	n.s. (p=0,310)	2424,40
D. E. ANEC.	379,26					

	F3 ANECOICO	F3 CELULAR	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2277,14	2465,55	8,27	133,23	n.s. (p=0,674)	4687,75
D. E. ANEC.	803,54					

	F4 ANECOICO	F4 CELULAR	% DIFERENCIA	D. E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	3007,47	2619,88	-12,89	274,07	No aplica	6447,59
D. E. ANEC.	1146,71					

ASOMBRO

	F0 ANECOICO	F0 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	239,94	330,17	37,61	63,8	No aplica	341,59
D. E. ANEC.	33,88					

	F1 ANECOICO	F1 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	764,43	845,18	10,56	57,1	n.s. (p=0,398)	1183,08
D. E. ANEC.	139,55					

	F2 ANECOICO	F2 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1356,59	1561,16	15,08	144,65	n.s. (p=0,400)	2853,77
D. E. ANEC.	499,06					

	F3 ANECOICO	F3 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1997,2	2304,05	15,36	216,98	n.s. (p=0,161)	3570,40
D. E. ANEC.	524,4					

	F4 ANECOICO	F4 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2734,72	2813,67	2,89	55,83	n.s. (p=0,753)	4188,26
D. E. ANEC.	484,51					

ATENUACIÓN

	F0 ANECOICO	F0 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	242,25				No aplica	338,37
D.E. ANEC.	32,04					

	F1 ANECOICO	F1 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	694,45	834,41	20,15	98,97	n.s. (p=0,062)	1093,78
D. E. ANEC.	133,11					

	F2 ANECOICO	F2 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1103,58	1787,26	61,95	483,44	* (p=0,017)	2177,76
D. E. ANEC.	358,06					

	F3 ANECOICO	F3 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1911,07	2405,57	25,88	349,66	n.s. (p=0,225)	3829,82
D. E. ANEC.	639,58					

	F4 ANECOICO	F4 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2540,92	2899,81	14,12	253,77	No aplica	4221,50
D. E. ANEC.	560,2					

NO CONCLUSION

	F0 ANECOICO	F0 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	269,17				No aplica	328,96
D. E. ANEC.	19,93					

	F1 ANECOICO	F1 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	672,91	882,86	31,2	148,46	* (p=0,028)	1103,84
D. E. ANEC.	143,64					

	F2 ANECOICO	F2 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1076,66	1658,06	54	411,11	* (p=0,017)	2117,16
D. E. ANEC.	346,83					

	F3 ANECOICO	F3 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	1642,68	2363,27	43,87	509,54	n.s. (p=0,09)	3455,20
D. E. ANEC.	604,17					

	F4 ANECOICO	F4 CELULAR	% DIFERENCIA	D.E.	SIGNIFICACIÓN	$\bar{x} + 3s$
Todas	2467,09	2922,36	18,45	321,93	n.s. (p=0,345)	5275,71
D. E. ANEC.	936,21					

* (significativo al 5%)

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

En el *canal reverberante*, en general, se observa una tendencia a la disminución en todos los parámetros y modalidades. En la modalidad de lectura el F0 y el F1 aumentan; en tanto que el F2, el F3 y el F4 disminuyen. En el resto de las modalidades todos los parámetros disminuyen sistemáticamente y en la modalidad de atenuación la disminución del F3 es significativa.

En el *canal telefónico* no existe una tendencia general que se aplique a todos los parámetros de todas las modalidades. Sólo la modalidad de lectura varía de manera estable, aumentando en todos sus parámetros. También el F1 aumenta de manera regular en todas las modalidades, y de manera significativa, en las modalidades de lectura y no conclusión. En el resto de las modalidades los cambios son variables.

En el *canal celular*, si bien no existe una tendencia general del tipo de cambio que se aplique a todos los parámetros de todas las modalidades, es posible observar algunas regularidades. En la modalidad de asombro hay un aumento de todos los parámetros; y también en las modalidades de atenuación y no conclusión, exceptuando –en estas dos– el F0 que no registra datos. Por su parte, el F1 y el F2 aumentan en todas las modalidades, y de manera significativa en la modalidad de no conclusión. En las dos modalidades restantes de lectura y evidencia el comportamiento es más variable.

Por último, al comparar los tres *canales reverberante, telefónico y celular* se evidencia que el F1 es el parámetro más estable en el tipo de cambio producido, aumentando en los tres canales en la modalidad de lectura. En los canales telefónico y celular aumenta en todas las modalidades. El resto de parámetros no se afecta de la misma manera, ni en todos los canales, ni en todas las modalidades.

