

Capoccioni, Dianora; Jiménez, Carlos; Del Pino, Paulino; Díaz, José A.
Procesamiento digital de señal fonética que permita contrastar entre las consonantes //
y /r/ en el dominio del tiempo en castellano
Ingeniería UC, Vol. 14, Núm. 2, agosto-sin mes, 2007, pp. 35-41
Universidad de Carabobo
Venezuela

Disponible en: <http://redalyc.uaemex.mx/src/inicio/ArtPdfRed.jsp?iCve=70711528006>



Ingeniería UC
ISSN (Versión impresa): 1316-6832
farteaga20@gmail.com
Universidad de Carabobo
Venezuela

¿Cómo citar?

Número completo

Más información del artículo

Página de la revista

Procesamiento digital de señal fonética que permita contrastar entre las consonantes /l/ y /r/ en el dominio del tiempo en castellano

Dianora Capoccioni, Carlos Jiménez, Paulino Del Pino, José A. Díaz
Facultad De Ingeniería. Escuela de Ingeniería Eléctrica. Universidad de Carabobo
Valencia. Venezuela.

Email: dcapocci@uc.edu.ve, cjimenez@uc.edu.ve, pdelpi@uc.edu.ve, jadiaz@uc.edu.ve

Resumen

La dislalia consiste en una alteración funcional del habla, caracterizada por omisiones, inconstancias o cambios de una consonante por otra, en este caso se considerará la pronunciación incorrecta y el intercambio de las consonantes /l/ por la /r/ y viceversa, paralambdalismo y pararrotacismo. Se hizo una investigación preliminar encuestando a personas de edades entre 17 y 56 años de ambos sexos, hallándose que el 29,68% de los encuestados presentaban dislalia tipo paralambdalismo y el 39,06% presentaban dislalia tipo pararrotacismo. Con el fin de automatizar terapias correctivas y diferenciar ambas consonantes por parte de un computador se aplicó procesamiento digital a la señal fonética para calcular los distintos parámetros, en el dominio temporal, e investigar cuáles de ellos caracterizan la /l/ y la /r/ al ser pronunciadas por la misma persona, estableciendo un patrón de pronunciación correcta. Se procedió a hacer el contraste y se observaron 13 parámetros numéricos (P1, P2, P3, ..., P13) que agrupan los valores alrededor de una media aceptablemente distante, y un conjunto de 5 parámetros de tipo gráfico (P14a, P14b, P14c, ..., P14e), todos ellos útiles en la diferenciación de ambas consonantes.

Palabras clave: Dislalia, procesamiento digital para pronunciación, contrastación de patrones, paralambdacismo, pararrotacismo.

Phonetic signal digital processing for contrasting the consonants /l/ and /r/ in time domain in Spanish language

Abstract

The dislaly is a speech impair characterized by omissions, changes of one sound for other or in inconsistencies in speech, in this case it will be taken in account the improper consonant /l/ and /r/ pronunciation as well as the interchange between them (paralambdalism and pararrotacism). It was conducted a preliminary search testing people who had ages between 17 and 56 years old, both gender. It was found 29.68% of them had dilaly type paralambdacism and 39.06% had dislaly type pararotacim. With the aim to automaticise corrective therapies and the pc abilities differentiation of both consonants, it was applied a digital signal processing to phonetic signals in order to calculate the different time domine parameters and search who of them were more suitable to characterize the \l\ and \r\ when they were pronounced by the same person, stablishing a patern of right pronunciation. It was proceeded to contrast data and was found 13 numerical parameters (P1, P2, P3, ..., P13), that were cluster around a mean acceptable value and a set of 5 graphical parameters (P14a, P14b, P14c, ..., P14e), all of them suitable to both consonant differentiation.

Keywords: Dislaly, digital processing in pronunciation, pronunciation paterns contrast, paralambdacism, pararotacism.

1. INTRODUCCIÓN

La dislalia consiste en una alteración funcional del habla caracterizada por omisiones, mala pronunciación de un sonido dentro de la palabra, cambios de un sonido por otro o inconstancias.

Los enfoques más comúnmente usados en la investigación del habla son: articulación, acústica y percepción auditiva. Los tres enfoques proveen ideas y herramientas para obtener mejores y más eficientes resultados en el reconocimiento del habla por parte de los humanos (RH) y el reconocimiento automático del

habla por parte de máquinas (RAH) [1-8].

