

UNIVERSIDAD DE CHILE
FACULTAD DE MEDICINA
ESCUELA DE FONOAUDIOLÓGÍA

ESTUDIO ACÚSTICO DE LA VARIACIÓN INTERLOCUTOR EN
SUJETOS HABLANTES NATIVOS DEL ESPAÑOL DE SANTIAGO DE
CHILE.

INTEGRANTES

Ignacio Cárdenas Pinto
Humberto Ceballos Cartes
Shang-Hsueh Lee
Wladimir Pavez Rocha
Claudio Terrisse Gutiérrez

TUTOR PRINCIPAL
Flgo. Luis Romero Romero

TUTOR(ES) ASOCIADO(S)
Lg. Christian Peñaloza

METODOLOGA
Prof. Ilse Lopez

Santiago - Chile
2010

**Dedicado a nuestros padres, cuyo apoyo fue imprescindible
durante toda la carrera.**

**A nuestros amigos que estuvieron allí para tendernos una mano en los
momentos difíciles, al profesor flgo. Luis Romero cuya instrucción y motivación
nos orientó durante esta ardua tarea.**

**Finalmente, al pequeño Tomás Ceballos, que nació y creció junto a este
seminario de título. Cada sonrisa, cada paso que él daba,
nos impulsaba a seguir.**

Gracias.

Resumen:

1.- Introducción:

La voz es un fenómeno comunicativo complejo propio del ser humano, el cual posee características fonético-acústicas que hacen posible su medición objetiva. Dichas características están dadas por condiciones fisiológicas y psicológicas individuales, lo cual convierte a la voz en un elemento útil en la identificación de sujetos. Sin embargo, el desarrollo de la fonética forense, disciplina encargada del reconocimiento de individuos a través de su voz, es aún incipiente en Chile, razón por la cual se desconoce en el contexto nacional la existencia de constantes y variaciones fonético-acústicas que permitan identificar a un sujeto y al mismo tiempo diferenciarlo de los demás.

La presente investigación pretende identificar variaciones y constantes fonético-acústicas en un grupo de sujetos hablantes nativos del español de Chile.

2.- Metodología:

Diseño investigación: Esta investigación se basa en estudio no experimental descriptivo de correlación trasversal a realizarse entre marzo y noviembre del año 2010.

Población: El grupo de estudio consta de 60 hablantes nativos chilenos, 30 hombres y 30 mujeres de 19 y 23 años de edad para homogeneizar la muestra, y que no padecen alguna patología vocal por la confiabilidad de la misma.

Instrumentos de recolección de datos: Para la recolección de datos se utilizará un instrumento con la estructura del apartado de los corpus que se usaron en la base de datos Ahumada-Gaudí, es decir, una versión extendida a 182 palabras del texto "El Abuelo",

Procedimiento: La revisión y el análisis de los corpus obtenidos se realizaron con el software de análisis espectral Praat.

Las variables a analizar son: la frecuencia fundamental (F0), los formantes medios (F1, F2, F3) y la distancia inter-formántica F0-F1; F2-F3; F0-F3 de las vocales /i/ /u/ /a/ y las consonantes nasales /n/ y /m/.

Resultados y Conclusiones:

Si bien el F0 es un parámetro muy estable tanto en el análisis interlocutor como en el intralocutor, no presta grandes antecedentes para el objetivo del presente seminario, por cuanto no nos permite realizar individuación de un sujeto, es decir, no permite distinguirlo de los demás hablantes de una lengua ni nos entrega datos de valor para caracterizar la voz de un sujeto desconocido, ya que el promedio de este valor se mezcla mucho con el de los hablantes que lo rodean y que tienen una lengua común.

En el fonema /a/, si comparamos las formantes F0, F1, F2, F3 y las respectivas diferencias interformánticas del alófono [a] átono v/s [á] tónico dentro de una palabra, observamos que los promedios y las desviaciones estándar son muy similares, por ende, concluimos que ambos alófonos, tanto tónico como átono podrían ser analizados de manera conjunta en un sujeto sin que esto altere el resultado final.

En los análisis de las mujeres encontramos que el parámetro más estable entre sujetos (interlocutor) es la F1 de la vocal /a/ tónica en base a la poca variación del promedio y una baja desviación estándar. Por otro lado, un dato relevante de mencionar es el F0 del fonema /a/ como un dato muy estable entre todos los sujetos de la muestra.

En los análisis de los hombres encontramos que el parámetro más estable entre sujetos (interlocutor) es la F1 de la vocal /a/ tónica en base a la poca variación del promedio y una baja desviación estándar.

Tabla de contenidos

1. Introducción.....	1
2. Marco Teórico	
2.1. La Voz.....	4
2.2. Aparato Fonador.....	5
2.3. Cavidades Supraglóticas.....	8
2.4. Resonancia y fonética acústica.....	9
2.5. Elementos Segmentales y Suprasegmentales.....	11
2.7. Identificación de Individuos.....	14
2.8. Corpus oral.....	15
2.6. Generalidades de la física acústica aplicada a la voz.....	17
2.7. Análisis de la voz en Praat.....	21
2.8. Coarticulación.....	21
3. Problema.....	22
4. Objetivo General.....	22
5. Objetivos Específicos.....	23
6. Variables	
6.1. F0.....	23
6.2. Formantes Medias.....	24
6.3. Fonemas /a/, /i/, /u/ tónicas.....	24
6.4. Fonemas /a/, /i/, /u/ átonas.....	25
6.5. Fonemas /m/ y /n/ trabantes.....	25
6.6. Fonemas /m/ y /n/ en articulación con vocales.....	25
7. Metodología	
7.1. Diseño de Investigación.....	26
7.2. Población y grupo en estudio.....	26
7.3. Instrumento de recolección de datos.....	26
7.4. Procedimientos.....	27
7.5. Pasos a seguir en la toma y análisis de muestras de voz/habla.....	29
8. Resultados.....	32

9. Conclusiones y discusiones.....	43
10. Bibliografía.....	48
11. Anexos.....	51

1. Introducción

Resulta interesante en el contexto actual la realización de un estudio acústico de la variación intra e interlocutor, en un intento por dilucidar patrones acústicos que sean confiables y estables, para diferenciar e identificar a los hablantes. Así comenzar a construir un fundamento teórico a la ejecución práctica de la fonética forense en Chile.

La voz es uno de los pilares esenciales de la comunicación humana. Esta es producida por la movilización de aire a través de la laringe, donde se realizan modificaciones en el flujo aéreo y se crea un sonido débil que necesita de amplificación y filtrado, que son llevados a cabo por los resonadores faringonasales (Jackson-Menaldi, 1992). Estos resonadores son un conjunto de cavidades e interrupciones en el camino del flujo aéreo (superior a la laringe), que filtran y amplifican armónicos de aquel sonido, resultando así una voz con características físico acústicas propias de la fisiología vocal. El estudio de los factores físico-acústicos de la voz humana esta a cargo de la fonética acústica, una rama de la lingüística que se dedica a investigar los componentes físicos de los sonidos del habla y que nos permite realizar un perfil fonético de las características de la voz (Rabiner y Schafer, 1978; Rose, P, 2003).

Las características acústicas de la voz han hecho factible la medición de sus componentes y la instauran como un fenómeno susceptible de ser caracterizable y objetivable. Por esta línea de pensamiento, se ha llegado a decir que existen elementos vocales que hacen diferenciables cada una de las voces, es decir, que hacen posible la diferenciación objetiva entre una voz y todas las demás.

Es conocida la facultad del ser humano de identificar voces, razón por la cual la voz probablemente ha sido utilizada como elemento de reconocimiento por miles de años. Sin embargo, los registros de su uso con fines legales son recientes, principalmente porque las tecnologías de la voz corresponden a invenciones modernas: herramientas tales como el teléfono, grabadoras de sonido y el espectrógrafo surgieron en los inicios del siglo XX. Estos instrumentos facilitaron el estudio de la voz y su utilización en el plano legal (Nolan, 2001; Erickson, 2005). Los

espectrógrafos son instrumentos (o software) que entregan gráficos donde se muestra la señal acústica de la voz en tres dimensiones: tiempo vs frecuencia vs intensidad llamados espectrogramas (Bonastre, 2001). En la década del cuarenta, estos fueron comparados con las huellas dactilares, razón por la cual surgió el término “voiceprint” o “huella de voz” (Erickson, 2005). Sin embargo, las huellas dactilares de un individuo corresponden al indicio directo de una característica física virtualmente invariable (a no ser por destrucción de la yema del dedo), mientras que la voz es un elemento que está lejos de serlo (Nolan, 2001).

La voz es un fenómeno inherentemente complejo, pues está determinado por variaciones que competen a distintas dimensiones, relacionadas por una parte al lenguaje mismo y su uso (determinado por el contexto), y por otra parte, a condiciones propias del individuo, asociadas a su estado de salud, tanto en el plano fisiológico o de los órganos que producen el habla, como también en lo psicológico y emocional (Bonastre, 2001; Nolan 2001; Rose, 2003).

Los órganos que producen la voz muestran gran flexibilidad, debido a que también corresponden a órganos en común; de la respiración y de la deglución (Nolan, 2001). El resultado de esta variabilidad es que, estrictamente en el sentido físico, ningún individuo puede decir dos enunciados exactamente de la misma manera (Nolan, 2001; Rose, 2003).

La voz es capaz de entregar gran cantidad de información respecto de su locutor, como su género, estado emocional y de salud (Rose, 2003). Esta información no está separada en diferentes canales sino que está toda mezclada en el mensaje. Aun así, la voz puede ser analizada a través de diferentes parámetros, los que pueden repetirse en distintos individuos dentro de una población, pero que en su conjunto tienen más probabilidades de ser asociados a un solo sujeto.

No existen estudios poblacionales que establezcan cuantos parámetros diferentes se deben considerar a la hora de determinar de manera confiable a un hablante en específico (Nolan, 2001). Asimismo, en Chile no existen estudios referentes a las características fonético acústicas del idioma en relación a su

utilización en el plano legal. Es por esta razón que los estudios de fonética acústica deberían tomar más importancia debido al uso que se les puede dar para la identificación de sujetos mediante evidencias en audio. Realizar un estudio de este tipo desde la perspectiva fonoaudiológica, resulta un avance en el área considerando que el profesional fonoaudiólogo posee un conocimiento de la voz, el habla y el lenguaje más diverso que otras disciplinas, además de un manejo biométrico, fisiológico y acústico que logra asociar de manera integral la lingüística, la audiología, la física acústica, la odontoestomatología y la foniatría. Entonces con este enfoque en las competencias del fonoaudiólogo, resulta pertinente realizar un análisis fonético-acústico de la voz y del habla más rico y con mayores consideraciones.

2. Marco Teórico

2.1. La Voz

La voz es un fenómeno físico-acústico inherentemente complejo, ya que está determinado por variaciones que competen a distintas dimensiones, relacionadas por una parte a condiciones propias del individuo, su estado de salud físico, y psicológico; y por otra, al sistema del lenguaje, pues su uso es eminentemente comunicativo, y varía de acuerdo al contexto (Bonastre, 2001; Nolan 2001; Rose, 2003). Es por esto que la voz es capaz de entregar gran cantidad de información acerca del locutor que la produce (Rose, 2003).

