

EXTRACCIÓN DE LR_s DE AUDIOS DE CASOS CERRADOS DE LA POLICÍA DE INVESTIGACIONES DE CHILE PDI

Claudia Rosas¹, Jorge Sommerhoff², César Sáez³, José Novoa⁴

¹Instituto de Lingüística y Literatura, Universidad Austral de Chile (UACH), Valdivia, Chile.
claudiarosas@uach.cl

²Instituto de Acústica, Universidad Austral de Chile (UACH), Valdivia, Chile. jsommerh@uach.cl

³Laboratorio de Criminalística, Policía de Investigaciones de Chile (PDI), Chile csaeze@investigaciones.cl

⁴Instituto de Acústica, Universidad Austral de Chile (UACH), Valdivia, Chile. djnovomix@gmail.com

Resumen

Con el fin de observar el peso de la evidencia en situaciones de casos reales, se extrajo el F2 y el F3 del fonema /a/ de muestras de 10 hablantes, procedentes de audios de casos cerrados de la Policía de Investigaciones de Chile. Para el análisis de cada hablante se consideró 25 muestras por sesión. Con estos datos se estructuró una población de referencia y se procedió a calcular los LR_s para el mismo hablante y diferente hablante haciendo uso de la fórmula de Kernel multivariado de Aitken mostrado en Rose (2006). Se calcularon 10 LR_s para el mismo hablante y 100 de diferentes hablantes con lo que se obtuvo resultados con un 40% y un 29% de error para el mismo y diferente hablante respectivamente.

Palabras-clave: Comparación forense de la voz, LR.

Abstract

In order to observe the strength-of-evidence in real case situations, samples of F2 and F3, of the phoneme / a /, have been extracted from audios of 10 speakers, from closed cases of the Investigative Police of Chile. For the analysis of each speaker 25 samples per session were considered. With this data, a reference population was structured and LR_s for the same speaker and different speaker were calculated the using Kernel multivariate formula shown in Rose Aitken (2006). 10 LR_s were calculated for the same speaker and 100 LR_s from different speakers, obtaining 40% and 29% of error for the same and different speaker, respectively.

Keywords: Forensic voice comparison, LR.

PACS no. 43.72.Uv

1 Introducción

El presente trabajo forma parte de un proyecto de investigación (Fondecyt 1110742) cuyo objetivo es elaborar una base de datos para la comparación forense de la voz acorde a la casuística pericial de Chile. Parte de la metodología exige observar las características de los audios de la PDI como referente para la formación de una base de datos ad hoc.

La metodología general diseñada para asegurar el logro del objetivo propuesto, en su fase experimental, considera tres etapas, cada una asociada a un objetivo y metodología específica, a saber: (1) recuperación y caracterización de los audios ingresados a la Policía de Investigaciones de Chile (2006-2010); en la que se recupera los audios de casos reales, clasifica y extrae las características fonopragmáticas, ambientales y de transmisión de las escuchas recuperadas, aplicando un método combinado, auditivo y acústico; (2) ampliación de la base de datos de acuerdo a la casuística nacional observada; destinada a la confección de una base de datos alternativa a partir de las características fonopragmáticas, ambientales y de transmisión más frecuentes, evidenciadas en la etapa anterior; y, finalmente, (3) análisis acústico del corpus para modelamientos automáticos; donde se analizan acústicamente las características de las muestras seleccionadas en las etapas anteriores, de acuerdo a métodos y técnicas reconocidas que los entornos para procesado de la señal proveen actualmente.

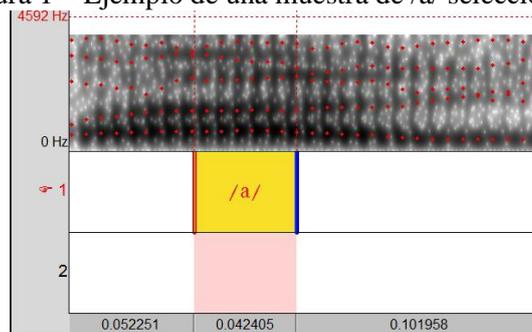
En esta oportunidad, específicamente, se requiere observar el comportamiento de los LRs aplicados a una muestra de audios registrados de la Policía de Investigaciones de Chile.