Esta investigación toca al primer y tercer enfoque, ya que, mediante procesamiento digital de la señal fonética en el dominio temporal, se pretende establecer los parámetros que permitan a una máquina contrastar dos consonantes pronunciadas o articuladas por una persona, motivo por el cual se hace necesario establecer cuales parámetros, evaluados a partir de la pronunciación de un ser humano y almacenados en una base de datos, le permiten al computador diferenciar una /r/ de una /l/.

El trabajo consistió en hacer una investigación exploratoria de los parámetros más adecuados en el establecimiento de un patrón de referencia que permita contrastar la pronunciación de ambas consonantes, /l/ y /r/ [9].

La recolección de la data fonética se tomó de una población de 64 personas, jóvenes y adultos de habla castellana entre 17 y 56 años que frecuentan la Universidad de Carabobo, el tipo de análisis que se aplicó a los datos fue en su primera etapa cualitativa, debido a la consideración de cuáles palabras, que contienen los fonemas descritos, son difíciles de pronunciar, entendiéndose por pronunciación incorrecta el uso de la /l/ en lugar de la sílaba /r/ (paralambdalismo) o el uso de la /r/ en lugar de la sílaba /l/ (pararrotacismo) [10-11].

Posteriormente se aplicó, a los parámetros seleccionados, un análisis cuantitativo considerando aquellos parámetros que presentaban diferencias al ser medidos en la pronunciación de la /l/ y en la pronunciación de la /r/, debido a esta dualidad de análisis de datos se puede decir que se aplicó un análisis cuali-cuantitativo [9].

Una vez conformadas las matrices, cuyos elementos son los parámetros obtenidos del procesamiento digital de las sílabas /la/ y /ra/, se procedió a hacer el contraste de esos parámetros, y establecer el grado de distorsión en el uso de la consonante indebida.

2. METODOLOGÍA O DESARROLLO DE LA INVESTIGACION

La señal fonética convertida en señal eléctrica mediante muestreo a 11025Hz, canal mono y resolu-

ción de 16 bit, fue adquirida y preprocesada suprimiendo ruido y eliminando los tiempos de silencio, iniciales y finales, de la sílaba para ser almacenada en un archivo .wav, único para cada sílaba, /la/ y /ra/, pronunciada por cada individuo o unidad experimental (UE) [12]. A cada uno de los archivos se le hizo procesamiento digital, que permitió registrar y graficar la data como un conjunto de amplitudes en función del tiempo $A(t)$, o arreglos unidimensionales $[n \times 1]$ con valores de amplitud de señal que han sido normalizadas entre -1 y +1, tal como muestra el sonograma de la Figura 1 [13].

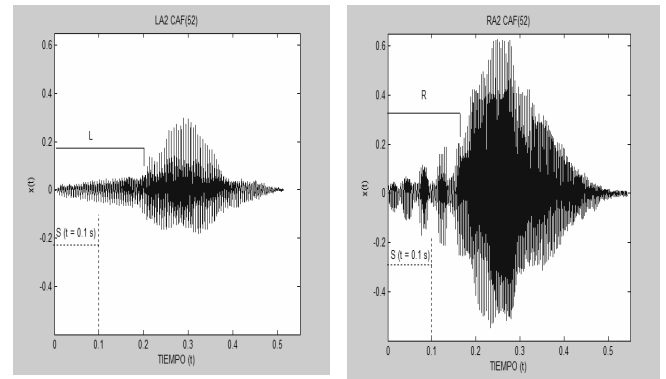


Figura 1. Sonogramas o gráficas correspondientes a las consonantes /la/ y /ra/ pronunciadas por la UE 2; C, correctamente pronunciada; por, A, adulto; F, femenino; de 52 años de edad.

Las distintas funciones o arreglos, $A(t)$, de cada una de las sílabas permitió obtener los parámetros que permitieron contrastar la pronunciación de una consonante u otra (dislalia por paralambdalismo y pararrotacismo), independientemente de: micrófono utilizado, sexo y edad, ya que entre humanos somos capaces de entender esas consonantes sin importar quién las emita siempre y cuando sean pronunciadas correctamente.

La Figura 1 muestra el sonograma de ambas sílabas pronunciada por una misma UE. Las consonantes tuvieron una longitud o duración en el tiempo promedio de duración de las sílabas fue de 0.14401 s para la consonante /la/ y 0.145300 s para /ra/, por lo que se garantiza la existencia de ambas en los primeros 0,102s de la grabación de la pronunciación de ambas consonantes, (en la redacción se dirá siempre 0.1s), este fue el intervalo de observación denominado como $S(t) = A(t = 0,1 \text{ s})$, donde: $n = 0.1 \text{ s} / (T) = 0,1 \text{ s} / (1/11025 \text{ s}^{-1}) = 0,1 / 9,07 \times 10^{-5} \text{ s} \approx 1103 \text{ muestras}$, ver Figura 1.