La voz humana se describe en los manuales de anatomía como una función secundaria dentro del aparato respiratorio (Bernal *et al*, 2000). La laringe en su origen se establece como una estructura de protección de las vías respiratorias inferiores. Posteriormente, con el transcurso de la evolución, adquiere la función fonatoria, que va ligada al desarrollo normal del ser humano, el cual da a lugar al lenguaje oral, comunicativo, simbólico y abstracto, tan variado como cantidad de culturas existen en el planeta. La expresión del lenguaje se concreta en el habla, que es la articulación de un mensaje voluntario ejecutado por el aparato bucofonatorio (Casado y Adrián, 2002).

La función fonatoria corresponde a la producción de ondas de sonido mediante la vibración de estructuras en el interior de la laringe (Morrison y Rammage, 1996) o la utilización inteligente de los mecanismos que la laringe emplea para la protección de las vías respiratorias bajas (García-Tapia y Cobeta, 1996). Es un proceso complejo, que se produce cuando la aducción de las cuerdas vocales ofrece la resistencia necesaria al flujo aéreo para poder entrar en vibración, de tal manera que la energía aerodinámica de corriente continua, el aire espirado, es convertida en energía acústica de corriente alterna (Casado *et al.*, 2002; Romero *et al.*, 2007).

En la fonación participan distintos sistemas, como el sistema neuro-muscular, respiratorio, de resonancia y articulatorio, y también hay que mencionar aquellos sistemas que inician y coordinan a los sistemas anteriores, que son los subsistemas neurológico central y periférico (Adrián *et al.*, 2002).

Sin embargo, la vibración de las cuerdas vocales produce un sonido que escuchado aisladamente se asemeja más a un zumbido de tonalidad grave (Casado y Adrian, 2002). Este fenómeno, constituye una parte del gran proceso que implica la producción de la voz, y solo una pequeña parte del accionar del sistema que está encargado de ello, el aparato fonador (Romero *et al.*, 2007).

El estudio de la voz, se comprende como uno mucho más amplio que el de la laringe y la sonorización del aire espirado, pues implica también la transmisión de un mensaje que depende del contexto. De hecho, en ausencia de laringe, también se puede producir voz como es el caso de los laringectomizados (Casado *et al.* 2002).

2.2 Aparato Fonador

A grandes rasgos, el aparato fonador se puede dividir en 3 partes (Bernal, 2000; Casado *et al.*, 2002): Cavidades infraglóticas, Cavidad glótica y Cavidades supraglóticas. Para su descripción, es necesario tener en cuenta que estas estructuras, a su vez, constituyen parte de otros sistemas, con otras funciones que les son propias (Romero y Villanueva, 2007 p.33). Esta es la razón por la cual en la producción de la voz están implicados los sistemas neuro-muscular, respiratorio, de resonancia y articulatorio, como también aquellos que inician y coordinan a los sistemas anteriores, que son los subsistemas neurológico central y periférico (Casado *et al.*, 2002).

Cavidades infraglóticas

También son llamadas tracto vocal infraglótico (Casado *et al.*, 2002). Corresponden al compartimiento anatómico ubicado debajo de las cuerdas vocales y que interviene en la producción de la voz (Casado y Adrian, 2002). Está compuesto por diafragma, pulmones, bronquios y tráquea (Bernal, p.25). Su función es proporcionar la corriente de aire espirada necesaria para producir el sonido (Bernal, p25). Estas estructuras generan la potencia o volumen de voz mediante la creación de una fuerza que se dirige en forma de corriente aérea hacia la glotis (Casado *et al.*,

2002). Dentro de este contexto, la voz es considerada como una espiración sonora y activa llamada soplo respiratorio (Casado *et al.*, 2002).

La forma en que las cavidades infraglóticas participan en la producción de la voz es la siguiente: la emisión del soplo fonatorio viene precedida por una inspiración, llevada a cabo principalmente por el músculo diafragma (Casado *et al.*, 2002). Este músculo, en conjunto con los intercostales externos (Romero *et al.*, 2007) controla la inspiración, pues mediante su contracción provoca el despliegue e hinchado pulmonar (Bernal, 2000). Inmediatamente después, se genera una espiración activa, ascendiendo el diafragma y descendiendo las costillas y el esternón, de manera que la dimensión del tórax se reduce (Casado *et al.*, 2002; Bernal, 2000). Los bronquios y la traquea, por otro lado, funcionan como canales de transmisión del flujo aéreo (Bernal, 2000).

Cavidad glótica

La cavidad glótica está formada por la laringe, en la cual se encuentran las cuerdas vocales, pliegues musculosos responsables de la producción de la vibración básica para la generación de la voz. En medio de dichos pliegues, existe un espacio triangular llamado glotis. Cuando el aire sale de los pulmones y pasa por la hendidura glótica, provoca la vibración de las cuerdas vocales, la que puede variar en frecuencia e intensidad según varíe la masa, longitud y tensión de las cuerdas vocales (Bernal, 2000).

La laringe está formada por cuatro unidades anatómicas básicas: esqueleto, musculatura extrínseca, musculatura intrínseca y mucosa laríngea (Casado *et al.*, 2002). Esta última, recubre a modo de película todas las demás estructuras.

El esqueleto de la laringe está conformado por un hueso suspensorio, el hioides (Casado *et al.*, 2002) y los cartílagos tiroideos, epiglotis, cricoides, aritenoides, corniculados y cuneiformes (Pratter-Swift, 1986 citado en Romero *et al.*, 2007). Los primeros 3 son impares y de mayor tamaño, mientras los otros 3, son pares y de menor tamaño (Romero *et al.*, 2007). Estos elementos a su vez se encuentran unidos por las fascias o membranas tirohioidea, cricotiroidea y cricotraqueal, lo que los convierte en una estructura anatómico-funcional (Casado *et al.*, 2002).

La musculatura extrínseca tiene directa relación con el esqueleto laríngeo, debido a la función suspensoria (Casado *et al.*, 2002), elevadora y descendora (Romero *et al.*, 2007) del mismo. Además, las uniones existentes entre los cartílagos laríngeos son flexibles, y su variación, cuando se eleva el esqueleto de la laringe, depende de estos músculos (Casado *et al.*, 2002). Se clasifican según su localización y función en infrahioideos, los cuales sujetan la laringe al tórax, y suprahioideos, los que sujetan la laringe a la mandíbula. Los primeros corresponden al esterno-tiroideo, tiro-hioideo, esterno-hioideo y omo-hioideo. Los músculos suprahioideos, por otra parte, son el milo-hioideo, geni-hioideo, digástrico, estilo-hioideo (Casado *et al.*, 2002) e hipogloso (Romero *et al.*, 2007).

La musculatura intrínseca de la laringe actúa sinérgicamente en la producción de la voz, la respiración, la deglución, y también ayuda al mecanismo muscular general del organismo, por medio del reflejo de cierre por esfuerzo glótico (Romero *et al.*, 2007). Está conformada por los músculos cricotiroideo, cricoaritenideo posterior, cricoaritenideo lateral, interartenoideo, tiroaritenoides superiores también llamados falsas cuerdas vocales o bandas ventriculares, y por último los músculos tiroaritenoides inferiores que constituyen las cuerdas vocales verdaderas o pliegues vocales (Romero *et al.*, 2007, Casado *et al.*, 2002).

Las cuerdas vocales están compuestas por varias capas de tejidos, las cuales difieren tanto histológicamente como en sus propiedades mecánicas y funcionales. Desde el punto de vista mecánico, en la cuerda vocal “se supone la existencia de áreas de distintas propiedades físicas, que van de la elasticidad superficial a la rigidez en profundidad, cuyo comportamiento mecánico es diferente y variable y regulado tanto activa como pasivamente por la acción muscular” (García-Tapia *et al.*, 1996).

Desde el punto de vista funcional, se ha propuesto la teoría de “cubierta-cuerpo” la cual explica la vibración de las cuerdas vocales al relacionarla con las propiedades mecánicas de sus capas y la interacción entre ellas. Esta teoría distingue 3 capas diferentes según su valor funcional. La primera es la “cubierta”, que está constituida por epitelio escamoso estratificado, y las capas superficial e intermedia de la lámina propia, que corresponden a fibras laxas y elásticas respectivamente. La segunda es la de “transición”, de rigidez intermedia, está conformada por las capas intermedia y profunda de la lámina propia, que están hechas de fibras elásticas y

colágenas respectivamente, llamadas en conjunto ligamento vocal. La tercera capa es el “cuerpo”, la cual es mecánicamente más rígida, y está compuesta por las fibras del músculo estriado tiroaritenideo inferior, llamado músculo vocal (Morrison *et al.*, 1996; Casado *et al.*, 2002; Romero *et al.*, 2007).

La vibración de la cuerda vocal genera una frecuencia fundamental (F0) o tono vocal (Prater y Swift, 1986, citado por Romero y Villanueva, 2007), además de una serie de armónicos que corresponden a múltiplos de dicha frecuencia fundamental. La frecuencia e intensidad de la vibración producida puede ser modificada por la masa, la longitud y la tensión de la cuerda vocal (Bernal *et al.*, 2000). Sin embargo se advierte que “la fonación producida a nivel de la glotis constituiría un sonido de baja intensidad, sin estructuras de resonancia adyacentes que lo amplificaran”. (Romero y Villanueva, 2007, p. 45)

2.3 Cavidades supraglóticas

Son las cavidades que se encuentran sobre las cuerdas vocales, también llamadas en su conjunto tracto vocal supraglótico (Casado *et al.*, 2002). Corresponden a estructuras anatómicas que derivan del tercer, cuarto y quinto arco branquial y constituyen el aparato resonancial; su función consiste en amplificar la señal vibratoria emitida por las cuerdas vocales a través de la resonancia (Romero *et al.*, 2007).

Existe consenso entre diversos autores en que son principalmente 3 las cavidades: faríngea, oral y nasal. Sin embargo existen otros que dan cuenta de una cuarta cavidad, que es la labial (Bernal, 2000). Además, se incluyen elementos articulatorios existentes en las cavidades faríngea y oral: labios, dientes, lengua, mejillas, paladar duro, alvéolos y paladar blando (Farías, 2007). Especial mención requiere el paladar blando (velo del paladar), gracias al cual se determina la resonancia a nivel de la cavidad nasal que carece de elementos móviles y cuya función es pasiva en la producción del habla (Bernal, 2000).

2.4 Resonancia y fonética acústica

Se advierte que “la resonancia es el fenómeno acústico gracias al cual una estructura vibratoria (fuente de sonido) excita el aire contenido en una cámara, aire que a su vez hace que las paredes de ésta vibren de modo similar”. Es decir, al producirse la vibración de las cuerdas vocales, las partículas de aire circundantes son excitadas, y su movimiento es transmitido a las paredes del sistema resonador, las cuales también vibran. Además, las vibraciones también se pueden transmitir por contacto directo de la fuente con la cámara (Romero *et al.*, 2007).

Los resonadores poseen una gama de frecuencias de vibración que les son propias de acuerdo con sus características físicas, lo que se denomina frecuencia natural (Villayandre, 2009). Estas características están determinadas por la forma, tensión y viscosidad de las paredes, que son variables (Romero *et al.*, 2007). En consecuencia, cada modificación de las cavidades supraglóticas implica un cambio de sus frecuencias naturales de vibración (Villayandre, 2009).