2 Análisis

2.1 Selección de los segmentos y extracción de los valores de F2 y F3.

En cada sesión se extrajo 25 muestras del fonema a las que se les calcularon mediante el uso del software PRAAT los formantes F2 y F3 generando un total de 50 datos por sesión y 100 valores por hablante. Para cada hablante se obtuvo 25 muestras. En cada sesión se extrajo 25 muestras del fonema a las que se les calcularon mediante el uso del software PRAAT los formantes F2 y F3 generando un total de 50 datos por sesión y 100 valores por hablante. Para la selección de cada muestra del fonema /a/ analizado se consideró su valor estacionario [1], además de un tiempo no inferior a 40 ms. (Ver figura 1).

Figura 1 – Ejemplo de una muestra de /a/ seleccionada



2.2 Obtención de la población de referencia y extracción de los LR

En Con esta información se estructuró una población de referencia y se procedió a calcular los LR o razones de verosimilitud para los casos del mismo y diferente hablante haciendo uso de la fórmula del kernel multivariado de Aitken mostrado en [2].

Numerador de LR=

$$(2\pi)^{-p} |D_1|^{-1/2} |D_2|^{-1/2} |C|^{-1/2} (mh^p)^{-1} |D_1^{-1} + D_2^{-1} + (h^2 C)^{-1}|^{-1/2} \\ \times \exp \left\{ -\frac{1}{2} (\bar{y}_1 - \bar{y}_2)^T (D_1 + D_2)^{-1} (\bar{y}_1 - \bar{y}_2) \right\} \\ \times \sum_{i=1}^m \exp \left[-\frac{1}{2} (y^* - \bar{x}_i)^T \left\{ (D_1^{-1} + D_2^{-1})^{-1} + (h^2 C) \right\}^{-1} (y^* - \bar{x}_i) \right]. \quad (1)$$

Denominador de LR=

$$(2\pi)^{-p} |C|^{-1} (mh^p)^{-2} \prod_{i=1}^2 \left[|D_i|^{-1/2} \right. \\ \left. + |D_i^{-1} + (h^2 C)^{-1}|^{-1/2} \times \sum_{i=1}^m \left\{ -\frac{1}{2} (\bar{y}_i - \bar{x}_i)^T (D_i + h^2 C)^{-1} (\bar{y}_i - \bar{x}_i) \right\} \right]. \quad (2)$$

Con

$$D_1 = n_1^{-1} U \quad (3)$$

$$D_2 = n_2^{-1} U \quad (4)$$

$$h = (4/(2p + 1))^{1/(p+4)} m^{-1/(p+4)} \quad (5)$$

$$y^* = (D_1^{-1} + D_2^{-1}) (D_1^{-1} \bar{y}_1 + D_2^{-1} \bar{y}_2) \quad (6)$$

Donde:

D_1 y D_2 = matrices de varianza/covarianza de las muestras indubitadas y dubitadas.

U y C = matrices de varianza y de covarianza entre hablantes

\bar{y}_1 y \bar{y}_2 = medias de las muestras indubitadas y dubitadas

n_1 y n_2 = número de replicas por hablante.

m = número de hablantes de la población de referencia,

p = número de variables correlativas por hablante.

\bar{x}_i = medias de cada sección de la población de referencia

Luego:

$$LR = \frac{\text{Numerador de LR}}{\text{Denominador de LR}} \quad (7)$$

3 Resultados

Se calculan 10 LRs para el mismo hablante y 100 de diferentes hablantes con lo que se obtuvo resultados con un 40% de error para el mismo hablante y un 29% para diferentes hablantes. (Ver Tabla 1).

Tabla 1 – Resultados de LRs (SS = mismo hablante; DS = diferente hablante).