Se tomó el valor absoluto de cada elemento del segmento $S(t)$ y al nuevo arreglo se le denominó $absS(t)$, mostrado en la Figura 2, y con sus elementos se calcularon un conjunto de parámetros, de los cuales cuatro fueron los que proporcionaron diferencias entre la pronunciación de las consonantes /l/ y /r/.

P1: número de amplitudes de valor mayor o igual al 50 por ciento de la máxima amplitud presente en el arreglo $absS(t)$.

P2: número de amplitudes de valor menor al 50 por ciento de la máxima amplitud presente en el arreglo $absS(t)$.

P3: número de amplitudes de valor mayor o igual a 75 por ciento de la máxima amplitud presente en el arreglo $absS(t)$.

P4: número de amplitudes de valor menor al 75 por ciento de la máxima amplitud presente en el arreglo $absS(t)$.

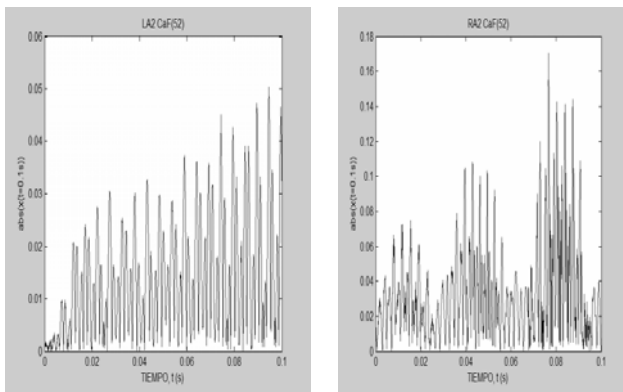


Figura 2 Gráfico correspondiente al valor absoluto de los primeros 0.1 s de las graficas correspondientes a las sílabas /la/ y /ra/ donde se garantiza la existencia de los sonidos de las consonantes /l/ y /r/ respectivamente.

El arreglo unidimensional, $absS(t)$ de la Figura 2, fue subdividido en 33 ventanas de 0.003 s de duración, tal como se muestra en la Figura 3, y en cada ventana se ha calculado el valor eficaz con las 34 muestras o amplitudes contenidas en cada ventana utilizando la ecuación (1).

$$[V_{RMS}]_n = \sqrt{\frac{\sum_{i=1}^{N_a} A_i^2}{N_a}} \quad (1)$$

donde: $n = 33$ y $N_a = 34$.

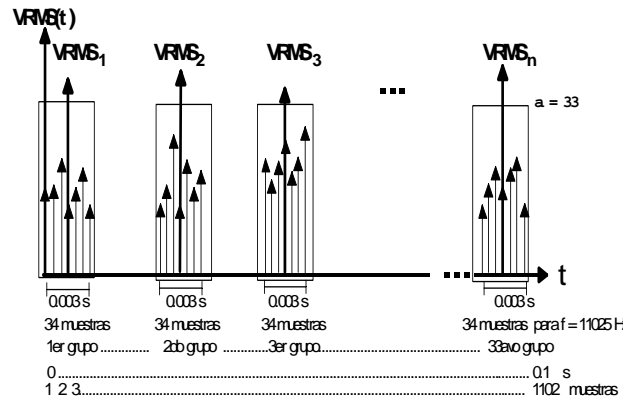


Figura 3. Esquema correspondiente al valor absoluto de los primeros 0.1 s de las gráficas correspondientes a las sílabas /la/ o /ra/ donde se garantiza la existencia de los sonidos de las consonantes /l/ y /r/ respectivamente.

Los 33 valores eficaces de cada consonante para cada UE fueron almacenados en los arreglos de 33 elementos $VRMS(t)$, Figura 4.

