



VI Congreso Iberoamericano de Acústica - FIA 2008  
Buenos Aires, 5, 6 y 7 de noviembre de 2008

FIA2008-A156

## **Diseño de una Metodología de Análisis de la Prosodia mediante el estudio de la Variaciones de la Frecuencia Fundamental “PITCH”, para la caracterización de voces de distintas regiones de Chile**

Guillermo Grez <sup>(a)</sup>  
Dr. Manuel Recuero <sup>(b)</sup>  
Pablo Alarcón <sup>(c)</sup>  
Daniel Opazo <sup>(c)</sup>

(a) Área de Sonido y Acústica Universidad Tecnológica de Chile INACAP, Brown Norte 290, Santiago, Chile. E-mail: [guillermogrez@vtr.net](mailto:guillermogrez@vtr.net)

(b) Investigador responsable, Grupo de Investigación en Instrumentación y Acústica Aplicada (I2A2), Universidad Politécnica de Madrid, Crta. Valencia Km-7 Madrid 28031, Madrid, España.

(c) Alumnos egresados carrera de Ingeniería en Sonido, Universidad Tecnológica de Chile INACAP, Brown Norte 290, Santiago, Chile.

### **Abstract**

This work fits into the area of Acoustic Phonetics and presents the development of a methodology for the analysis of the Prosodia by studying variations in fundamental frequency, effective, reliable that allows to obtain an approach to melodic Chilean pattern.

For this study it was used a Database of Chilean Voices, accomplished by telephone interface containing 380 speakers with recordings of phrases Phonetically Balanced, Vowels, Series and Spontaneous speaks.

The methodology analysis design developed at the discretion of the scientific and statistical comparison of the change in the fundamental frequency, level of intonation and stress of the speakers from different areas of Chile. It also presents the development, through the database, to the Chilean formants letter, which identifies similarities and differences, through an automated measuring vocal parameters.

### **Resumen**

Este trabajo se enmarca en el área de la Fonética Acústica y presenta el desarrollo de una metodología para Análisis de la Prosodia mediante el estudio de las variaciones de la frecuencia fundamental, efectivo, confiable y que permita obtener un acercamiento al patrón melódico chileno.

El diseño de la metodología de análisis se desarrolla bajo el criterio de la comparación estadística y científica de la variación de la frecuencia fundamental, a nivel de Entonación y Acentuación de los hablantes por zonas de Chile.

Se presenta además la elaboración, mediante la base de datos, de la carta de formantes chilena, en la cual se identifican similitudes y diferencias existentes, a través de una automatización de la medición de los parámetros vocales.

## 1. Introducción

### 1.1. Fonética:

Es la ciencia que estudia los sonidos desde el punto de vista de su producción, transmisión y percepción. “La fonética se refiere a los sonidos en el habla, incluyendo su producción acústica y los procesos físicos y fisiológicos de emisión y articulación involucrados”. Este trabajo se basa en la *Fonética Acústica o Física* que describe las propiedades físicas de los sonidos transmitidos como Ondas Sonoras. Utiliza los métodos descriptivos de la Acústica.

### 1.2. Producción de sonidos del lenguaje desde punto de vista acústico.

El conducto vocal es esencialmente un tubo cerrado en un extremo (larínge) y abierto por el otro lado (tracto vocal, fosas nasales); este conducto puede cambiar su volumen por medio de los movimientos de los labios, en su parte superior, o de la lengua. Puede crear turbulencias o explosiones y siempre funciona como un resonador, el cual cumple la función de responder selectivamente –como un filtro acústico- a las frecuencias vibratorias que coinciden con la suya. Los órganos articulatorios cambian de posición, formando en la región supraglótica cavidades de formas y volúmenes diferentes (articulación de la voz); al ser estas cavidades (resonadores) diferentes, la frecuencia de respuesta es distinta, y en consecuencia el espectro de la onda que sale del resonador. De este modo se crean las distintas ondas sonoras que conforman el lenguaje. Los máximos de la función de transferencia del conducto vocal originan las formantes del espectro de sonido, que pueden ser definidos como las resonancias del conducto vocal. Las frecuencias se denominan frecuencias formantes. Las frecuencias de los armónicos de la onda sonora compleja dependen de la frecuencia fundamental y las frecuencias de los formantes dependen de los resonadores, aunque en más de un caso pueden coincidir. Las vocales pueden distinguirse desde el punto de vista acústico, de acuerdo a qué frecuencias se producen las resonancias del tracto, es decir, según sus formantes.

### 1.3. Prosodia

Es uno de los principales rasgos fonéticos que varía localmente dentro de una misma lengua. Se realiza -principalmente- mediante la evaluación del habla espontánea. La prosodia es una rama de la lingüística que analiza y representa formalmente aquellos elementos suprasegmentales o prosódicos de la expresión oral, tales como *el acento*, que suele presentarse sobre una sílaba y que depende de la fuerza de espiración y, por tanto de la velocidad con que el aire es expulsado; *el tono*, donde intervienen los músculos laríngeos en la tensión voluntaria de las cuerdas vocales para producir un número mayor o menos de vibraciones laríngeas; y *la duración*, que es la prolongación de la articulación de un sonido.