Para que se produzca la resonancia en el tracto vocal es preciso que la frecuencia natural de vibración de la cavidad resonadora se asemeje a la de la fuente de sonido. Es por esto que la onda sonora glotal es sometida a una serie de alteraciones, pues el resonador de acuerdo a su morfología, enfatiza o amplifica las frecuencias que coinciden con la frecuencia natural del cuerpo resonador, y por otra parte, desestima o filtra aquellas que no coinciden (Villayandre, 2009).

El proceso de amplificación y filtración de la frecuencia fundamental y sus armónicos produce la amplificación de varios grupos o bandas de frecuencias las cuales se denominan formantes, que cubren un rango de frecuencias que va desde los 100 Hz hasta los 3500 Hz. Esto se debe a que las resonancias de alta frecuencia son atenuadas por el tracto vocal, el cual tiende a actuar como un filtro pasa-bajo con una caída de aproximadamente 12 dB por octava (Bernal *et al.*, 2000).

Las formantes se relacionan con un resonador en especial, que de acuerdo a sus cualidades vibratorias favorece cierto rango de frecuencias. De esta manera, la primera formante o F1 tiene directa relación con la faringe; la segunda formante o F2 se asocia con la cavidad oral; y F3 con la conformación de los labios (Bernal *et al.*,

2000). Otros autores afirman que el tercer formante, además entrega información sobre la longitud del tracto vocal (Marrero *et al*, 2003).

Desde un punto de vista fonético acústico, las vocales son los sonidos emitidos por la sola vibración de las cuerdas vocales sin ningún obstáculo entre la laringe y las aberturas oral y nasal. Son siempre sonidos de carácter sonoro y por consiguiente de espectro bien definido (Bernal *et al.*, 2000).

Las consonantes, por el contrario, se emiten interponiendo alguna constricción formada por los elementos articulatorios. Los sonidos correspondientes a las consonantes pueden ser sonoros o no, y dependen de la vibración de los pliegues vocales (Bernal *et al.*, 2000).

Para el análisis espectrográfico, es esencial el estudio de las formantes de las vocales puesto que entregan mayor, y más clara información a nivel de frecuencia e intensidad, lo que se debe a la ausencia de obstrucciones en el tracto vocal. Esto es apreciable visualmente en el espectrograma, lo que hace más fácil su observación y análisis (Bernal *et al.*, 2000).

Los formantes 1 y 2 de las vocales son los que permiten su mejor caracterización y se relacionan con las dimensiones del movimiento lingual. El primer formante se relaciona con la dimensión alto-bajo y el segundo con la anterior-posterior (Ruiz y Soto-Barba, 2005). En cuanto a sus valores, fluctúan entre los siguientes rangos:

Vocal	Hombres		Mujeres	
	F1	F2	F1	F2
I	265	2318	241	2835
E	454	1995	492	2252
A	657	1215	664	1168
O	475	888	511	981
U	294	669	243	629

Tabla 1. Extraída de Román 2010

La energía acústica queda de manifiesto en el siguiente espectrograma, en las zonas más sombreadas.

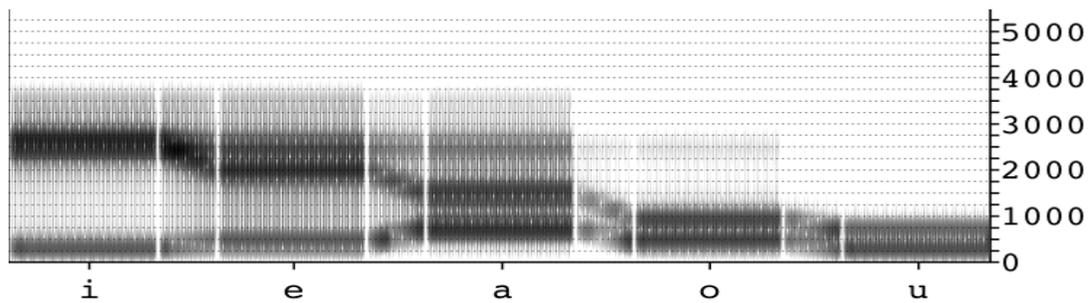


Figura 1. Espectrograma de las 5 vocales españolas sintetizadas a partir de los datos de Ruiz y Soto-Barba (2005). Las franjas negras más cercanas al eje de las abscisas corresponden al primer formante (F1) de cada una de las vocales. Por otra parte, las franjas negras que están sobre las anteriores, corresponden a al segundo formante (F2) En algunas vocales es distinguible además un tercer formante (F3).

2.5. Elementos Segmentales y Suprasegmentales

El análisis fonético del habla se basa principalmente en el estudio de los elementos segmentales del sistema fonético y fonológico de una lengua (Llisterri, 2004). A modo de glosario, en el español de Chile podemos encontrar los siguientes elementos segmentales:

- **Sonido:** es la unidad de habla, producto de las vibraciones acústicas que se propagan por el espacio como resultado de una operación articulatoria producida en los órganos fonatorios, y se caracteriza por su materialización de los fonemas de forma individual y particular. (Benito, 2005)
- **Fonema:** Es la unidad de percepción mental de un sonido, caracterizada por la cualidad contrastiva que es capaz de producir cambios de significados. Es objeto de estudio de la fonología que lo clasificó en dos grandes unidades: fonemas vocálicos y fonemas consonánticos. (Frías, 2001; Hualde, 2001)

- **Alófonos (o fonos):** Son aquellos sonidos concretos que constituyen variantes de un único fonema. Son objeto de estudio de la fonética (Frías, 2001; Hualde, 2001)
- **Grafía:** Representación escrita de un fonema en la lectura, no existe una correspondencia biunívoca de cada elemento fónico. (Llisterri, 2003; Benito, 2005)
- **Alfabeto fonético:** conjunto de representaciones escritas (grafía de cada fonema) de los fonemas donde existe una correspondencia biunívoca entre un sonido y un símbolo, llamados también ortográfico (Llisterri, 2003).

Los fonemas consonánticos son clasificados según tres parámetros principales:

1. Punto articulatorio: Es el lugar específico de la cavidad oral donde se articula un determinado fonema. En la articulación del sonido, el paso de aire de los pulmones se ve total o parcialmente impedido por los órganos articulatorios activos y pasivos. (Frías, 2001; Hualde, 2001)

2. Modo articulatorio: Es la posición que adoptan los órganos articulatorios en cuanto al grado y regularidad de apertura en el momento del paso del aire, en el inicio o durante la ejecución del fonema (Frías, 2001; Hualde, 2001).

3. Sonoridad: Es el producto de la actividad vibratoria de las cuerdas vocales ante el paso del aire, durante la ejecución de los fonemas sonoros. En los fonemas áfonos, los pliegues vocales no ofrecen resistencia al paso del aire (Hualde, 2001).

Los fonemas vocálicos se clasifican de acuerdo a 3 parámetros: altura del dorso de la lengua (vocales altas, media y baja); desplazamiento lingual en la cavidad bucal (vocales anteriores, central y posteriores); y posición de los labios (vocales redondeadas y no redondeadas). (Hualde, 2001)

La voz como fenómeno complejo no sólo incluye los elementos segmentales, sino que para un correcto análisis acústico debe incluir otras variables que pueden

modificar patrones en el espectrograma, como por ejemplo la frecuencia fundamental. Estos fenómenos fonético-fonológicos que actúan simultáneamente sobre más de un segmento (la sílaba, la palabra, el grupo fónico, etc.) reciben el nombre de elementos suprasegmentales (Iribar *n.d.*). Llisterri los organiza en:

- 1) Melodía: Variación de la altura tonal a lo largo de un enunciado (Llisterri, 2003)
- 2) Acento: Sensación perceptiva que pone de relieve una sílaba sobre el resto de las sílabas de la palabra (Gil, 2007).
- 3) Tono: Es el grado de tensión y distensión (vibración) que provoca el flujo en los repertorios de patrones vocálicos (Sabido, 2002).
- 4) Ritmo: Sensación perceptiva provocada por la sucesión de determinados elementos en periodos regulares de tiempo (Gil, 2007).
- 5) Entonación: Sensación perceptiva que produce, fundamentalmente, las variaciones de tono a lo largo de un enunciado (Gil, 2007).
- 6) Pausas: Silencio o vocalización intercalados en el discurso (Gil, 2007).
- 7) Velocidad de elocución: Número de elementos fónicos (sonidos y pausas) que se pronuncian en una unidad de tiempo determinada (Gil, 2007).
- 8) Calidad de la voz: Características globales debidas a los ajustes laríngeos y de las cavidades supraglóticas de un hablante (Llisterri, 2010).

Estos elementos se relacionan en el análisis acústico con: La frecuencia fundamental (melodía, tono, acento), la configuración espectral (accento, calidad de la voz), la amplitud (accento, ritmo) y el tiempo (accento, pausas, velocidad de elocución y ritmo) (Llisterri, 2010).

El estudio de los elementos suprasegmentales constituye la prosodia (Llisterri, Iribar).

La prosodia tiene por objeto un dominio muy amplio que comprende el estudio de diversos fenómenos asociados al acento, al ritmo y a la entonación, así como a sus manifestaciones físicas producto de las variaciones de la duración, de la frecuencia fundamental y de la intensidad (Universidad de los Andes Merida Venezuela, Laboratorio de Fonética).

Para el análisis acústico es de suma importancia el integrar la información prosódica ya que las tres áreas básicas de las tecnologías del habla -síntesis, reconocimiento y los sistemas de diálogo- requieren estos datos para alcanzar sus objetivos. La mayor importancia de la prosodia en el análisis acústico para la identificación de un sujeto se basa en la variación que provoca en la F0 fundamentalmente (Zañartu, 2003).

2.6. Identificación de individuos, variables/constantes interlocutor e intralocutor

Es relevante el estudio de los patrones fonéticos que tengan la menor variabilidad intralocutor y la mayor variabilidad interlocutor, porque con ello podemos obtener canones objetivos para identificar locutores con un alto grado de fiabilidad (Rose, 2003, Wolf en Marrero *et al*, 2007). En cuanto a esto, hay muchos patrones en estudio dentro de los cuales algunos importantes son:

- Parámetros referidos a la fuente y a los resonadores (Stevens, 1971, en Marrero *et al.*, 2007).
- El espectro a largo plazo, el vector de los formantes vocálicos, el vector temporal y el vector de F0 (Hollien, 1990 y 1991; en Marrero *et al*, 2007)

Los que son aprobados mayoritariamente (Marrero *et al.*, 2007) son el estudio de los formantes medios F1, F2 y F3 (este último proporciona información sobre la longitud del tracto vocal), cuyas características ofrecen mayor estabilidad en las vocales extremas /a/, /i/, /u/, y la dificultad de no poder alterar las distancias entre formantes a voluntad (Wolf, 1972 en Marrero *et al.*, 2007). Por otra parte, en un estudio preliminar de Marrero y Battaner realizado en el año 2003, centrado en la variación interlocutor, se establecen como objeto de estudio los siguientes parámetros:

- Pitch (F0): Frecuencia Fundamental, valores medios
- LTAS (Long term average spectre o Espectro a largo plazo): Intensidad Media (dB) y Espectro medio (Hz) .
- Segmentos vocálicos: Parámetros: distancia F0 / F1 / F2 / F3 / F0-F1, distancia /F2-F3, y distancia / F0-F3.
- Fonemas nasales: Parámetros: F1 / F2 / F3 en coarticulación con vocales

- VOT: Duración del inicio, parecido al ataque vocal.
- Fonemas fricativos de alta resonancia: Parámetros: Turbulencia al comienzo y en el punto máximo.