$$VRMS(t) = [VRMS_1 \ VRMS_2 \ VRMS_3 \ \dots \ VRMS_{33}] \quad (2)$$

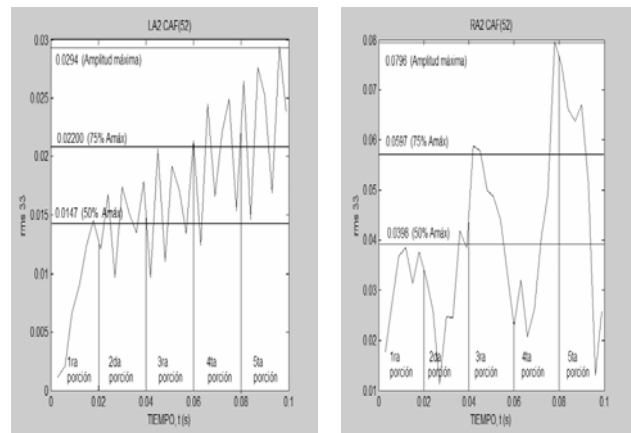


Figura 4. Gráfico del arreglo de 33 elementos, $VRMS(t)$, correspondiente al valor eficaz del valor absoluto de los primeros 0.1 s de duración de las sílabas /la/ y /ra/ reducido a 33 elementos. Se indican las marcas para la máxima amplitud presente así como la marca del 75 por ciento de la máxima amplitud y del 50 por ciento de la máxima amplitud. El área bajo la curva ha sido separada en 5 porciones.

De todos los parámetros calculados con los elementos de $VRMS(t)$, aquellos que mostraron diferencias entre la pronunciación de las consonantes /l/ y /r/

cias entre la pronunciación de las consonantes /l/ y /r/ fueron cinco parámetros.

P5: número de amplitudes mayores o iguales al 50 por ciento de la máxima amplitud presente en el arreglo VRMS(t).

P6: número de amplitudes menores al 50 por ciento de la máxima amplitud presente en VRMS(t).

P7: coeficiente de modulación (Amáx / Amín) en la 2da porción de VRMS(t).

P8: coeficiente de modulación en la 3ra porción de VRMS(t).

P9: coeficiente de modulación en la 5ta porción de VRMS(t).

Mediante cálculo de valor eficaz, ecuación (3), se simplificó el perfil de 33 elementos, VRMS(t), a un perfil de ocho elementos,

$$VCOMP(t) = [VCOMP_1 \ VCOMP_2 \ VCOMP_3 \dots VCOMP_8] \quad (3)$$

el proceso se esquematiza en la Figura 5 y el resultado en la Figura 6.

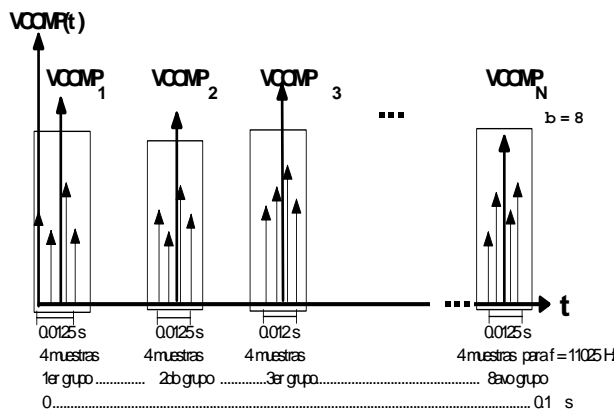


Figura 5. VCOMP(t): esquema correspondiente al valor absoluto de los primeros 0.1 s de las gráficas correspondientes a las sílabas /la/ o /ra/. Se elaboraron 8 ventanas sobre los 33 valores para calcular el valor eficaz a grupos de 4 amplitudes, (5 amplitudes para el ultimo intervalo) que componían el perfil generado por VRMS(t).

$$[V_{COMP}]_n = \sqrt{\frac{\sum_{i=1}^{N_b} V_{RMS_i}^2}{N_b}} \quad (4)$$

donde: n = 8 elementos del arreglo

N_b = número de elementos que se están promediando

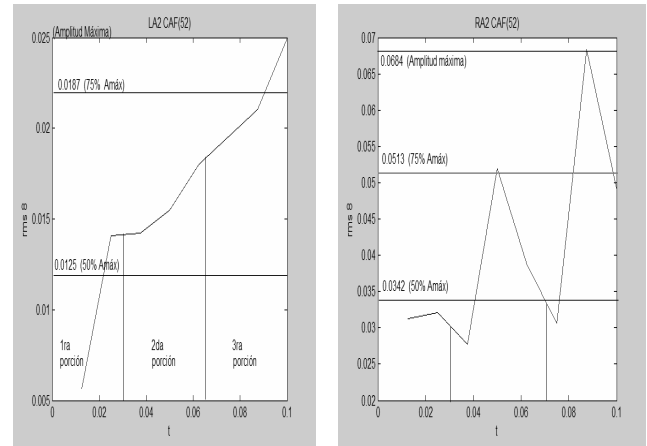


Figura 6. Gráfico del arreglo unidimensional de 8 elementos VCOMP(t). Se indican las marcas para la máxima amplitud presente así como la marca del 75% de la máxima amplitud y del 50% de la máxima amplitud. El área bajo esa curva ha sido separada en 3 porciones.