#### 1.3.1 El Acento

El acento se manifiesta como un contraste entre unidades acentuadas y unidades inacentuadas, pero acústicamente lo hace por medio de la melodía. Este, no tiene un único parámetro acústico, aunque son tres los que intervienen en la sílaba tónica con diferente grado de importancia. Se destaca la elevación y la duración de la frecuencia fundamental ( $f_0$ ), aunque es posible que sea una combinación de ambas la que proporciona la prominencia de la sílaba acentuada frente a sus vecinas inacentuadas. La amplitud no parece jugar un papel demasiado relevante en esta prominencia.

#### 1.3.2 Tono y Entonación

En el acento existen cambios de tono, pero son debidos a la fuerza de expulsión de aire. De manera voluntaria y sin una mayor fuerza de expulsión del aire se puede elevar o disminuir la frecuencia de vibración de las cuerdas vocales a través de la tensión. Nuestras lenguas no suelen utilizar fonológicamente estos medios tonales, pero sí utilizan la entonación; es decir, las variaciones frecuenciales a lo largo de un enunciado, que es el resultado de la integración de la melodía y el acento

#### 1.3.3 La Melodía

Consiste acústicamente en la curva que describe el primer armónico o tono fundamental ( $f_0$ ) en el grupo melódico, que es el conjunto de tonos que se suceden entre dos pausas. Suele coincidir con una unidad sintáctica: oración, proposición, vocativo, etc. Para el estudio de los contornos melódicos se suele tomar como unidad un grupo fónico. Existen diversos métodos para la eliminación de la

variabilidad individual y toda aquella que no sea relevante. El procedimiento se conoce como “estilización o normalización” de las curvas.

### 1.3.4 La Duración

Existe una duración intrínseca de cada sonido que depende o de su propia naturaleza; por ej. las vocales bajas suelen tener mayor duración que las vocales altas; las consonantes sordas suelen tener mayor duración que las vocales altas; y también suelen durar más que las sonoras correspondientes y las fricativas más que las oclusivas. En castellano la mayor o menor duración de las vocales depende de la sílaba tónica o átona respectivamente.

## 2 Análisis de Pitch

El Pitch de un sonido simple, es una variable subjetiva que nos indica si el sonido es alto o bajo. Desde el punto de vista musical, nos indica dónde se situaría el sonido dentro de una escala musical. Como parámetro objetivo nos evalúa la repetición del frente de ondas del sonido. Para una componente pura, el Pitch nos proporciona como valor objetivo la frecuencia. Para sonidos complejos su valor corresponde normalmente con el de la frecuencia fundamental ( $f_0$ ) cuando el espectro es una serie de armónicos de una  $f_0$  presente. Para determinar el Pitch se han elaborado muchos algoritmos. En esencia siguen dos planteamientos posibles correspondientes a los dos modelos básicos del mecanismo del oído para extraer el Pitch. Uno consiste en un análisis en el dominio temporal. Es decir sobre la onda de sonido se produce el análisis (teoría de la periodicidad). El otro procedimiento extrae el Pitch después de la descomposición en frecuencias que realiza la cóclea (teoría de la frecuencia). La mayoría de los algoritmos desarrollados trabajan a partir del análisis espectral del sonido, habiendo evidencias a favor de una teoría de la frecuencia (Duihuis, 1982). En el dibujo de cualquier técnica de análisis de voz que involucre llevar el análisis sobre unas pocas décimas de segundos, ocurre un proceso de promedios, particularmente en donde el espectro cambia rápidamente, con consecuentes perturbaciones de propiedades espectrales de variación rápida. En este tipo de análisis, el espectro computacional es un producto de la excitación espectral y del espectro del tracto vocal, y ambos no están listos por separados. Así, el efecto es de distorsión de la envolvente espectral del tracto vocal. Esto es notorio si el espectro excitado tiene antiresonancias (ceros) cercanas a las resonancias (polos) del tracto vocal. Ambos problemas pueden ser sobrellevados por el análisis sincrónico de Pitch en donde periodos individuales de señal de voz se identifican y analizan en forma aislada.

## 3 Desarrollo Experimental

A través de una base de datos de 380 locutores (Frases, Identificación, Prosodia, Serie de Números y Vocales), se busca estudiar la prosodia evaluando las variaciones que existen en el tiempo de la frecuencia fundamental sobre el relato de habla espontánea y texto leído por cada locutor, con el fin de encontrar las diferencias fonéticas propias del lenguaje para las distintas zonas de nuestro país.

### 3.1 Metodología y Análisis de la Variación de Pitch

El locutor efectuó un breve comentario de la forma que celebran en la ciudad que habita, las festividades del 18 de Septiembre (fiestas patrias en Chile). Este comentario fue efectuado de la forma más espontánea posible por la persona. Al revisar el total de las muestras, se seleccionó la frase que más se repetía entre los locutores “*dieciocho de septiembre*”, por lo tanto, se editó y separó entre hombres y mujeres por zonas, en “**Dieciocho**” y “**Septiembre**”.