Se pudieron obtener algunos resultados interesantes, como que las variaciones en el comienzo de la [s] (variación fonética o alófono del fonema fricativo /s/), pueden ser de más de 1000 Hz dentro del mismo contexto o que las diferencias entre los triángulos vocálicos de dos sujetos son de más de 400 Hz en F2 y 100 Hz en el F1. Las distancias entre los formantes también mostraron un importante rango de variación, especialmente entre F2-F3, aunque manteniendo idéntico el tipo de estímulo (Marrero *et al.*, 2003).

Otras diferencias significativas que se han obtenido en estudios de esta línea, son: 1) Los valores medios de F0 presentan variaciones significativas en función del estilo de habla; 2) La duración de los grupos fónicos y su número de segmentos, predicen de manera significativa la desviación típica del F0; 3) Las resonancias en zonas de altas frecuencias y la F0 permiten diferenciar en todos los segmentos vocálicos un mayor número de sujetos que los valores de F1 y F2; 4) La duración del grupo fónico presenta variaciones significativas en función del estilo de habla (Marrero *et al.*, 2007).

2.7. Corpus oral

El “corpus” siempre ha sido una herramienta útil en la fonoaudiología ya sea en el área de lenguaje infantil (evaluación cualitativa en niños con TEL), en el área de voz (grabación para análisis acústico con software) o en el habla de pacientes neurológicos (evaluación del lenguaje de pacientes afásicos o análisis acústico de disártricos),

En este tipo de investigación se hace necesario la creación de un corpus para el desarrollo y la evaluación de un sistema de reconocimiento del habla, principalmente para la identificación de individuos (Llisterri, 2005). Así, investigaciones en España han utilizado un conocido corpus denominado Ahumada-Gaudí (Ortega, 2000) como muestra para el análisis fonético acústico en estudios como el proyecto

VILE, dedicado al análisis acústico de la variación inter e intra locutor en español (Marrero *et al.*, 2003).

El corpus Ahumada fue diseñado para incluir muchas de las fuentes de la variabilidad entre hablantes, permitiendo centrarse en ellos y estudiar sus efectos subyacentes en los sistemas de verificación de hablantes (Ortega, 2000). En ese sentido a los hablantes se les solicitaron las siguientes tareas de elicitación:

- a) 24 números aislados, descartando los primeros y últimos 2 de ellos debido a consideraciones prosódicas. Los 20 números restantes consistían en dos repeticiones de los dígitos del 0 al 9 aisladamente.
- b) 10 cadenas de dígitos que constan de 10 números cada una, siendo las primeras 5 cadenas de dígitos idénticos para todos los hablantes a través de todas las sesiones de grabación, y las últimas 5 cadenas son específicas para cada locutor y sesión.
- c) 10 declaraciones fonológicamente balanceadas de 8-12 palabras de largo. Estas declaraciones fueron idénticas para todos los hablantes a través de todas las sesiones
- d) 1 texto silábica y fonológicamente balanceado (de 182 palabras y más de 1 minuto de duración) leído a una velocidad normal de lectura (160 palabras por minuto aproximadamente). este texto fue modificado (para evitar condicionamiento) para todos los hablantes a través de todas las sesiones.
- e) 2 repeticiones del texto modificado que se usó en el ítem anterior, pidiéndole al hablante que lo lea con velocidad rápida y lenta (esta tarea solo se pidió en las sesiones 1.3 y 5 donde las grabaciones in situ fueron satisfactorias.
- f) 1 texto específico, diferente de locutor en locutor y de sesión en sesión. Este texto será elegido al azar de novelas o periódicos, y al menos 1 minuto de este tipo de habla es requerido.
- g) Más de 1 minuto de habla espontánea pidiéndole a cada hablante que describa (sin pausas largas ni dudas) cualquier cosa que quieran. Hay disponibles algunas pinturas y fotografías, y temas como “describir sus últimas vacaciones”, “describir el lugar donde vive/nació”, etc. que han sido sugeridos también.

La primera duda que se plantea al leer estos apartados corresponde a la definición de “fonológicamente equilibrado”. Esto tiene directa relación a que el corpus debe tener una frecuencia de aparición de alófonos o de fonemas que refleja la norma de una lengua (Llisterri, 2005). Una investigación de la frecuencia de fonemas en el español de Chile (Perez, 2003) entrega una posible solución para generar frases fonológicamente equilibradas para la norma chilena, es decir, el porcentaje de cada fonema en nuestras oraciones debe ser proporcional a la frecuencia de la norma de Chile que se presenta en la siguiente tabla:

Tabla 1. Distribución porcentual de la frecuencia de los fonemas del español

/ i /	7.46	/ f /	0.75
/ e /	14.13	/ s /	9.61
/ a /	12.31	/ x /	0.74
/ o /	9.28	/ ch /	0.32
/ u /	3.05	/ y /	0.69
/ p /	2.58	/ m /	2.62
/ t /	4.92	/ n /	7.78
/ k /	3.94	/ ñ /	0.24
/ b /	1.92	/ r /	6.19
/ d /	4.84	/ rr /	0.64
/ g /	0.94	/ l /	5.05

Tabla 2. Extraída Pérez 2003

2.6. Generalidades de la física acústica aplicada a la voz

Características acústicas de los sonidos

El sonido es una onda mecánica longitudinal proveniente de una fuente de vibración mecánica que se propaga a través de un medio elástico. La vibración es periódica cuando presenta una repetición continua a intervalos determinados, por ejemplo, al golpear un diapasón generará una vibración periódica conformando una forma sinusoidal. (Jackson-Menaldi, 1992)

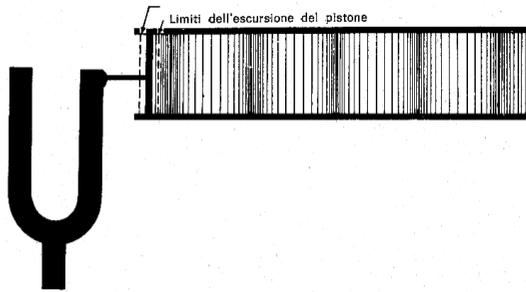
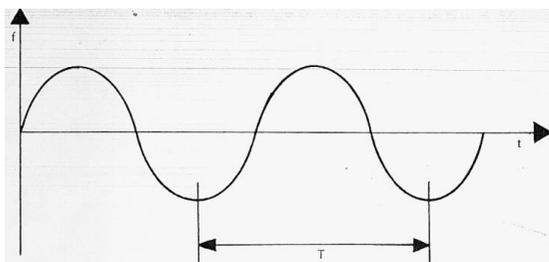


Figura 2. Onda sonora sinusoidal registrada de un diapasón vibrante. (Llisterri, 2010)

En el campo de la labor fonoaudiológica, son de interés las frecuencias (Hz) percibidas por el ser humano como sonido, las cuales fluctúan entre 20 y 20.000 Hertz (Bernal *et al.*, 2000).

Los movimientos recurrentes (que son cíclicos o frecuenciales), pueden dividirse en simples, complejos, periódicos y no periódicos. De tal forma, se entiende al movimiento recurrente simple o movimiento armónico simple como movimiento continuo que constituye la base de todos los sonidos de habla y es percibido como tono puro. Éste, a su vez, está determinado por el parámetro A que es la amplitud de la onda, el parámetro T o período, que es el tiempo en segundos en que una partícula demora en dar una vuelta completa, y el parámetro F, que es la frecuencia o el números de vueltas en un segundo expresado en Hz. Los ejemplos clásicos serían las ondas producidas por el diapasón y el audiómetro. (Jackson-Menaldi, 1992; Llisterri, 2010)



Frecuencia(f) , amplitud(A) y duración(t)

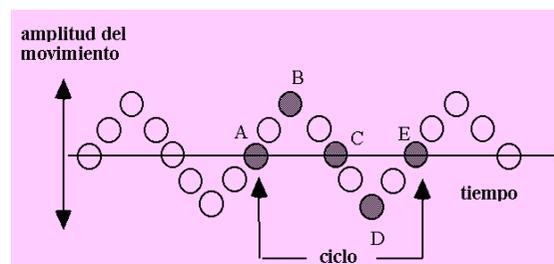


Figura 3 y 4. Movimiento armónico simple de la partícula. (Llisterri, 2010)

En cuanto al sonido del habla, como suma de movimientos armónicos simples de distinta frecuencia o movimiento armónico complejo, comprendido gracias a la visualización de los coeficientes de las funciones sinusoidales de los armónicos simples (transformada de Fourier), es resultado de una serie de movimientos vibratorios al ritmo de abertura y cierre de las cuerdas vocales en el momento del paso de partículas de aire que entran en vibración a nivel del tracto vocal, o simplemente, la glotis. Estas ondas sonoras complejas periódicas son producto de la vibración glotal y se pueden definir en función de la amplitud (que corresponde a la intensidad del sonido), la frecuencia fundamental y/o los formantes, y la duración del tiempo en el cual se lleva a cabo el movimiento. (Jackson-Menaldi, 1992)

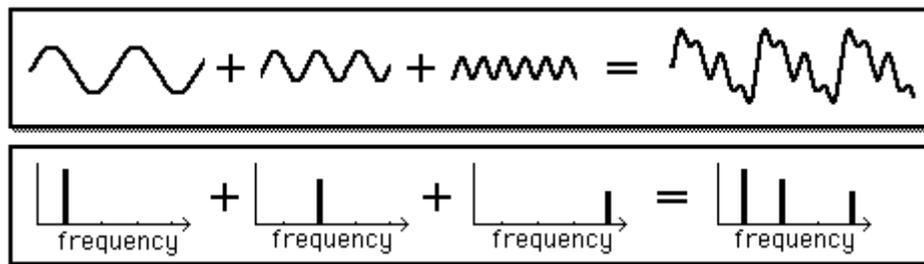


Figura 5. Onda sonora compuesta como suma de ondas sonoras simples. (Listerri, 2010)

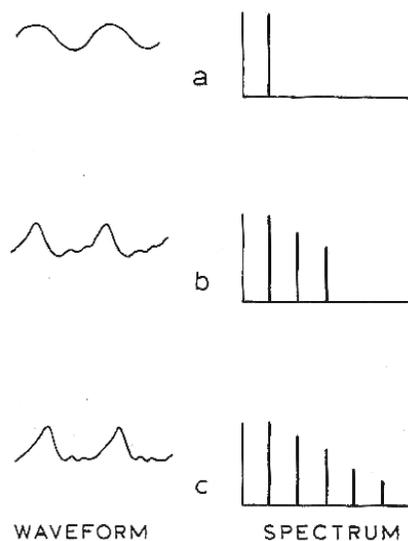


Figura 6. Esquema de comparación entre “waveform” y “spectrum” de diferentes frecuencias donde se observa: a) Tono puro, b) y c) ondas compuestas. (Listerri, 2010)