Los parámetros calculados sobre los arreglos unidimensionales VCOMP(t), y que permitieron contrastar las consonantes /l/ y /r/, fueron cuatro:

P10: número de muestras de amplitud mayor o igual al 50% del máximo valor presente en VCOMP(t).

P11: número de muestras de amplitud menor al 50% del máximo valor presente en el arreglo VCOMP(t).

P12: coeficiente de modulación (Amáx / Amín) en la 2da porción de VCOMP(t).

P13: coeficiente de modulación (Amáx / Amín) en la 3ra porción de VCOMP(t).

Finalmente los cinco parámetros gráficos P14a, P14b, P14c, P14d, P14e, los cuales se refieren al conjunto de los cinco sonogramas obtenidos para cada UE, los cuales permitieron diferenciar, por simple inspección visual, la diferencia entre la pronunciación de las consonantes /l/ y /r/. Fig. 7

P14a: perfil generado por las sílabas /la/ o /ra/.

P14b: perfil generado por las amplitudes obtenidas durante los primeros 0.1 s de pronunciación de la sílaba donde se garantiza la presencia de la consonante, S(t).

P14c: perfil del valor absoluto del segmento S(t) denominado absS(t).

Pd: perfil generado por los 33 valores eficaces o elementos del arreglo VRMS(t).

P14e: perfil construido con los 8 elementos del arreglo VCOMP(t). Ver ejemplo de la UE Nro 2 de la Figura 7.

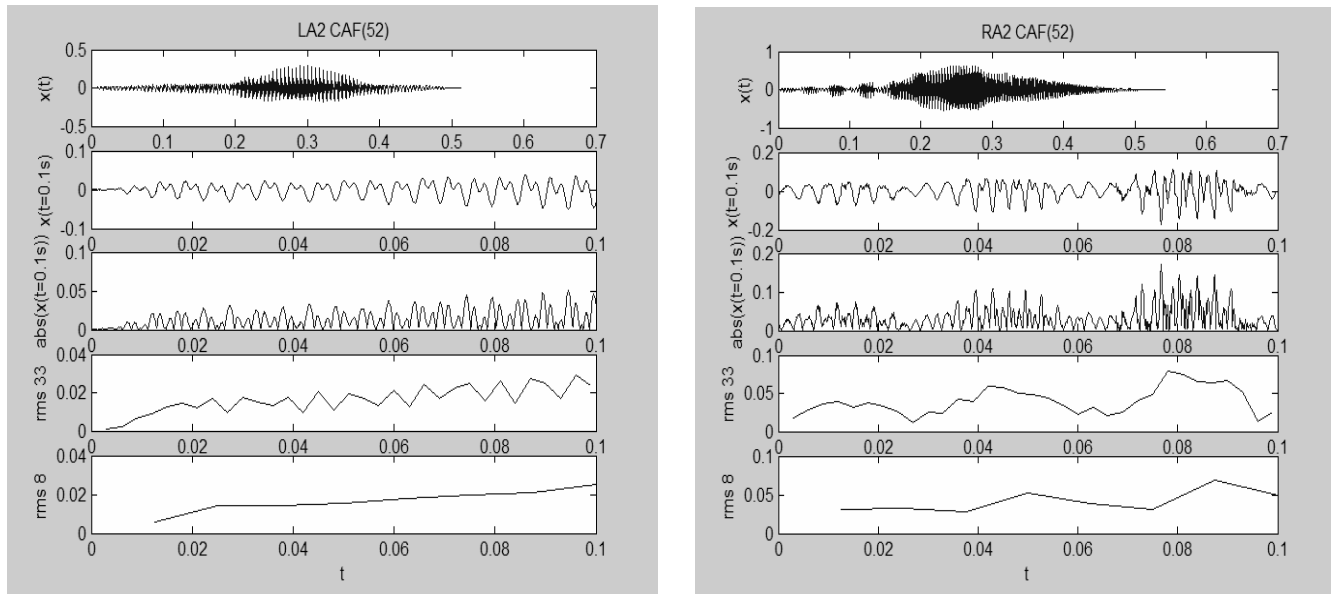


Figura 7. Conjunto de los cinco sonogramas obtenidos de la señal fonética emitida por la UE Nro 2. (a) al pronunciar ambas sílabas /la/ y /ra/.(b) al tomar el primer segmento de 0.1 s, S(t). (c) al considerar el valor absoluto del segmento, absS(t). (d) al calcular el valor eficaz de absS(t) reducido a 33 amplitudes o elementos, [VRMS(t)]. (e) al calcular valor ae eficaz de [VRMS(t)] reducido a ocho elementos. En cada grafico se evidencia, por simple inspección visual, la diferencia entre la pronunciación de las consonantes /l/ y /r/ .