Con el fin de obtener más y mejores resultados, también se seleccionó palabras de frases leídas por los locutores. De la frase: “**Las personas que habitan la región metropolitana padecen frecuentemente de enfermedades respiratorias**”, se extrajo por separado las palabras “**metropolitana**” y “**padecen**”. Se seleccionó la palabra “**padecen**”, ya que el verbo siempre “corta” aire, y es ahí, donde la gente manifiesta diferencias de acentuación, que se aprende en entornos familiares, y por ende, posiblemente regionales. La palabra “**metropolitana**” se seleccionó para completar el estudio de manera más acabada. Mediante el software editor de audio “*SOUNDFORGE*”, se cargaron las muestras, y se procedió a editar cada región de audio. No se aplicó ningún tipo de procesador, solo se eliminó la información no utilizable y se aplicó a las muestras un pequeño “fade in” y “fade out” de milisegundos, para eliminar cortes abruptos e información errónea en los datos.

Debido a la diferencia temporal de las muestras, se procedió a alinearlas en el tiempo mediante una herramienta específica del software editor de audio “*PROTOOLS*” llamada “*TIME COMPRESSION EXPANTION*”, la cuál comprime (elimina información redundante) o expande (interpola datos) una muestra de audio, modificando su tiempo pero sin variar su tono fundamental o Pitch. Con el método adoptado, se crearon sesiones en *Protools*, separadas por palabra y género, en donde se promediaron las duraciones de todas las muestras, obteniendo el alineamiento temporal.

Mediante el programa computacional “*PRAAT*”, utilizado por especialistas fonéticos para editar y analizar señales del habla humana, se obtienen los cálculos, análisis, gráficos y extracción de datos que competen en gran parte este estudio. Este programa es capaz de realizar variados cálculos dentro del área de la fonética, como por ejemplo, Análisis de Formantes, Voz, Intensidad y Pitch. El procedimiento de los análisis, gráficos y extracción de los datos se hacen a partir del archivo de audio cargado que se desea analizar. El programa muestra la curva de Pitch y da la opción para extraer el contorno de Pitch (Pitch Tier), en puntos de tiempo y frecuencia, sin la información expresada/inexpresada de la muestra. Por ejemplo, si PitchTier contiene dos puntos, 150 Hz. en un tiempo de 0.5 segundos y 200 Hz en un tiempo de 1.5 segundos, entonces esto debe ser interpretado como un contorno de Pitch que es constante en 150 Hz para todos los tiempos antes de 0.5 segundos, constantes en 200 Hz para siempre después de 1.5 segundos, y directamente interpolados para tiempos entre 0.5 y 1.5 segundos (p. ej. 170 Hz en 0.7 segundos, 210 Hz en 1.1 segundos, etcétera).

Luego, se extrae un documento de texto (“.txt”) de los datos de Pitch (Frecuencia en el tiempo), el cual se puede abrir, en Excel. Los datos ya extraídos, son de tipo texto y clase “Pitch Tier”. Cada punto, tiene un tiempo en milisegundos y un valor de frecuencia en Hz. Este procedimiento se realizó para el total de las muestras a analizar, obteniendo así, una planilla de datos para cada una de éstas.

Se utilizó un código “*MATLAB*” diseñado por Andrés Barrera, ingeniero civil acústico, para extraer del software “*PRAAT*”, únicamente los datos de tiempo y frecuencia de cada muestra. Para obtener el promedio de la variación de la frecuencia fundamental por zonas de nuestro país, se importan los documentos tipo texto de cada una de las muestras ordenadas por el código, en una hoja de cálculo en Excel, obteniendo una planilla de 2 columnas con los tiempos y frecuencias de todas las muestras, se ordenan de manera ascendente, creando una tercera columna en la planilla, con el promedio de las frecuencias por cada punto de los tiempos. Luego, se grafica el promedio de la variación de la frecuencia fundamental en el tiempo de las tres zonas superpuestas en un solo gráfico promedio, Se estudia cada palabra por sílaba o tramo también, debido a que el análisis global no es tan objetivo porque el comportamiento es irregular en toda la palabra.

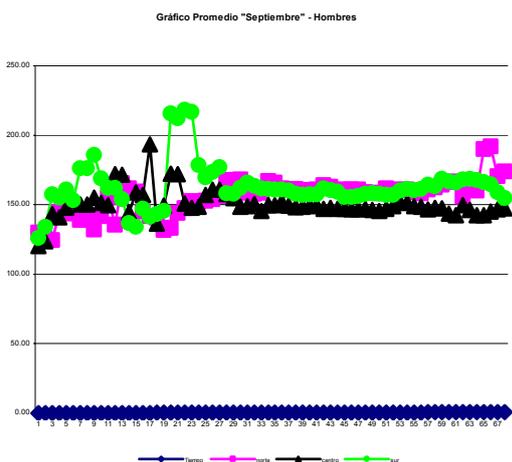


Figura 1. Gráfico Promedio, 3 zonas. “Septiembre” – Hombres.

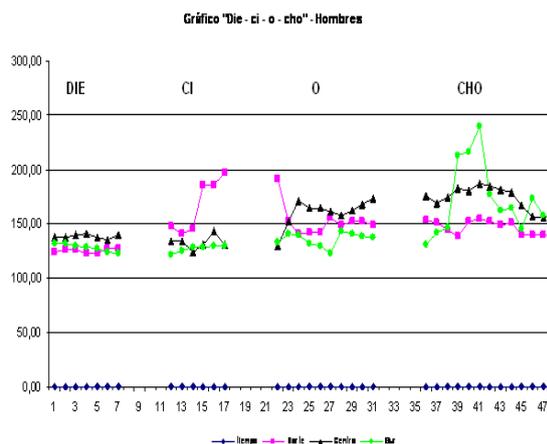


Figura 2. Gráfico “Die-ci- o-cho” – Hombres.