La caracterización de la onda sonora compleja del sonido del habla se atribuye a la frecuencia fundamental (F0), que corresponde al ciclo por segundo de abertura y cierre de las cuerdas vocales; los armónicos, que consisten en los componentes restantes de la onda sonora compleja y que poseen una frecuencia en múltiplo entero de la frecuencia fundamental; la resonancia, que es la modificación de la amplitud de los armónicos realizada por las cavidades supraglóticas; los formantes, que son frecuencias amplificadas por la resonancia; y la amplitud del sonido o la intensidad en dB.(Llisterri, 2010)

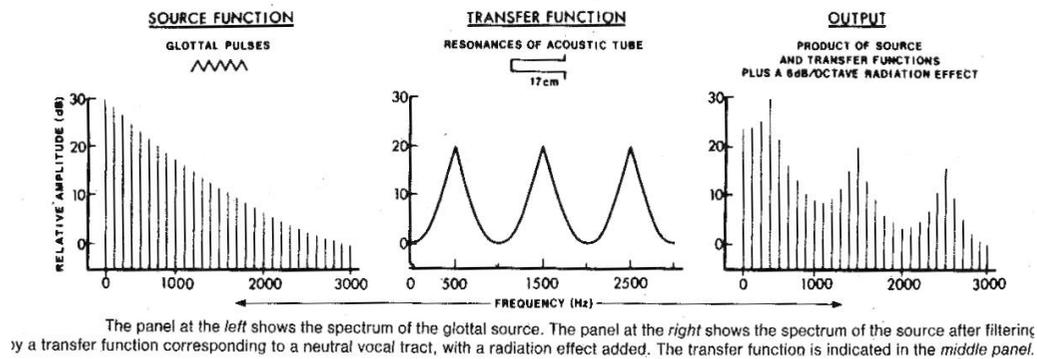


Figura 7. Se observa el espectro de ondas de la fuente glotal (a la izquierda), la que es filtrada por el tracto vocal (centro), lo que da una onda resultante o "output" (derecha).

Para facilitar el estudio acústico del habla, se han configurado varias representaciones gráficas con el fin de mostrar las características o parámetros de la voz hablada. Así como el oscilograma en el que se registra y representa gráficamente la evolución (fluctuación) de la amplitud a lo largo del tiempo. (Llisterri, 2010)



Figura 8. Oscilograma. (Llisterri, 2010)

Y el espectrograma que es una representación gráfica de la variaciones de la frecuencia y la intensidad mediante los niveles de colores o grises del sonido en que se muestran a lo largo del tiempo.(Llisterri,2010)

2.7. Análisis de la voz en Praat

Praat es un programa desarrollado por Boersma and Weenink, cuya última versión es del 2008, y que permite hacer análisis acústicos de la voz del tipo espectral mediante el procesamiento informático de la señal de la voz (Bernal *et al.*, 2000). La conversión digital se realiza mediante algoritmos matemáticos procesados en un ordenador, el cual entrega valores en Hz, oscilogramas y espectrogramas (Román, 2008).

Los análisis en Praat ofrecen opciones muy variadas, dentro de estas es útil la obtención de características fonéticas, cuyo análisis sirve para la identificación de sujetos.

Al trabajar con este software, una vez cargada o generada una muestra. Se abrirá un abanico de posibilidades en las cuales nos ofrece señalarnos los formantes y la línea melódica, configurar el ancho de banda, ajustar la resolución frecuencial y temporal, y hacer zoom entre otras cosas atinentes para un estudio acústico (Román, 2008).

2.8. Coarticulación

Se entiende por coarticulación a los cambios en la articulación de un segmento de habla dependiendo de lo que le precede (coarticulación retrógrada), y de los segmentos siguientes (coarticulación anterógrada) (Cohen & Massaro, 1993). La coarticulación es uno de los factores que hace más compleja la segmentación de la onda sonora en unidades lingüísticas y su posterior análisis (Llisterri, n.d.) por lo que resulta un desafío a la hora de establecer qué segmento del espectrograma se debe analizar.

Rose indica que la coarticulación es un factor individual, por lo que cada persona tendría su propia estrategia para coarticular un segmento y que incluso puede ser estudiado como un elemento que produce diferencia inter-locutores (Rose, 2002). Por lo mismo, es iluso pensar que se pueden dar valores arbitrarios a la cantidad de tiempo (o ciclos) que se vayan a analizar en el espectrograma porque un sujeto puede tener un alto grado de coarticulación y otro mínimo (Rose, 2002).

3. Problema

El problema es que se desconoce si es posible establecer en Chile la existencia de constantes y variaciones fonético / acústicas descritas en la literatura extranjera, que permitan diferenciar a un individuo del resto de una población, es decir, encontrar parámetros que nos permitan reconocer la muestra de voz de un individuo dentro de un grupo de muestras. El por qué de este problema nace en la fonética forense actual que se lleva a cabo en la sección sonido de la PDI de Chile, que considera muchas variables, pero no toma en cuenta que estas nacieron bajo el contexto de otros registros de habla, u otras variantes del español distintas a la nuestra. Entonces es de importancia comenzar con los argumentos teóricos para nuestra norma, aunque sea solo con un aproximación cuyo valor se centra en la ratificación mediante datos empíricos.

4. Objetivo general

Identificar las variaciones/constantes fonético acústicas interlocutor en un grupo de 20 sujetos hablantes nativos chilenos pertenecientes a la Región Metropolitana de Santiago de Chile, entre los meses de junio y noviembre del año 2010.

5. Objetivos específicos

Todos los objetivos están planteados en modalidad lectura y están clasificados para hombres y mujeres por separado.

1. Identificar las variaciones/constantes fonético-acústicos interlocutor de los formantes medios y F0 de los alófonos de los fonemas /a/, /i/, /u/ tónicos, y /m/ y /n/ trabantes.
2. Identificar las variaciones/constantes fonético-acústicos interlocutor de los formantes medios y F0 de los alófonos de los fonemas /a/, /i/, /u/ átonos, y /m/ y /n/ en articulación con vocal.
3. Caracterizar las variaciones/constantes fonético-acústicos interlocutor de los formantes medios y F0 de los alófonos de los fonemas /a/, /i/, /u/ tónicos, y /m/ y /n/ trabantes.
4. Caracterizar las variaciones/constantes fonético-acústicos interlocutor de los formantes medios y F0 de los alófonos de los fonemas /a/, /i/, /u/ átonos, y /m/ y /n/ en articulación con vocal.

6. Variables

6.1. Frecuencia fundamental (F0)

Definición nominal: es el tono fundamental producido por una fuente de sonido, se mide en Hertz.

Definición conceptual: La frecuencia fundamental de la voz es el producto del movimiento de las cuerdas vocales debido al efecto Bernoulli y a la estructura fisiológica de la cuerda explicada por la teoría de cubierta-cuerpo propuesta por Hirano.

Operacional: En el contexto del estudio, es el tono más grave de la voz producido por un sujeto y puede apreciarse en el espectrograma como energía acumulada en la parte más cercana al eje X.

Es una variable simple numérica continua independiente.

6.2. Formantes medias (F1, F2, F3)

Definición nominal: Cada uno de los rasgos identificables de un sonido o de un fonema.

Definición conceptual: Corresponden al resultado del filtrado de los armónicos de la F0 producidos por la cuerda vocal. Las estructuras encargadas de dicho filtrado son las cavidades de resonancia: cavidad oral, nasal, faríngea y los labios.

Operacional: En el contexto del estudio, corresponden a estructuras de frecuencias que se encuentran en un valor sobre la frecuencia fundamental y que se observan en el espectrograma como zonas ennegrecidas en frecuencias más altas que la F0.

Son una variable simple numérica continua independiente.

6.3. Fonemas /a/, /i/, /u/ tónicos

Definición nominal: Corresponde a una variación de los fonemas que tiene acentuación en el texto que se utilizó para las muestras.

Definición conceptual: Son variaciones alofónicas de un fonema que se caracterizan por tener acento y se reflejan en una mayor energía e intensidad en el espectrograma.

Operacional: En el estudio corresponden a una variable que crea una división o categoría de análisis diferente, por posibles discrepancias a establecerse con los fonemas átonos.

Es una categoría de fonemas que por su acumulación de energía debiese presentar características más estables o diferenciables que los fonemas átonos.

6.4. Fonemas /a/, /i/, /u/ átonos

Definición nominal: Corresponde a una variación de los fonemas que no tiene acentuación en el texto que se utilizó para las muestras

Definición conceptual: Son variaciones alofónicas de un fonema que se encuentran fuera de una sílaba tónica o en un contexto en el cual el acento no las resalta en la palabra o sintagma.

Operacional: En el estudio son una variable que crea una categoría que se distingue de los fonemas tónicos puesto que en el análisis espectral, estos fonemas deberían tener menos intensidad y energía.

6.5. Fonemas /m/, /n/ trabantes

Definición nominal: Corresponde a una variación alofónica de los fonemas /m/ y /n/ que genera una división o categoría de análisis diferente a la articulada con vocal.

Definición conceptual: Son la variante alofónica de los fonemas nasales que se ubican en posición final o coda silábica.

Operacional: Dentro del Estudio se distinguen por sus rasgos marcados en el análisis espectral y proporcionan una categoría de análisis distinta de los fonemas /m/ y /n/ articulados con vocales.

6.6. Fonemas /m/, /n/ en articulación con vocal

Definición nominal: Es una variación alofónica de los fonemas nasales /m/ y /n/ que genera una categoría de análisis aparte de la variación en trabante.

Definición conceptual: Son una variante de los fonemas /m/ y /n/ que se suceden de una vocal, como en las sílabas directas por ejemplo.

Operacional: En el estudio proporcionan una categoría de alófonos de los fonemas /m/ y /n/ cuya unidad de análisis lleva contigua una vocal que se puede utilizar como clave visual en el espectrograma.

7. Metodología

7.1. Diseño de la investigación:

La investigación es un estudio no experimental descriptivo (estudio de campo) de correlación transversal a realizarse entre marzo y noviembre del año 2010.

7.2. Población y grupo en estudio:

El grupo en estudio constó de 20 hablantes, 8 hombres y 12 mujeres para homogeneizar la muestra en una proporción similar a la del universo que corresponde a la gente de la ciudad de Santiago, donde encontramos una población estimada al 30 de Junio del 2009 según el censo del 2002 correspondiente a 3.324.216 Hombres y 3.490.414 Mujeres (Instituto Nacional de Estadísticas, 2009), lo que nos da una relación de 1:1.05 respectivamente. El criterio de inclusión, fué ser hablante nativo del español de Chile, y no padecer alguna patología vocal orgánico u organicofuncional, ya que las estructuras formánticas son un parámetro alterado en las disfonías, que puede mejorar con tratamiento y que causaría variación en los resultados en distintos momentos evaluativos. (Shaheen & Nelson, 2009); y por factibilidad (heterogeneidad de la muestra y compromiso) ser estudiante de la Universidad de Chile. Para ser seleccionados se hizo un llamado colectivo a participar en la investigación vía e-mail. La edad considerada de los sujetos fué entre 19 y 23 años.