3. RESULTADOS

Los parámetros evaluados para cada UE fueron los elementos dispuestos en dos columnas, una para la pronunciación de la consonante /l/, conformando el arreglo [L], y otra para la pronunciación de la consonante /r/, conformando el arreglo [R].

Se procedió a generar las columnas [L-R] y $(((L-R)/L) \times 100)$ para contrastar los elementos de cada par de columnas [L] y [R].

Los parámetros que presentaron diferencias notables entre los elementos fueron los escogidos como aquellos que permitieran diferenciar la pronunciación de ambas consonantes y al ser utilizados para construir los histogramas de frecuencias de cada par de columnas [L] y [R], presentaron agrupación de los valores alrededor de medias distanciadas.

Se les aplicó análisis de varianza de un factor, con significancia estadística $\alpha = 0.05$, [6] a ambas columnas, y se consideraron como parámetros válidos, para discernir pronunciación entre ambas consonantes, aquellos donde el factor $F \gg \gg F_{crítico}$.

Los parámetros gráficos de las consonantes /l/ y /r/ presentaron diferencias sistemáticas para todas

las UE.

A continuación, en la Tabla 1, se resaltan con negritas el conjunto de resultados que evidenciaron mayores diferencias entre ambas consonantes.

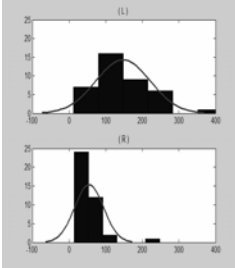
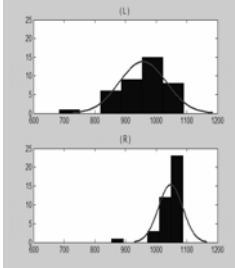
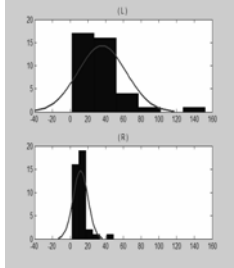
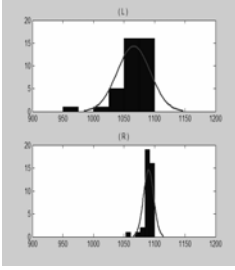
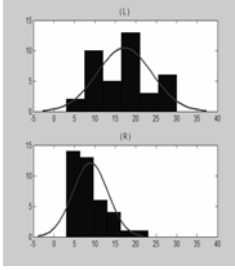
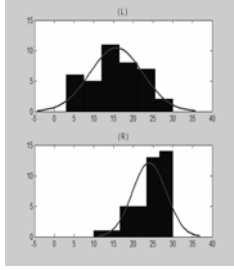
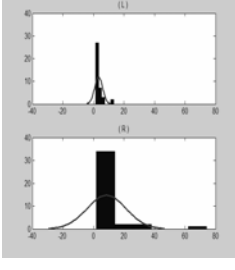
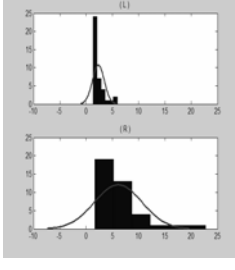
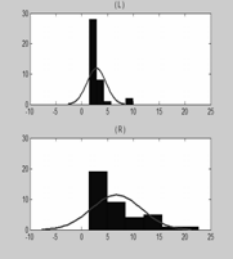
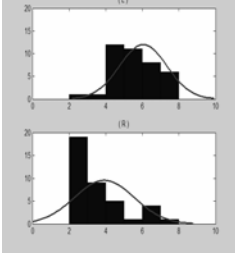
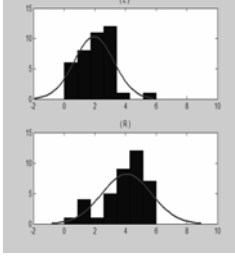
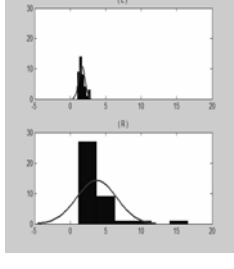
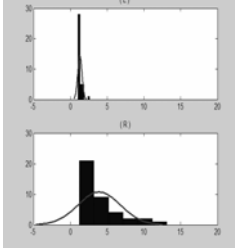
4. CONCLUSIONES

La contrastación efectuada entre el par de arreglos unidimensionales (columnas) que contenían los 69 parámetros, calculados para las consonantes /l/ y /r/, mostró 13 parámetros numéricos y cinco parámetros gráficos, para cada UE, todos ellos permiten diferenciar la consonante /l/ de la consonante /r/.