Mediante el software “*SPSS*” de Análisis Objetivo Estadístico, se importan los datos de frecuencia de las 3 zonas en columnas separadas. Se ocupa un Test estadístico de comparación de medias, de muestras pareadas o “*Paired – Samples T Test*”. Se obtiene como dato de salida, los resultados estadísticos del Test pareado, el cual se analiza e interpreta en este estudio.

**Tabla 1.** Resultado Estadístico “SPSS”

Paired Sample Test		
“Septiembre” Hombres	df	Sig. (2-tailed)
Pair 1 Centro-Sur	35	,729
Pair 1 Norte-Sur	35	,048
Pair 1 Norte-Centro	35	,821

### 3.2 Metodología y análisis de la Entonación.

La entonación se manifiesta fonéticamente mediante la melodía cuyo relato físico es la fo; por tanto, la relación entre entonación y melodía es equivalente a la mantenida entre fonemas y fonos. Una forma de normalizar las curvas melódicas de una frase pronunciadas es la siguiente: se toma nota de la frecuencia fundamental solamente de las vocales de cada sílaba. Una vez analizadas, tomamos el primer valor como referencia: es decir, ese primer valor será el valor cero de un porcentaje; el 100% será la octava de ese valor (se obtiene multiplicando por 2). Los cálculos posteriores se harán realizando una regla de 3.

De la frase leída por los locutores: “*Las personas que habitan la región metropolitana, padecen frecuentemente de enfermedades respiratorias*”, se extrajo la frase: “**enfermedades respiratorias**” para analizar y comparar. A partir de los datos de frecuencia, se estilizan en porcentajes y luego se reduce a una sola escala útil, en semitonos, las diferencias tonales que existan; se propone que sólo movimientos tonales superiores a más de 1.5 semitonos tendrían valor lingüístico. Para calcular los semitonos se emplea la siguiente fórmula:

$$(12/\text{LOG}_{10}(2)) \times (\text{LOG}_{10}(F0 \text{ primera medición} / F0 \text{ segunda medición})) \quad (1)$$

**Tabla 2.** Muestra 131 – Zona Norte – Mujeres.

131		EN	FER	ME	DA	DES	RES	PI	RA	TO	RIAS
	HZ	222	197	193	210	226	223	232	232	165	136
	%	0	-11,2	-13	-5,4	1,8	0,45	4,5	4,5	-25,9	-38,7
	ST	0	-2,1	-2,4	-1	0,3	0,1	0,8	0,8	-5,1	-8,5

### 3.3 Metodología y análisis de la Acentuación.

Mediante un análisis de energía, se normaliza a 0db el nivel máximo (la sílaba acentuada “DE”), de la palabra “**PADECEN**” extraída de la frase “**metropolitana padecen**”.

Mediante Software “*PRAAT*” se extraen valores de nivel y temporales de todas las muestras.

Se trabajó con el total de muestras de la base de datos, exceptuando aquellas muestras que no entregaron información de intensidad, debido al método de registro utilizado.

Se obtuvieron los valores de nivel y de tiempo del total de las muestras, para así promediar y graficar diferencias.

## 4 Desarrollo de la Carta de Formantes Nacional

Las resonancias del tracto vocal se denominan formantes. Son propiedades físicas del tracto vocal que producen amplificaciones en el espectro vocálico. En las vocales, *el primer formante (F1)* generalmente controla la amplitud del sonido y depende de la forma de la cavidad faríngea: cuanto más estrecha mayor frecuencia y viceversa. *El segundo formante (F2)* controla la inteligibilidad del sonido y se supedita a la posición de la lengua: si se eleva en la parte anterior, la frecuencia subirá en relación directa con la altura y la anterioridad alcanzadas; si es en la parte posterior, descenderá en relación inversa con la altura. *El tercer formante (F3)* está relacionado con la acción de los labios. Es más alto si éstos están estirados y más bajo si están redondeados.

En la carta de formantes, el eje de las abscisas (desde los 500 Hz. a los 4000 Hz.) lleva los valores de F2; sobre el eje de las ordenadas (desde los 200 Hz. hasta los 1000 Hz.), lleva los valores de F1.

En el eje de las ordenadas se observa el grado de abertura del conducto vocal, es decir, existe una relación constante y directa entre la abertura bucal y el nivel de frecuencias representado en el eje de las ordenadas. Por el contrario, el eje de las abscisas indica la longitud de la cavidad bucal, es decir existe una relación constante e inversa entre la longitud de la cavidad bucal anterior y el nivel de frecuencias representado en el eje de las abscisas.

Con el fin de construir un triángulo vocálico nacional, se utilizó un código “*MATLAB*”, diseñado por Víctor Espinoza, que automáticamente detecta la fo, reconociendo las frecuencias formantes.