7.3. Instrumentos de recolección de datos

Para la recolección de datos (muestras de habla) se utilizó un instrumento con la estructura del apartado “d)” de los corpus que se usaron en la base de datos Ahumada-Gaudí, es decir una versión extendida a 182 palabras del texto “El Abuelo” (Ver anexo 1) que es usado habitualmente en la evaluación de pacientes con trastornos de habla. La modificación de este corpus fue revisada por el profesional Lingüista Christian Peñaloza que es docente de la escuela de Fonoaudiología de la Universidad de Chile. La finalidad de este corpus fue el de permitir el registro de

muestras de habla equilibradas en parámetros fonético acústicos. El registro de la toma de muestras se llevó a cabo con un computador, el software Praat, y un micrófono dentro de la cámara silente del laboratorio de fonética de la escuela de Fonoaudiología de la Universidad de Chile. La revisión de los corpus obtenidos se hizo con el software de análisis espectral Praat.

7.4. Procedimientos

Se seleccionaron 42 sujetos que respondieron al llamado colectivo hecho en la Universidad de Chile, luego mediante un consentimiento informado (ver anexo 2) se les explicó el rol que cumplirían en la investigación y se les pidió autorización para grabar sus muestras .

Las muestras de voz fueron digitalizadas y posteriormente analizadas mediante el software Praat, lo que es explicado en detalle en el punto 7.5.

Del corpus “El Abuelo” descrito en el apartado 7.4, se desprende la siguiente tabla que indica el número de alófonos o variaciones de fonemas que fueron analizados por cada muestra (en total se analizaron 20 muestras).

	1º párrafo	2º párrafo	3º párrafo	4º párrafo	Total
a tónicas	10	10	5	9	34
a átonas	21	17	9	14	61
a Total	31	27	14	23	95
i tónicas	3	3	1	10	17
i átonas	13	10	7	11	41
i Total	16	13	8	21	58
u tónicas	3	2	1	1	7
u átonas	6	13	3	8	30
u Total	9	15	4	9	37
m oclusiva trabante m en articulación con vocal	2 8	2 2	3 0	1 3	8 13
m totales	10	4	3	4	21
n oclusiva o trabante n en articulación con vocal	9 10	9 2	7 2	8 3	33 17
n totales	19	11	9	11	50
Total alófonos	85	70	38	68	261

En esta investigación se utilizó la información espectrográfica de los alófonos que no estuvieran coarticulados, independiente del número de ciclos que se encontraran y de la longitud temporal de dicho alófono, con la salvedad de que perceptivo-acústicamente este fuese distinguible como un sonido del habla del español de Chile y no como un ruido.

Los datos de los alófonos analizados fueron vaciados en planillas de Excel cuyo diseño distribuía a cada fonema en una página diferente. A su vez en cada página, existía la división entre “átona y tónica” para las variaciones de los fonemas vocálicos /a/, /i/, /u/ y la división entre “trabante y en articulación con una vocal” para los fonemas nasales /m/ y /n/.

Para el posterior análisis fueron considerados:

- Los promedios y la desviación estándar de los formantes F0, F1, F2, F3 de los alófonos de los fonemas vocálicos tónicos y átonos de cada uno de los individuos.
- Los promedios y la desviación estándar de los segmentos o intervalos entre los formantes F0-F1, F1-F2, F2-F3 y F0-F3 de los fonemas vocálicos tónicos y átonos de cada uno de los individuos.
- Los promedios y la desviación estándar de los formantes F0, F1, F2, F3 de los alófonos de los fonemas nasales trabantes y en articulación con vocal de cada uno de los individuos.
- Los promedios y la desviación estándar de los segmentos o intervalos entre los formantes F0-F1, F1-F2, F2-F3 y F0-F3 de los fonemas nasales trabantes y en articulación con vocal de cada uno de los individuos.

Estos datos obtenidos fueron vaciados a tablas generales para la integración de los datos de todos los individuos, la creación de gráficos y la descripción del comportamiento de las variables en estudio.

7.5. Pasos a seguir en la toma y análisis de muestras de voz/habla.

En este apartado se describe en una pauta práctica el paso a paso para tomar y analizar una muestra. Esta pauta fue seguida en cada una de las muestras que se obtuvieron.

Los pasos a seguir son:

- Pedir al sujeto que va a participar en el estudio que lea, apruebe y firme el consentimiento informado.
- Chequear del correcto funcionamiento del instrumental y del cableado.
- Colocar el micrófono en posición perpendicular a la altura de la boca y a una distancia de 5 a 7 centímetros para evitar saturación.
- Dar las instrucciones: que lea a velocidad homologable a su habla conversacional y que continúe la lectura a pesar de los errores.
- Abrir el programa Praat, En la ventana Praat Objects apretar New, Record Stereo Sound, Record.
- Pedir al sujeto que lea el texto, “El Abuelo (extendido)”
- Una vez terminada la lectura, presionar Stop, en la parte superior de la ventana ir a File, marcar write to wav y después seleccionar la carpeta donde se guardará el archivo, luego colocar nombre en la ventanilla Name (nombre y apellido, para propósitos administrativos) y después presionar Save to list & Close.
- En el caso de que la muestra este guardada en el ordenador desde antes, en la ventana Praat objects, ir a la parte superior de la venta a y presionar en Read, luego en read from file. Al seleccionar Read from file se abre una ventana que permite navegar por el disco duro del ordenador y seleccionar el fichero que contiene la señal que se desea analizar.

- Una vez seleccionado el fichero que contiene la señal sonora que se desea analizar, se elige, en el menú de la derecha, la opción Edit.
- Con ello se abre una ventana de análisis y de edición en la que se muestra el oscilograma (forma de onda) en la parte superior y el espectrograma en la parte inferior. En esta ventana hay que presionar sobre el menú superior de la pantalla en Spectrum y poner un ticket en show spectrogram, ya que ver el espectrograma será de utilidad después a la hora de distinguir formantes y seleccionar tramos no coarticulados de los alófonos, luego presionar en Spectrum settings, Window length y asignar el valor de 0,03. esto ayuda al análisis perceptivo visual que se tiene del espectrograma y que también ayuda a la selección de los alófonos.
- Posteriormente, se debe identificar los fonemas y sus formantes. Para esto hay que presionar dentro de la ventana, en la tercera barra que dice "Total duration", lo que hace que se reproduzca la cadena de audio.
- Se puede estimar el lapso temporal en el que irán apareciendo los fonemas, marcando el espacio temporal con el cursor del mouse, haciendo clic derecho sobre la ventana en los puntos donde se estime conveniente al observar el espectrograma.
- Para tener una mayor resolución frecuencial, dentro de la misma ventana (cuyo nombre es el que se etiquetó en pasos anteriores) hay que presionar sobre el menú de la esquina izquierda inferior en el botón "sel".
- Para hacer más zoom, se puede volver a seleccionar con el cursor del mouse el lapso temporal e ir presionando el botón "sel", y si se quiere volver atrás, presionar el botón "bak" (que se encuentra al lado del botón sel) y en caso de que no se pueda volver más atrás, presionar el botón "out" (también está junto a los otros dos botones anteriores).
- Luego se distingue el o los fonemas que se quieren analizar viendo la regularidad de las ondas y la energía del espectrograma, a la vez que se

escucha (presionando en “visible part”) la parte seleccionada para cerciorarse de que se está trabajando con el fonema (alófono) correcto sin partículas de otros alófonos adyacentes (coarticulación).

- Por último, para obtener los formantes, dentro de la selección se presiona los botones del teclado: F1 para formante 1, F2 para formante 2, F3 para formante 3 y F5 para el Formante 0 (o frecuencia fundamental o “pitch”).
- Los datos de los formantes (en Hertz) obtenidos en el paso anterior se guardarán en tablas Excel para el posterior análisis de resultados.

8. Resultados

Para los resultados se graficaron los datos obtenidos del análisis acústico de las muestras. Se crearon gráficos separados por género y otro para la totalidad de la muestra, mostrando los promedios de las formantes de cada vocal tónica y atona y nasales aproximantes y trabantes.

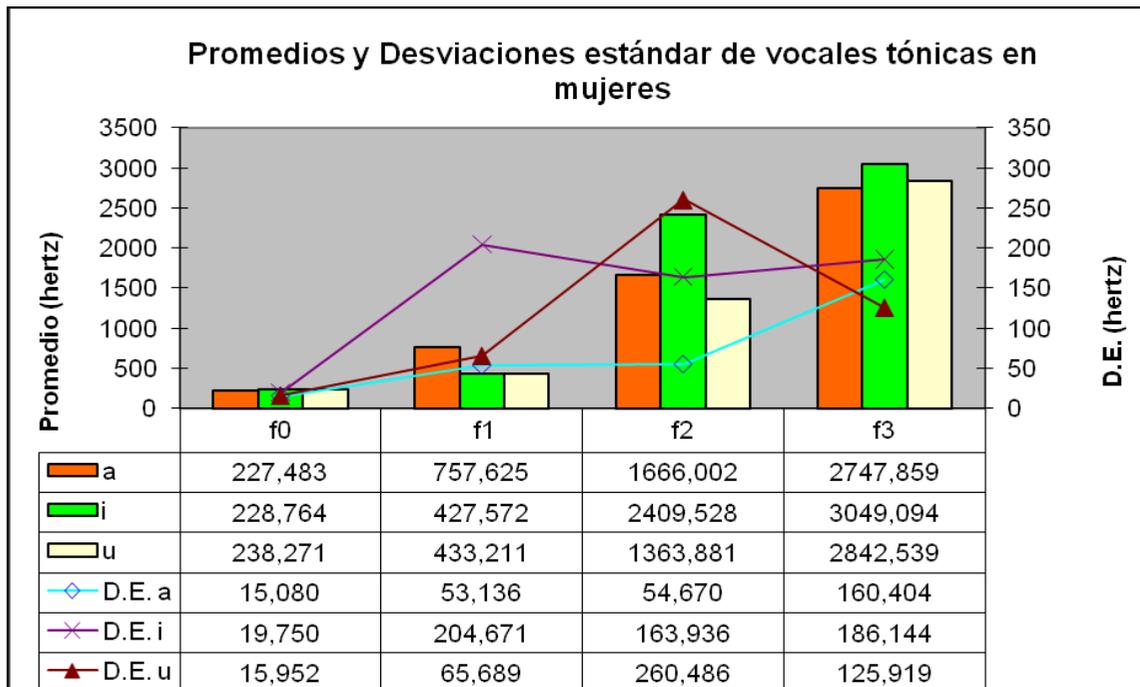


Gráfico 1. Promedio y desviaciones estándar de las frecuencias F0, F1, F2 y F3 vocales tónicas en mujeres.

Se observa un promedio mayor de la vocal /i/ tónica en el F2 y F3, mientras en el F1 se encuentra un promedio mayor de la vocal /a/tónica. A su vez, se observa que en los promedios de la F0 hay pocas variaciones entre vocales tónicas.

En relación a las desviaciones estándar en F0 se presenta la menor variación entre vocales tónicas. En F1 y F2 se altera esta similitud entre desviaciones estándar de las vocales tónicas en donde en F1 aumenta la D.E. de la /i/ tónica y en F2 aumenta la de /u/, éste último formante presenta una dispersión de datos mayor entre

las 3 vocales, mientras en las del F3 se observan más concentradas, menos que en F0.