RESULTADOS									
Índice	LR		Índice	LR		Índice	LR		Índice
	SS	DS		SS	DS		SS	DS	
1	1,7952	7,26E-09	34		6,7806	67		5,8651	
2	5,5716	5,5317	35		3,47E-08	68		2,62E-29	
3	0,0353	1,4306	36		1,15E-17	69		1,61E-12	
4	5,745	1,34E-23	37		6,3652	70		7,5088	
5	0,0043	1,92E-04	38		1,43E-07	71		0,0119	
6	6,0411	8,2411	39		5,3645	72		2,31E-17	
7	1,2645	6,3652	40		2,97E-09	73		5,60E-16	
8	3,56E-05	1,43E-07	41		0,0013	74		4,90E-11	
9	0,1161	1,76E-16	42		0,0018	75		2,02E-14	
10	4,4643	5,5317	43		3,4925	76		6,55E-15	
11		2,95E-13	44		3,47E-08	77		1,61E-12	
12		0,2527	45		3,25E-07	78		0,002	
13		6,15E-26	46		0,3236	79		8,36E-13	
14		7,91E-06	47		1,76E-16	80		0,0019	
15		6,0847	48		2,4173	81		5,0326	
16		7,1154	49		1,26E-18	82		5,4422	
17		2,97E-09	50		4,72E-10	83		0,0806	
18		1,26E-18	51		5,5032	84		0,0045	
19		1,34E-23	52		0,0067	85		6,76E-04	
20		0,0425	53		1,15E-17	86		0,0119	
21		6,15E-26	54		7,37E-15	87		2,20E-14	
22		1,50E-15	55		0,3236	88		0,0019	
23		6,47E-06	56		6,3815	89		0,0208	
24		3,34E-25	57		6,069	90		3,11E-06	
25		6,80E-21	58		8,1774	91		1,06E-05	
26		0,0018	59		2,16E-33	92		2,04E-14	
27		5,5032	60		2,02E-14	93		3,36E-18	
28		8,2411	61		8,014	94		1,51E-18	
29		3,31E-12	62		0,0045	95		5,60E-16	
30		6,0847	63		5,47E-20	96		0,2542	
31		1,1071	64		3,36E-18	97		5,4422	
32		3,34E-25	65		6,187	98		2,01E-16	
33		7,72E-05	66		8,1774	99		1,06E-05	
						100		12,7966	

4 Conclusiones

Los resultados de los LR arrojan valores inferiores a 1 del peso de la evidencia para las comparaciones del mismo hablante, con un porcentaje de error de un 40% y valores superiores a 1 del peso de la evidencia para los distintos hablantes comparados, con un 29% de error. Los valores de las comparaciones están arrojando tendencias, es decir, se percibe, al menos, discriminación entre los audios comparados –lo cual es bueno- aún con los porcentajes. Sin embargo un sistema es "robusto" mientras más grandes son los números LR "sobre 1" y "bajo 1" (peso s de la evidencia). De la experiencia realizada se hubiera esperado, para las comparaciones con el mismo hablante, valores muy superiores a 1 (miles, ojalá) pero eso es parte de las mejoras a estudiar y trabajar.

Agradecimientos

Se agradece al Consejo de Investigación Científica y Tecnológica del Gobierno de Chile que, a través de su programa Fondecyt, financia el proyecto bajo el cual se inscribe el presente trabajo.

Referencias

- [1] Hernáez, I.;Etxebarria, B.; Olabe, J.C.; Etxebarria, P.; Gaminde, I. Estudio de las trayectorias de los formantes: síntesis de las aproximantes β , δ , γ , *Phonetica, Trabajos de fonética experimental*. Bilbao, España: Universidad de Deusto, 1995, pp. 101-127.
- [2] Rose, P. Accounting for Correlation in Linguistic-Acoustic Likelihood Ratio-based Forensic Speaker Discrimination. 2006 IEEE Odyssey - *The Speaker and Language Recognition Workshop*, 2006, pp.1-8.
- [3] Brümmer, N.; Preez, J. Application-independent evaluation of speaker detection. *Computer Speech and Language* 20, 2006, pp. 230–275.