El código entrega como resultado, planillas Excel de los valores de las formantes, tabla de promedios, desviaciones y un gráfico formante, en el cuál se encuentra la primera y segunda formante de todos los locutores estudiados y se logra una idea de la pronunciación vocálica de Chile:

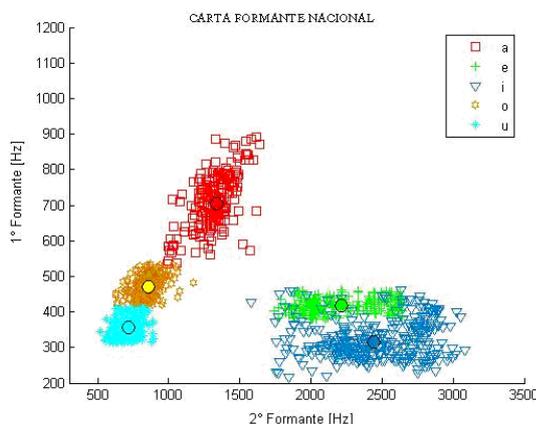


Figura 3. Carta Formante Nacional.

## 5 Análisis Variación Frecuencia Fundamental

### 5.1 Habla Espontánea y Texto Leído

En el análisis de resultados, se obtuvieron gráficos de valores promedio de la variación de la frecuencia fundamental en el tiempo en conjunto con los resultados estadísticos que entrega el software “SPSS”. Se analiza por sílabas y en su totalidad la palabra “Septiembre” y “Dieciocho” de Hombres y Mujeres para habla espontánea y la palabra “Metropolitana” y “Padecen” de Hombres y Mujeres para texto leído. Existen variaciones de la frecuencia fundamental, que en el análisis estadístico, se objetivizan a través de los resultados que se obtienen de los valores estadísticos que muestra el Test efectuado en “SPSS”, de la diferencia de las medias, como por ejemplo, el parámetro “SIG. (2 –TAILED)”, el cual refleja que las comparaciones entre zonas pareadas, no son iguales. Claro está, que debemos considerar que es relevante la ubicación geográfica de los locutores, ya que el estudio se realizó por zonas de nuestro país y a la vez por diferencia de sexo. Se puede decir, que existe una menor igualdad de variación de la frecuencia en Mujeres que en Hombres, ya que la diferencia de las medias hace que el valor de la diferenciación sea más cercano entre ellas como muestran los resultados del “SPSS”. El análisis de la palabra de texto leído “Padecen”, refleja en el gráfico promedio que la variación de la frecuencia fundamental subjetivamente es menor que en la palabra “Septiembre”. Se aprecia en el análisis estadístico objetivo, que la diferencia de las medias es menor a la de habla espontánea y también los valores de “SIG (2-TAILED)”. Debido a que la lectura de esta palabra estaba dentro de una frase, y fue leída en forma natural con el fin de que fuera dicha de la mejor manera fonéticamente balanceada, es razón de que se encuentren menos diferencias de las variaciones de frecuencia entre las zonas.

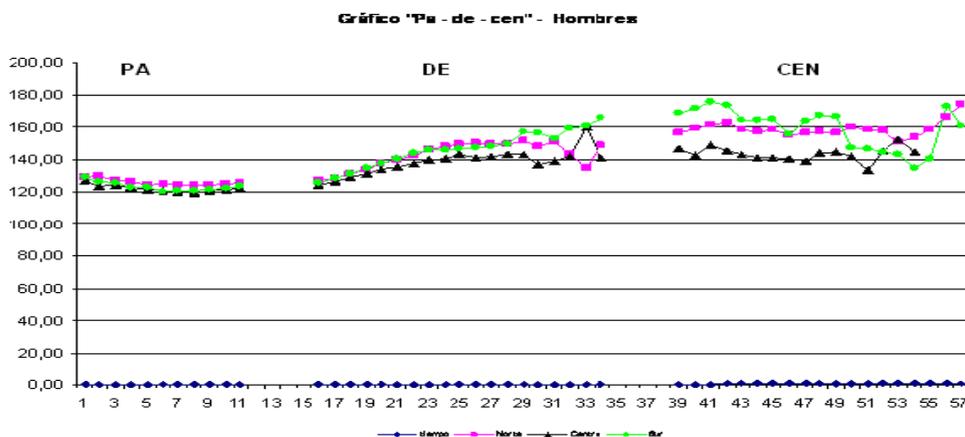


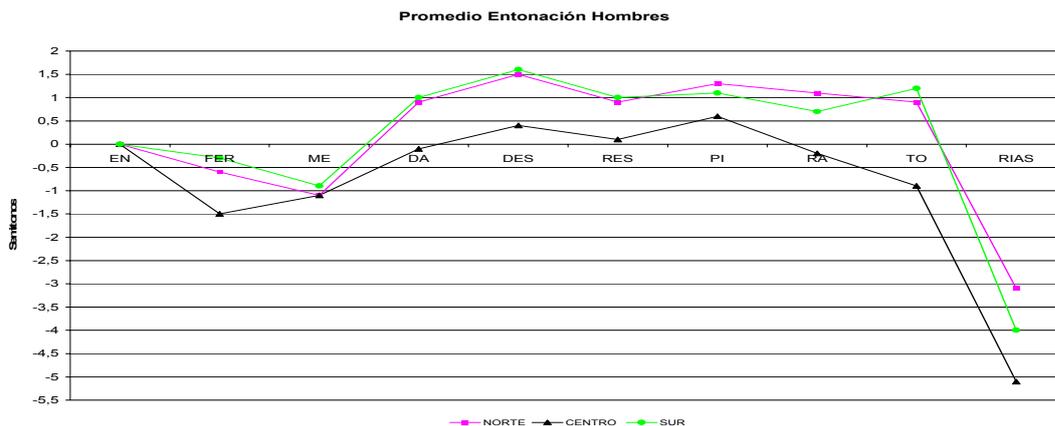
Figura 4. Gráfico “PA – DE – CEN” – Hombres.