En la /a/ tónica la D.E se mantiene en un valor inferior a las otras vocales.

En el gráfico 2 se muestran los datos obtenidos de las vocales átonas. En este gráfico no se evidencian mayores diferencias con el anterior, sin embargo, se puede observar que las desviaciones estándar del F1 están menos dispersas que la del gráfico 1 mientras que en las del F3 se presentan más dispersas. A su vez, las D.E de la vocal /u/ átona están más distanciadas con las otras presentando mayores incrementos de valores de D.E en el F2 y F3, mientras que en las D.E. de la vocal /a/ e /i/ no se observan mayores aumentos de las mismas mostrando un relieve más plano. F0 se muestra estable y con pocas variaciones entre vocales tanto en promedios de los formantes como en las desviaciones estándar al igual que como se presenta en las vocales tónicas.

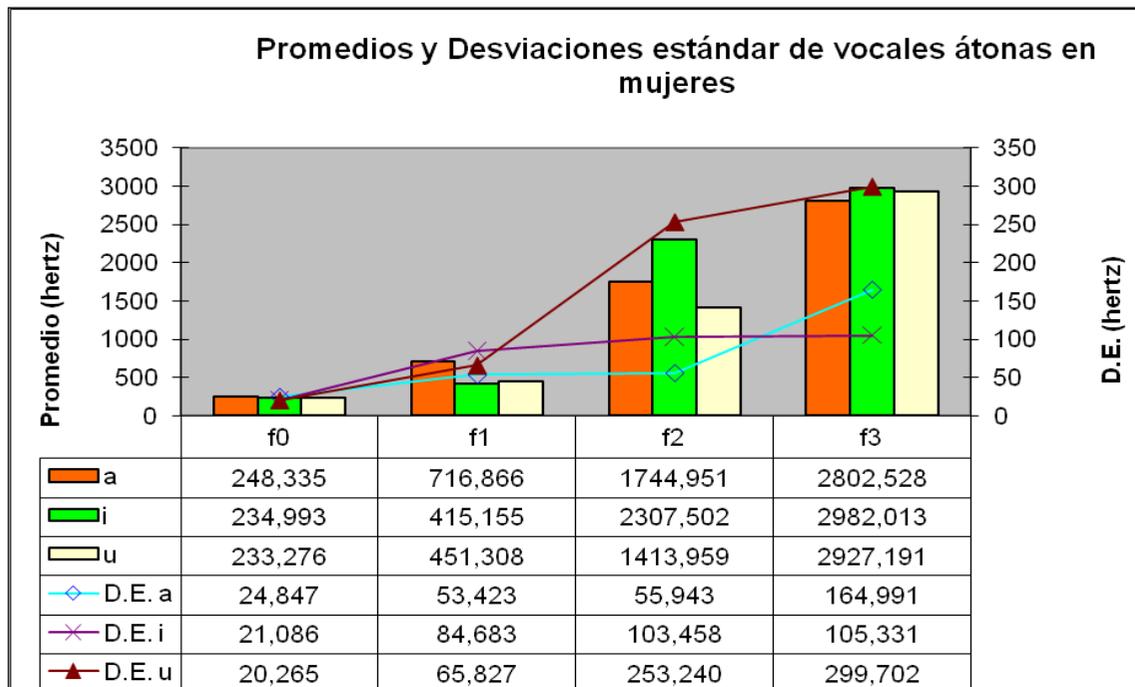


Gráfico 2. Promedios y desviaciones estándar de las frecuencias F0, F1, F2 y F3 de vocales átonas en mujeres.

Tanto vocales tónicas como atonas presentan valores similares en promedio de los formantes siendo posible diferenciarlos a través de los cambios en las desviaciones estándar, específicamente en F1 y F3.

También se graficaron los datos de consonantes nasales /m/ y /n/, los cuales se pueden observar en los gráficos 3 y 4.

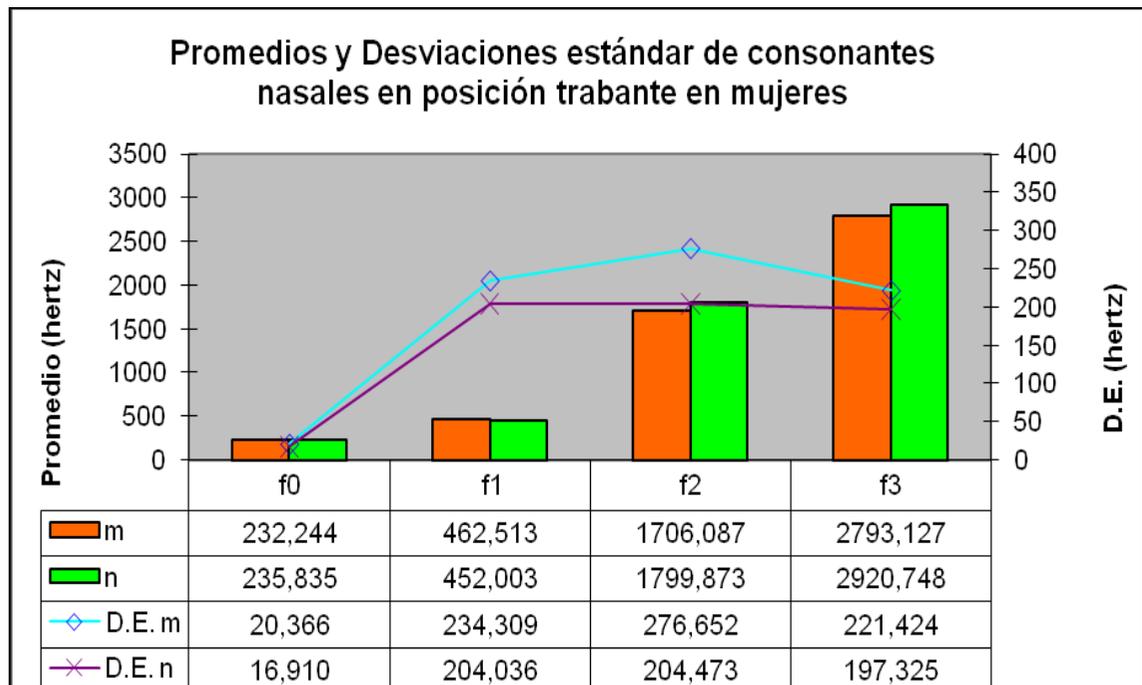


Gráfico 3. Promedios y desviaciones estándar de las frecuencias F0, F1, F2 y F3 de consonantes nasales en posición trabante en mujeres.

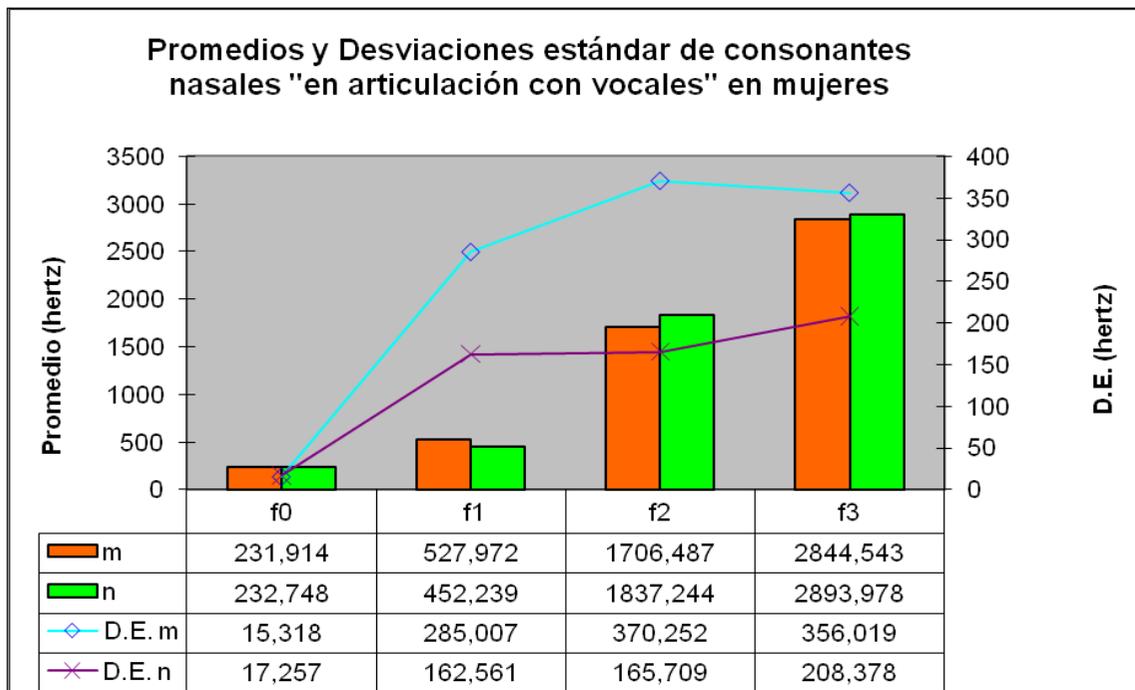
Los promedios de las formantes de las consonantes nasales tanto en posición trabante como en articulación con vocales en mujeres se aumentan en proporciones similares, siendo valores mayores en la F2 y F3.

A su vez, dentro de las desviaciones estándar de las consonantes nasales en posición trabante se observan unas F0 más estables, lo que ocurre similar en las nasales en articulación con vocales; mientras que en F1, F2 y F3 se observa una meseta en los valores de desviación estándar de /n/, teniendo /m/ más variaciones de desviación estándar entre formantes.

Las consonantes nasales en articulación con vocales se diferencian de las trabantes principalmente en los valores de las desviaciones estándar de F1, F2 y F3,

como con /n/ en articulación con vocales que a diferencia de la posición trabante no presenta meseta en estos valores.

En /m/ de articulación con vocales existe una mayor desviación estándar en F1,



F2 y F3 respecto a su posición trabante.

Gráfico 4. Promedios y desviaciones estándar de las frecuencias F0, F1, F2 y F3 de consonantes nasales en articulación con vocales en mujeres.

En relación a los promedios y desviaciones estándar de vocales tónicas en hombres, se observa un valor promedio aumentado del alófono /i/ en F2 en comparación con F1 al igual que los de las mujeres, además los promedios de F0 son similares entre ellos. Se observa en la /u/ una F3 más aguda tanto en tónica como en átona que las otras dos vocales a diferencia de las mujeres donde se presenta el F3 más agudo en alófono /i/.

Dentro de las desviaciones estándar, los valores del F0 están relativamente similares entre vocales, y pocas variaciones en la F1 y F3, sin embargo se observa una mayor variabilidad en la F2, especialmente en la D.E. de /u/ tónica.