5. DISCUSIÓN

Algunos de nuestros resultados pueden ser comparados con resultados de estudios previos como los de Künzel (2001) y Byrne y Foulkes (2004).

Künzel compara las variaciones que introduce el canal de transmisión telefónica sobre los formantes F1 y F2. Su experimento se aplica a las vocales del inglés en mujeres y hombres. Los resultados, con diferencias entre vocales y sexo, muestran que el canal telefónico produce un aumento sistemático y significativo del F1, en tanto que el F2 varía de manera variable y no significativa. Nuestros resultados del canal telefónico en la modalidad comparable –de lectura– coinciden con los de Künzel.

El estudio de Byrne y Foulkes se refiere a los efectos que produce la transmisión celular sobre los formantes F1, F2 y F3. Su experimento reproduce el diseño de Künzel (2001) en hombres y mujeres. Sus resultados, con diferencias entre categorías de vocales y variable sexual, muestran que el canal celular aumenta todos los formantes considerados, especialmente el F1, que alcanza el mayor número de diferencias significativas. Del mismo modo, nuestros resultados del canal celular, en la modalidad de lectura, coinciden con los de los de Byrne y Foulkes, si bien, sólo a modo de indicio, ya que las diferencias encontradas en el F1 no son significativas.

6. CONCLUSIÓN

El análisis de los resultados, que representan una cuarta parte del corpus total –las emisiones de una de las cuatro actrices– del proyecto de investigación en curso, muestra que las situaciones fonopragmáticas y ambientales (canales de transmisión) causan variaciones en las propiedades acústicas de la señal. Los efectos que estos factores producen son complejos; se relacionan con distintos aspectos: parámetro estudiado sobre el que actúan (F0, F1, F2, F3 y F4), orientación del cambio (aumento versus disminución) y significancia (significativo versus no significativo). 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 comparación con estudios previos, similares al nuestro, aunque con diferencias metodológicas, sugiere algunas coincidencias. Principalmente, los efectos del teléfono coinciden con los encontrados por Künzel (2001) para el alemán, y secundariamente, aunque no de manera significativa, con los del celular de Byrne y Foulkes (2004) para el idioma inglés.

Consideramos que la variabilidad comentada en este estudio debe ser tenida en cuenta. Esto es muy importante para el análisis pericial de voces del español de Chile donde no contamos con pautas de referencia en los términos descritos ni para esta variante en particular ni para el español en general.

OBRAS CITADAS

- Baldwin, J., P. French. 1990. *Forensic Phonetics*. Londres: Printer Published Limited.
- Battaner, E. *et al.*, eds. 2004. “Vile: Estudio acústico de la variación inter e intralocutor en español”, en 6° Congreso de Lingüística General. Santiago de Compostela, 3-7 de mayo de 2004. Area de Lingüística Xeral, Universidade de Santiago de Compostela.
- Braun, A. y T. Rietveld. 1995. “The influence of Smoking Habits on Perceived Age”. *ICPhS* 95 (2): 294-297.
- Broeders, A.P.A. 1995. “The role of Automatic Speaker Recognition Techniques in Forensic Investigation”. *ICPhS* 95 (3): 154-163.
- Byrne, C. y P. Foulkes. 2004. “The mobile phone effect on vowel formants”. *International Journal of Speech, Language and the Law* (Formerly Forensic Linguistics), XI (1): 83-102.
- Compton, A. J. 1963. “Effects of filtering and vocal duration upon the identification of speakers aurally”. *Journal of the Acoustical Society of America*. 35: 1748-1752.
- De Figuereido, R. M. y S. L. Olivier. 1995. “Speaker Identification Using a Spectral Moments Metric with the Voiceless Fricative /s/”. *ICPhS* 95 (3): 286-289.
- French, J. P. 1998. “Mr Akbar’s nearest ear versus the Lombard reflex: a case study for forensic phonetics”, *Forensic Linguistics*, 5 (2): 58-68.
- Hidalgo Navarro, Antonio. 2002. *Comentario fónico de textos coloquiales*. Madrid: Arco Libros, S. L., p. 29.
- Hirson, A., French, J. P. y Howard, D. 1995. “Speech fundamental frequency over the telephone and face-to-face: some implications for forensic phonetics”, in J. Windsor Lewis (ed.), *Studies in General and English Phonetics in Honour of Professor J.D. O’Connor*, London: Routledge, 230-40.