**Tabla 3. Resultados “SPSS”**

Pareid Samples Test			Pareid Samples Test			Pareid Samples Test		
Sílaba “Pa”	df	Sig.	Sílaba “De”	df	Sig.	Sílaba “Cen”	df	Sig.
hombres		(2-tailed)	hombres		(2-tailed)	hombres		(2-tailed)
Pair 1 Centro-Sur	10	,013	Pair 1 Centro-Sur	18	,000	Pair 1 Centro-Sur	18	,007
Pair 1 Norte-Sur	10	,000	Pair 1 Norte-Sur	18	,064	Pair 1 Norte-Sur	18	,926
Pair 1 Norte-Centro	10	,000	Pair 1 Norte-Centro	18	,024	Pair 1 Norte-Centro	18	,009

**5.2 Entonación**

Se obtuvieron valores promedios por zona y sexo, de diferencia tonal en semitonos. Se grafican por zona y sexo la Variación en Semitonos, para visualizar las diferencias tonales entre las zonas.



**Figura 5. Gráfico Promedio Entonación en semitonos – Mujeres.**

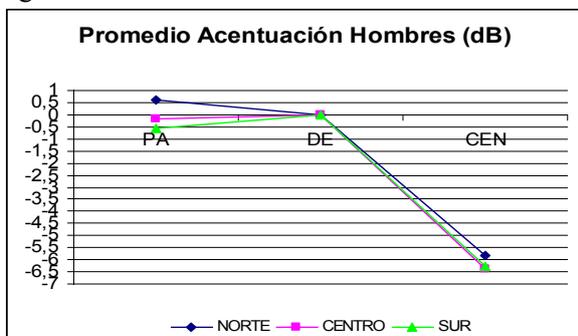
**5.3 Acentuación**

Primero se normaliza en dB. la energía de la palabra, obteniendo los datos para promediar y posteriormente se promedia en % los datos del tiempo en que se pronuncia cada sílaba de la palabra.

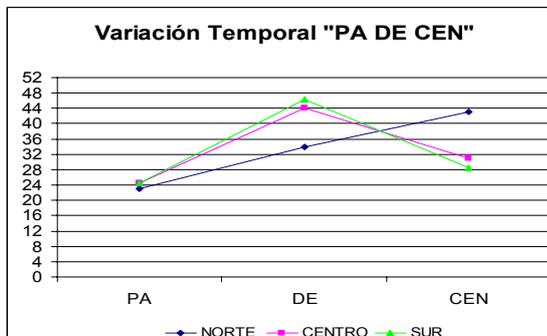
**Tabla 4. Promedios de energía y resultados temporales de la palabra “PADECEN”.**

Hombres		PA	DE	CEN	Mujeres		PA	DE	CEN
dB.	NORTE	0,61	0	-5,81	dB.	NORTE	-0,6	0	-3,34
	CENTRO	-0,19	0	-6,41		CENTRO	-0,6	0	-3,9
	SUR	-0,54	0	-6,29		SUR	1,83	0	-0,9
(%).	NORTE	23	34	43	T (%)	NORTE	19,88	31,34	48,5
	CENTRO	24,53	43,99	31,02		CENTRO	19,84	29,93	49,83
	SUR	24,43	46,36	28,42		SUR	23,1	37	39,5

Los resultados gráficos de acentuación en nivel y tiempo de las 3 zonas divididas por sexo son los siguientes:



**Figura 6. Gráfico Promedio Acentuación–Hombres**



**Figura 7. Gráfico Variación Temporal – Hombres.**

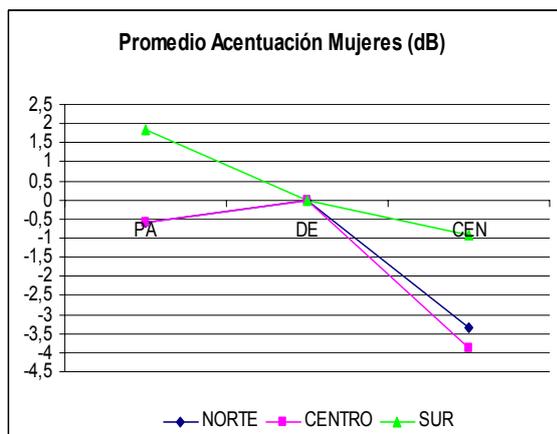


Figura 8. Gráfico Promedio Acentuación.

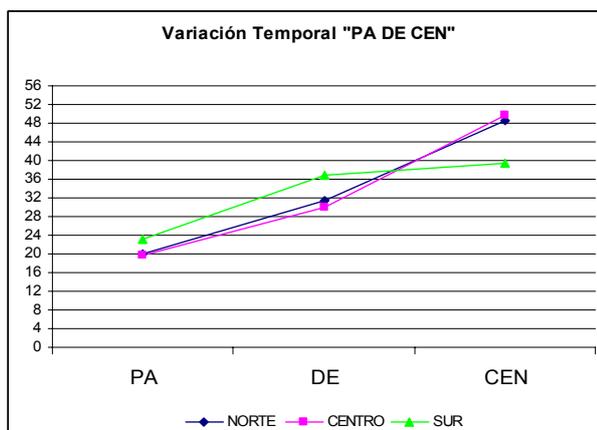


Figura 9. Gráfico Variación Temporal Mujeres.