Las diferencias de las características presentadas por las gráficas o sonogramas de ambas consonantes, construidas con los valores numéricos de los arreglos unidimensionales obtenidos del preprocesamiento y procesamiento digital del sonido, son contundentes.

La consonante /l/ cuando es bien pronunciada y perfectamente audible presenta un perfil con variación constante y con aumento sostenido de amplitud a medida que transcurre el tiempo, es decir, presenta un perfil creciente con poca modulación para cada UE.

Tabla 1. Resultados que permitieron evidenciar las diferencias existentes al pronunciar ambas consonantes /l/ y /r/. Los histogramas de frecuencia corresponden a una distribución normal de los elementos de los arreglos unidimensionales correspondientes a los parámetros de cada UE, [L] y [R]. Los tres primeros resultados, pos[L-R], neg[L-R] y ceros[L-R], fueron calculados contrastando los pares de columnas o arreglos unidimensionales cuyos elementos son los parámetros de cada UE, los tres últimos se obtuvieron al aplicar análisis de varianza de un factor con significancia estadística $\alpha = 0.05$

	<p>P1 [nupaxdd_50pc_S] pos[L-R] = 35.00 neg[L-R]=4: 10,26% ceros[L-R] = 0.00 F = 47.86 P = 1.25E-09 Fcritico = 3.97</p>		<p>P2 [ndwaxdd_50pc_S] pos[L-R]=4: 10,26% neg[L-R] = 35.00 ceros[L-R] = 0.00 F = 43.28 P = 5.41E-09 Fcritico = 3.97</p>		<p>P3 [nupbxdd_75pc_S] pos[L-R] = 34.00 neg[L-R]=5: 12,82% ceros[L-R] = 0.00 F = 28.39 P = 9.79E-07 Fcritico = 3.97</p>
	<p>P4 [ndwbxdd_75pc_S] pos[L-R]=5: 12,8% neg[L-R] = 34.00 ceros[L-R] = 0.00 F = 28.39 P = 9.79E-07 Fcritico = 3.97</p>		<p>P5 [nup33_50pc] pos[L-R] = 32.00 neg[L-R]=4: 0,26% ceros[L-R]=3:7,69% F = 42.88 P = 6.15E-09 Fcritico = 3.97</p>		<p>P6 [ndw33_50pc] pos[L-R]=4: 10,26% neg[L-R] = 32.00 ceros[L-R]=3:7,69% F = 42.88 P = 6.15E-09 Fcritico = 3.97</p>
	<p>P7 [CfMc33_5P_2] pos[L-R]=8: 0,51% neg[L-R] = 31.00 ceros[L-R] = 0.00 F = 6.15 P = 0.02 Fcritico = 3.97</p>		<p>P8 [CfMc33_5P_3] pos[L-R]=3: 7,69% neg[L-R] = 36.00 ceros[L-R] = 0.00 F = 25.88 P = 2.55E-06 Fcritico = 3.97</p>		<p>P9 [CfMc33_5P_5] pos[L-R]=6: 15,4% neg[L-R] = 33.00 ceros[L-R] = 0.00 F = 21.71 P = 13.30E-06 Fcritico = 3.97</p>
	<p>P10 [nup8_50pc] pos[L-R] = 30.00 neg[L-R]= 5: 2,82% ceros[L-R]=4: 10,26% F = 40.76 P = 1.24E-08</p>		<p>P11 [ndw8_50pc] pos[L-R]=5: 12,82% neg[L-R] = 30.00 ceros[L-R]=4: 10,26% F = 40.76 P = 1.24E-08 Fcritico = 3.97</p>		<p>P12 [CfMdCOMP_2] pos[L-R]=5: 12,82% neg[L-R] = 34.00 ceros[L-R] = 0.00 F = 21.23 P = 16.19E-06 Fc = 3.97</p>
			<p>P13 [CfMdCOMP_3] pos[L-R]=1: 2,56% neg[L-R] = 38.00 ceros[L-R] = 0.00 F = 32.30 P = 2.32E-07 Fc = 3.97</p>		

El análisis de varianza de un factor, con significancia estadística $\alpha = 0.05$, permitió considerar que ambos arreglos contienen elementos distintos, tal como se observa en la Tabla 1, donde se puede ver el resultado del parámetro P1 donde $F = 47.86 \gg \gg F_c = 3.97$ y probabilidad de confusión $P = 1.25E-09$.