- Hollien, H., Majewski W. y E. T. Doherty. 1982. "Perceptual identification of voice under normal, stress and disguise speaking conditions". *Journal of Phonetics*. 10: 139-148.
- Homayounpour, M.M. y G. Chollet. 1995. "A Study of Intra- and Inter- Speaker Variability in Voices of Twins for Speaker Verification". *ICPhS 95* (3): 289-301.
- Kreiman, J. y G. Papçun. 1991. "Comparing discrimination and recognition of unfamiliar voices". *Speech Communication* 10: 265-275.
- Künzel, H. J. 1995. "Field procedures in forensic speaker recognition", in J. Windsor Lewis (ed.), *Studies in General and English Phonetics in Honour of professor J. D. O'Connor*, London: Routledge, 68-84.
- 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.
- Kuwabara, H. y T. Takagi. 1991. "Acoustic parameters of voice individuality and voice-quality control by analysis-synthesis method". *Speech Communication* 10: 491-495.
- Ladefoged, P. 2003. "Validity of voice identification". *Journal of the Acoustical Society of America*. Vol. 114 (4) Pt.2.
- Nolan, F. J. 1983. *The Phonetic Bases of Speaker Recognition*, Cambridge: Cambridge University press.
- . 1995. "Can the definition of each speaker be expected to come from the laboratory in the next decades?". *ICPhS 95* (3): 130-137.
- Pisoni, D. B. 1993. "Long-term spectral memory in speech perception: some new findings on talker variability, speaking rate and perceptual learning". *Speech Communication* 13: 109-125.
- Pollack, I., Pickett J.M. y W. Sumbly. 1954. "On the identification of speakers by voice". *Journal of the Acoustical Society of America* 35: 354-358.
- Rosas A., C. 2005. "La fonética en la Universidad Austral de Chile: revista general". *Documentos Lingüísticos y Literarios* 28: 86-88.
- Rosas, C. y J. Sommerhoff. 2006. "Variabilidad idiofónica en español como herramienta forense". *Estudios Filológicos* 41: 235-249.
- Rosas, C. y J. Sommerhoff. 2008. La variación acústica del hablante: aplicaciones forenses. *Actas del VI Congreso Iberoamericano de Acústica FIA 2008*, Buenos Aires, p. 25.
- Rose, P.J. 2003. "The technical comparison of forensic voice samples", in I. Freckelton and H. Selby (eds), *Expert Evidence*, Sydney: Thomson Lawbook Company, Ch., p. 99.
- Sommerhoff, J. 2003. "Aplicación de la medida de disimilitud de Itakura a la medición de la inteligibilidad de la palabra". Proyecto de Investigación DID S-2003-58. Instituto de Acústica. Universidad Austral de Chile.
- Sommerhoff J., C. Rosas. 2003. Informe pericial Defensoría Penal Pública de Caldera y Chañaral.
- Sommerhoff J., C. Rosas. 2003. Informe pericial Defensoría Penal Pública La Serena.
- Sommerhoff J., C. Rosas. 2004. Informe pericial Defensoría Penal Pública La Serena.
- Sommerhoff J., C. Rosas. 2004. Informe pericial Defensoría Penal Pública Coronel-Lota.
- Summers, W. V. et al. eds. 1988. "Effects of the Acoustical Society of America, 84: 917-28.
- Summers, W. V. et al. eds. 1989. "An addendum to "Effects of noise on speech production: acoustic and perceptual analyses"", *Journal of the Acoustical Society of America*, 86:1717-21.
- Stevens, K. 1971. "Sources of Inter- and intra-speaker variability in the acoustic properties of speech sounds". *Proceedings of the Seventh International Congress of Phonetic Sciences*. Ed. por R. Charboneau y A. Rigault. The Hague: Mouton: 206-232.

- Stevens, K. *et al.* eds. 1968. "Speaker identification and authentication: a comparison of spectrographic and auditory presentation of speech materials". *Journal of the Acoustical Society of America* 44: 1596-1607.
- Tosi, O. *et al.* eds. 1972. "Experiment on voice identification". *Journal of the Acoustical Society of America* 51: 2030-2043.
- Van Dommelen, W. 1995. "Speaker and Listener sex for speaker height and weight identification". *ICPhS* 95 (3): 738-741.
- Van Rie, J. y R. Bezooijen. 1995. "Perceptual characteristics of voice quality in dutch males and females from 9 to years". *ICPhS* 95 (3): 290-293.
- Wells, J. C. 1982. *Accents of English* (3 vols.), Cambridge: Cambridge University Press.