## 6 Conclusiones

El hecho de ocupar una base de datos, realizada mediante interfaz telefónico, contribuyó a que los resultados de promedios no incluya la totalidad del registro de los locutores. Sin embargo, se logró abarcar gran una cantidad de muestras para nuestro estudio.

El hecho de no ocupar un método conocido como el de Goldstein o el de Terhardt, fue porque la meta era realizar un método de diseño de una metodología comparativa de la variación de frecuencia fundamental estadística y científicamente, y además, los métodos descritos se basan en el estudio espectral de análisis de Fourier y una ventana Hamming, los cuales pretenden determinar la frecuencia fundamental a partir de consideraciones psicoacústicas. El software "PRAAT", puede variar la frecuencia de muestreo, como también los parámetros a medir sin depender de la duración de las muestras. También, el desarrollo de nuestro método se basa en algoritmos simples y de una estructura mecánica del estudio de Pitch, a través del uso de software menos complejos en los cuales se obtienen las siguientes ventajas:

- Procesamiento de grandes cantidades de datos, en menor tiempo y costo para investigaciones futuras.
- Entrega fácil de datos útiles para diferentes áreas del estudio de la voz, patológicas, acústica forense, y otros.
- Es un proceso mecánico, ya que sólo se tiene que seguir los pasos y no da cabida a errores, por lo tanto es de fácil manejo para cualquier persona medianamente entendida en el tema.
- El uso de software de fácil aplicación y de interfaces más amigables.

### 6.1 Variación de la Frecuencia Fundamental

Considerando las tres zonas analizadas, es posible darse cuenta, tanto a través de los gráficos promedio como en los gráficos segmentados por sílabas, que existe una tendencia, para los tres casos, de conservar un patrón melódico final.

#### 6.1.1 Carta de Formantes Chilena

En el gráfico de formantes, se aprecia "el triangulo vocálico nacional", el cuál está formado por los puntos de colores distintos destacados en forma circular que representan el promedio de todos los puntos de cada una de las vocales que están en orden de izquierda a derecha, que son: u, o, a, e, i.

Se percató que en la mayoría de los locutores hombres existe una mayor densidad de frecuencias armónicas por banda de frecuencias que las mujeres. En el gráfico se observa la concentración de todos los hablantes en las vocales "o", "u" y en menor medida, pero no menos despreciable, en la vocal "a", esto nos dice que en Chile se habla de forma en que se destacan estas vocales. En cambio no así con las vocales "e", "i", en las cuales existe mayor diferencia de formantes.

La base de datos se realizó a través de una interfaz telefónica, donde el ancho de banda es limitado, por eso estimamos que el método de medición y construcción de "el triangulo vocálico" es correcto,

pero no es certero por la calidad de las muestras, pero si se asemeja a una realidad nacional, por lo tanto, se deja abierta la invitación a usar este método para encontrar una carta formante más precisa.

### 6.1.2 Habla Espontánea

- Del análisis de las 3 zonas para la palabra “*Septiembre*”, es posible darse cuenta que existe una diferencia de la variación de la frecuencia. En los gráficos se observan diferencias de variación considerables en la primera parte de la palabra, en la zona centro comienza con frecuencias más altas que en las zonas norte y sur. Luego, al finalizar la palabra, tienen la misma tendencia, pero en distintas frecuencias; la zona norte finaliza con una variación en frecuencias más alta que en las demás zonas.
- De los resultados obtenidos en las tablas que nos entregó el SPSS, podemos decir objetivamente, a través de los valores de SIG. (2 –TAILED), que en las 3 zonas se habla diferente unas de otras, en los hombres hay más diferencias entre las zonas que las mujeres al decir la misma palabra.
- Se puede decir que hay menos diferencias entre las zonas NORTE-SUR, en hombres y mujeres, ya que el valor del parámetro SIG. (2 –TAILED), es distinto de cero, y cabe recalcar que entre las zonas NORTE-SUR, sólo de mujeres, hablan muy parecido en la variación de la frecuencia ya que este valor es mayor que 0.05, que es nuestro valor límite para saber si es diferente o parecida la comparación.

### 6.1.3 Texto Leído

- Podemos concluir que en los gráficos hay diferencias subjetivas. En el análisis de la palabra “PADECEN”, el locutor se concentra más en lo que está diciendo, por lo tanto la variación de la frecuencia será menor a una palabra de habla espontánea. Se nota más la variación en el gráfico de locutores de hombres que mujeres, esto nos dice que las mujeres chilenas hablan más parecido que los hombres.
- De los resultados obtenidos en la tabla de datos del software SPSS, podemos afirmar que entre las zonas NORTE – SUR los hombres hablan más parecido que entre las demás zonas ya que el valor SIG. (2 –TAILED) alcanza a los .518, en cambio entre las mujeres las zonas que más hablan en forma similar son entre la zona NORTE-CENTRO.
- En el análisis por sílaba, podemos destacar que entre los hombres de las zonas NORTE-CENTRO en las sílabas “DE” y “CEN” son muy similares en la variación de la frecuencia fundamental, ya que nos arrojan valores de SIG. (2 –TAILED) de .064 y .926 respectivamente.