El estudio estadístico de los elementos de los arreglos [L] y [R] evidenció que las medias de cada conjunto de elementos, están lo suficientemente alejadas como para considerar que ambos arreglos contienen elementos que se agrupan en rangos distintos, tal como se observa en la Tabla 1.

Por otra parte, el perfil de la consonante /r/, cuando se pronuncia correctamente y se puede escuchar claramente, muestra un perfil con variaciones pronunciadas presentando valles y crestas muy marcadas a consecuencia de una mayor modulación.

Los valores de amplitud del sonido correspondientes a la consonante /l/ proporcionan un perfil o sonograma parecido para todas las UE, más aún si son pronunciadas por personas que guardan vínculo familiar cercano. Lo mismo se puede concluir para la consonante /r/.

Se ha observado que el conjunto padre e hijos varones así como el conjunto madre e hijas presentan perfiles con forma muy parecidas.

Los parámetros P1, P2, P3, P4, P8, P12, y P13, resaltados en negritas, fueron los que permitieron diferenciar con mayor confiabilidad las consonantes /l/ y /r/. Los perfiles de las cinco gráficas generadas para ambas sílabas permiten diferenciarlas con seguridad absoluta visualmente, por lo que con un ambiente confiable para la contrastación de los perfiles, que utilice redes neuronales utilizando el modelo de Kohonen, se pueden diferenciar por aproximación las formas y perfiles de cada pronunciación. [2, 14, 15].

5. REFERENCIAS

- [1] Higuera Miguel. "Espacio Logcopédico.com Trastornos metabólicos y trastornos de lenguaje. Argentina. www.espaciologopedico.com/recursos
- [2] Quero R. José Manuel. "Redes Neuronales. Aprendizaje Competitivo". Dpto. de Ingeniería Electrónica de la Universidad de Sevilla." www.gte.us.es/ASIGN/SEA/TranspRN/
- [3] Velásquez F. José Luis. "Técnicas y tipos de muestreo". <http://www.psico.uniovi.es>.
- [4] Casacubierta Francisco y Vidal Enrique. "Reconocimiento Automático del habla." Marcombo Boixareu Editores. 1987. Barcelona. España.
- [5] Flores E. Andrés. "Reconocimiento de Palabras (DTW). Cuantificación vectorial. Grupo de Procesamiento Digital de Señales. ." Pontificia Universidad Católica de Perú. www.alek.pucp.edu.pe/~dflores/tesis/
- [6] Bardaji Ana, Cruz N. Ana y Cruz N Nuria "El Rotacismo" www.imaginarium.es/vIE/
- [7] Canavos Geoge "Probabilidad y estadística. Aplicaciones y métodos." McGrawHill/ Interamericana de México, S.A. México 1988
- [8] Proakis John G. y Manolakis Dimitris G. "Tratamiento digital de señales." Prentice Hall, Inc. 2000. Madrid.
- [9] Hernández Sampieri Roberto, Fernandez Collado Carlos y Baptista Lucio Pilar. "Metodología de la Investigación." McGraw Hill. 3ra edición 2003. México D.F.
- [10] Perelló J. y Tortosa F. "Sordomudez." Ed. Científico-Médica, Audiofoniatría y Logopedia". Barcelona, España, 3ª Edición 1978. www.medynet.com/usuarios/PrevInfad/Audicion
- [11] www.pspedagogia.com. Psicología y educación para padres. <http://www.psicopedagogia.com/disartria>
- [12] Etter Dolores M. "Solución de problemas de Ingeniería con Matlab." PrinticeHall. 1997. México.
- [13] Childers Donald G. "Speech processing and synthesis toolboxes." John Wiley & Sons, Inc. 1999. New Cork, USA
- [14] Hilera José y Martínez Victor. "Redes neuronales artificiales, fundamentos, modelos y aplicaciones." Addison Wesley Iberoamericana, S.A. 1995 SERIE.
- [15] Clemente E, Vargas A, Olivier A, Kirschning I, Cervantes O. "Entrenamiento y Evaluación de reconocedores de Voz de Propósito General basados en Redes Neuronales feed forward y Modelos Ocultos de Markov." Depto. Ingeniería en Sistemas Computacionales Universidad de las Américas- -Puebla, Cholula, Puebla, México. 1998