### 6.1.4 Entonación

- La entonación en la frase de análisis “Enfermedades Respiratorias”, se manifiesta con un significado de afirmación que comienza con una elevación de tono desde la fase inicial, se mantiene uniforme en la fase media hasta la última sílaba tónica, desde la que desciende en cadencia hasta el final.
- Aunque existen variaciones tonales dentro de una misma zona, ésta es propia del rasgo que añade al enunciado un significado.
- Entre las zonas, existe una diferencia tonal de más de 1,5 semitonos en la fase final de la curva, que empieza en el último elemento tónico (“TO”) de la frase, hasta su término. En los Hombres, entre el Norte y Centro existe una diferencia de 1,8 y 2 semitonos, y entre la zona Centro y Sur hay 2,1 semitonos. En las Mujeres, esta diferencia se da entre la zona Norte y Sur con 1,2 st., y entre la zona Centro y Sur es de 1,6 y 2,4 st.

### 6.1.5 Acentuación

- La palabra de análisis “PADECEN” es tónica, o sea que posee una sílaba tónica y es trisílaba, las dos primeras sílabas son abiertas (terminan en vocal) y la última es cerrada (termina en consonante). También según el lugar que ocupa la sílaba tónica, la palabra lleva su acento prosódico en la penúltima sílaba.
- El nivel en la primera sílaba tanto en los Hombres como en las Mujeres, no es significativa, ya que los valores son despreciables. Aún con estos valores, se puede decir que el nivel de energía en la primera sílaba, en los hombres disminuye en el centro y sur, mientras que

aumenta en el norte, y en las mujeres, disminuye en el norte y centro, mientras que en el sur aumenta.

- En la última sílaba “CEN”, se encuentra una disminución de nivel en los hombres de 5 dB. en la zona norte y 6 dB. en la zona centro y sur, mientras que en las mujeres en la zona norte y centro es de alrededor de 3 dB. Estos valores son considerables, lo que indica que existe en hombres y mujeres una disminución de nivel en la sílaba no acentuada.
- La variación temporal de las sílabas difiere entre las zonas. En la zona norte se tardan en pronunciar más la última sílaba, mientras que en el centro y sur, se tardan más tiempo en la segunda sílaba “DE”. Las Mujeres en todas las zonas tardan más tiempo en pronunciar la última sílaba, en mayor o menor medida.

## **7 Alcances del Estudio de la Variación del Pitch en Chile**

Las palabras se han ido desvirtuando desde la colonización a través de la influencia indígena, en como se han pronunciado distintos fonemas desde la península ibérica hasta el “chilenismo”, lo que influye directamente en la pronunciación, y así en la variación de la frecuencia, acentuación y entonación.

Los valores obtenidos científicamente, nos indican que en la zona sur por ejemplo, existe una mayor variabilidad de la frecuencia fundamental que en el resto de las zonas de Chile. El rasgo melódico nos indica, que la expresión estudiada, presenta una variación considerable, lo que confirma la creencia popular de que se habla “cantado”. En tanto, la zona norte presenta las mismas características de variación por sobre la zona centro, lo que nos indica que en la zona centro se habla de manera más “plana”.

En los análisis de acentuación existe una variación de energía poco considerable, pero no despreciable, que nos indica que en todas las zonas la sílaba acentuada prosódicamente no es necesariamente la de mayor nivel respecto a las demás sílabas.

Las futuras aplicaciones de nuestro estudio son las siguientes:

- Realización del mismo método de estudio en los distintos países de habla hispana.
- Obtención de las diferencias del patrón melódico entre países.
- Centralizar este método en regiones y ciudades (diferentes estratos sociales, edad, etc.)
- Estandarización de resultados para la acústica forense.

A modo de reflexión, en este estudio se logra objetivizar algunos parámetros relativos del patrón melódico chileno. Mediante este estudio certero y representativo logramos caracterizar la voz chilena. Se propone realizar este método estadístico mediante una base de datos fonética de mejores condiciones acústicas y sistemas electroacústicos.

## **8 Referencias**

- Cortés, Francisco; Vásquez, Juan (2004). “Construcción de una base de datos fonética de voces chilenas”; Trabajo Tesis, Santiago, Chile, 2004.
- Recuero, Manuel; Grez, Guillermo (2004). “Construcción de una base de datos de voces chilenas por interfaz telefónico”; Trabajo de Suficiencia Investigadora; Universidad Politécnica de Madrid, España.
- Ortega, Javier; González, Joaquín. “Procesado de Voz: Análisis y Parametrización de Señal”; Universidad Politécnica de Madrid, España.
- Espinoza, Víctor (2005). “Automatización en la Medición de Parámetros Vocales”; Revista Sonido y Acústica Universidad Pérez Rosales, Volumen2, Número 8, Agosto. Santiago, Chile.
- Martínez, Eugenio. Análisis Espectrográfico de los sonidos del habla.
- Romero, José; Cerda, Salvador; “Uso del análisis de multirresolución para calcular el Pitch de señales en presencia de ruido”; LAB. Acústica. Dep. Física. U. Valencia.
- Martínez, Eugenio. “El sonido en la comunicación humana, introducción a la fonética”