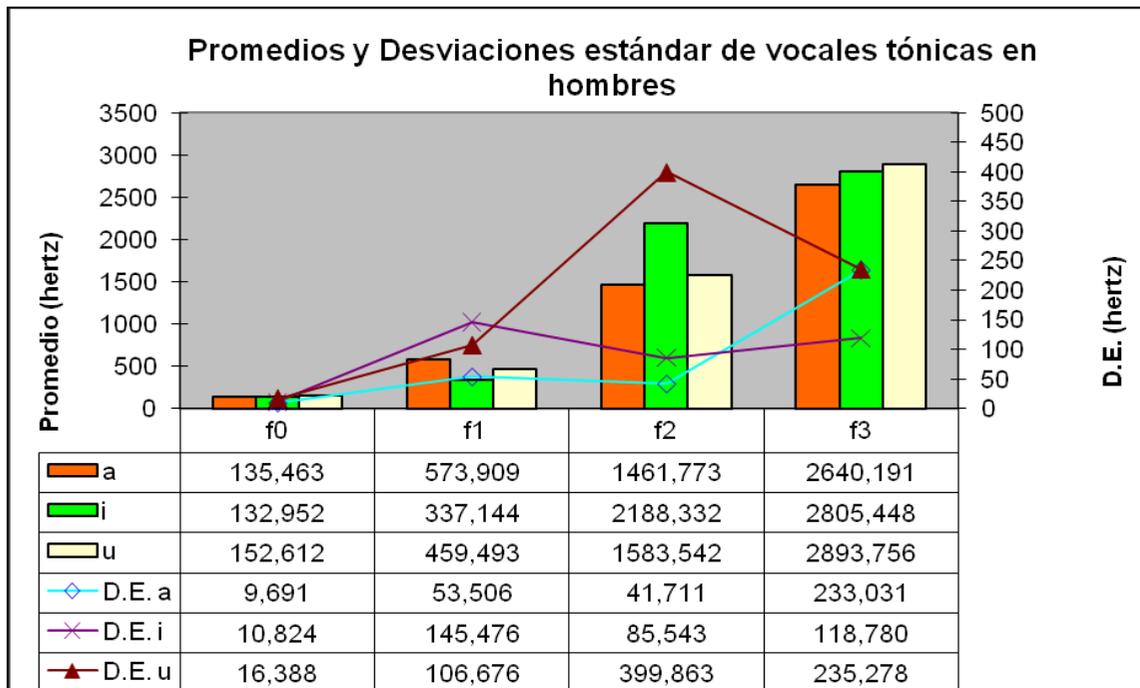


Gráfico 5. Promedios y desviaciones estándar de las frecuencias F0, F1, F2 y F3 de vocales tónicas en hombres.

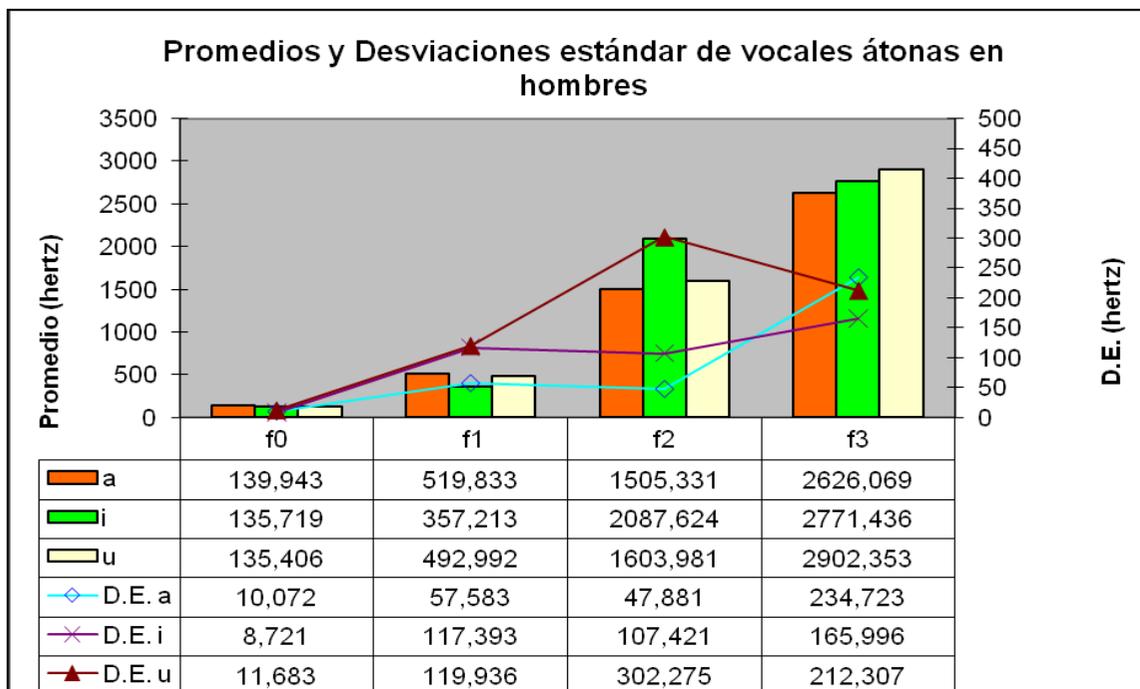


Gráfico 6. Promedios y desviaciones estándar de las frecuencias F0, F1, F2 y F3 de vocales átonas en hombres.

Las vocales átonas en hombres tienen una F0 y F1 con valores de promedio similares a las tónicas; además, en particular, se observa un descenso de D.E. en F3 de la vocal /u/ átona en F3.

En general, las vocales tanto átonas como tónicas se disponen de manera similar en ambos géneros, variando el valor promedio de las formantes hacia agudas o graves según sea mujer u hombre respectivamente.

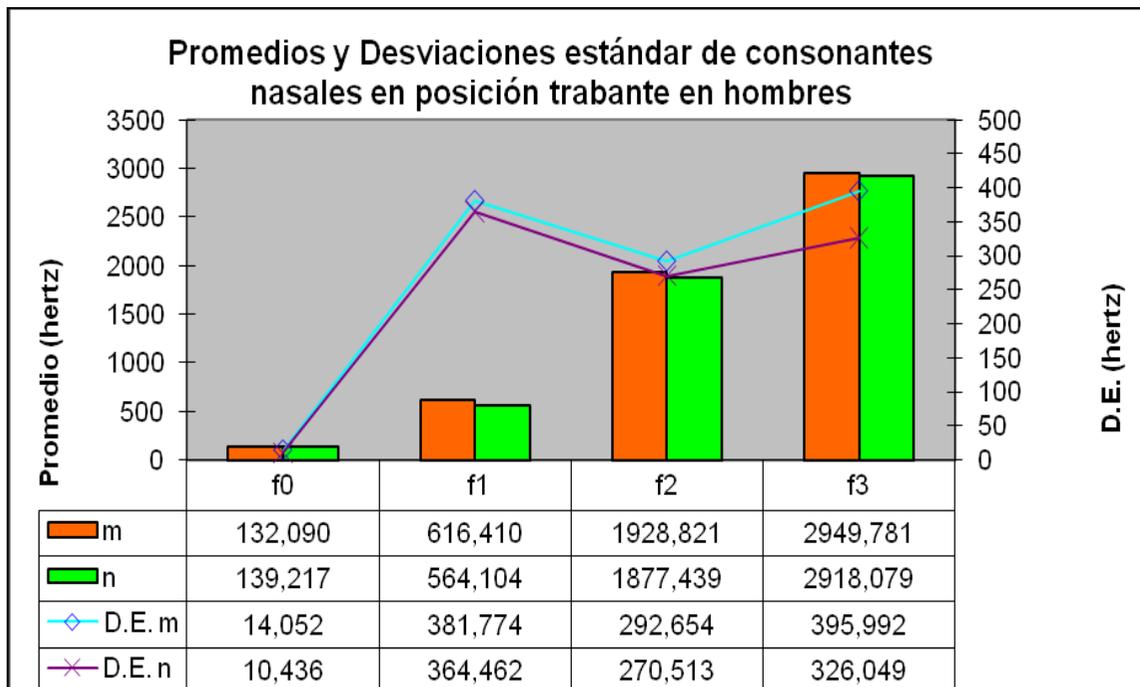


Gráfico 7. Promedios y desviaciones estándar de las frecuencias de F0, F1, F2 y F3 de consonantes nasales en posición trabante en hombres.

En el gráfico 7 se observa un mayor incremento de los promedios de los formantes de las nasales tanto en posición trabante como en articulación con vocales en F2 y F3. Cabe destacar que los promedios de la /m/ en articulación con vocales son mayores que la /n/ en articulación con vocales.

Dentro de las desviaciones estándar se observa un valor muy bajo en la F0, y un aumento abrupto en la F1 en las nasales en posición trabante y en articulación con vocales. A su vez se observa un leve descenso de las D.E. en la F2 en ambos casos.

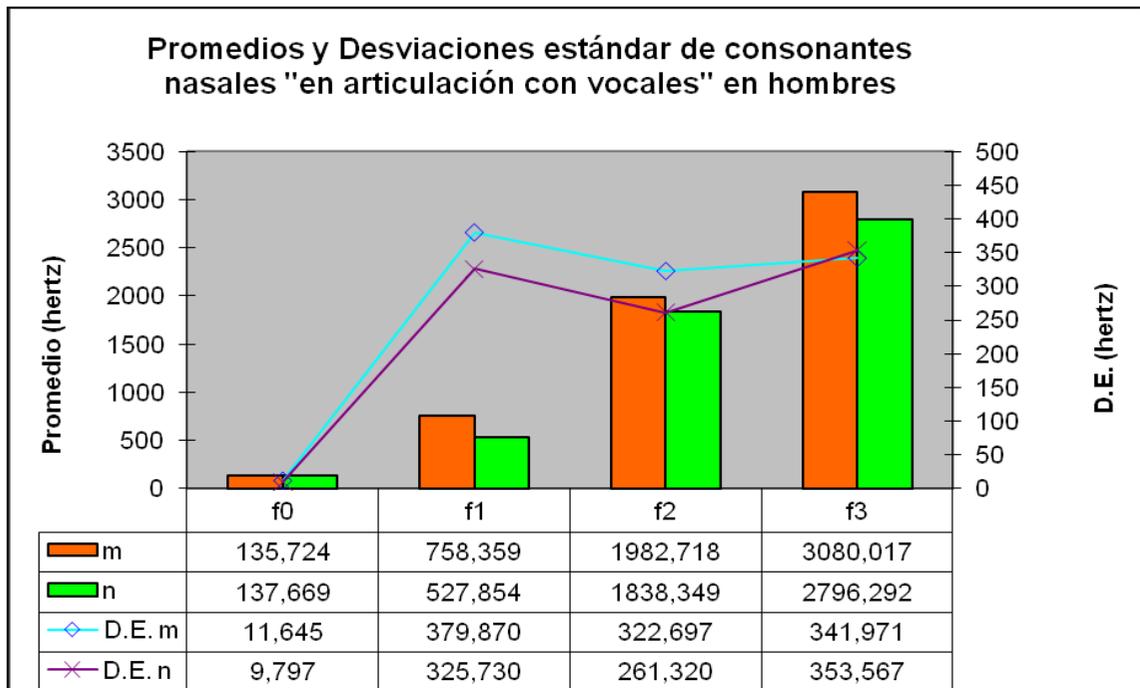


Gráfico 8. Promedios y desviaciones estándar de las frecuencias F0, F1, F2 y F3 de consonantes nasales en articulación con vocales en hombres.

El valor que presenta la menor desviación estándar es la F0 en la vocal atona /i/ con 8,721hz y a su vez el dato con mayor desviación estándar es la F2 en la vocal tónica /u/ con 399,863hz.

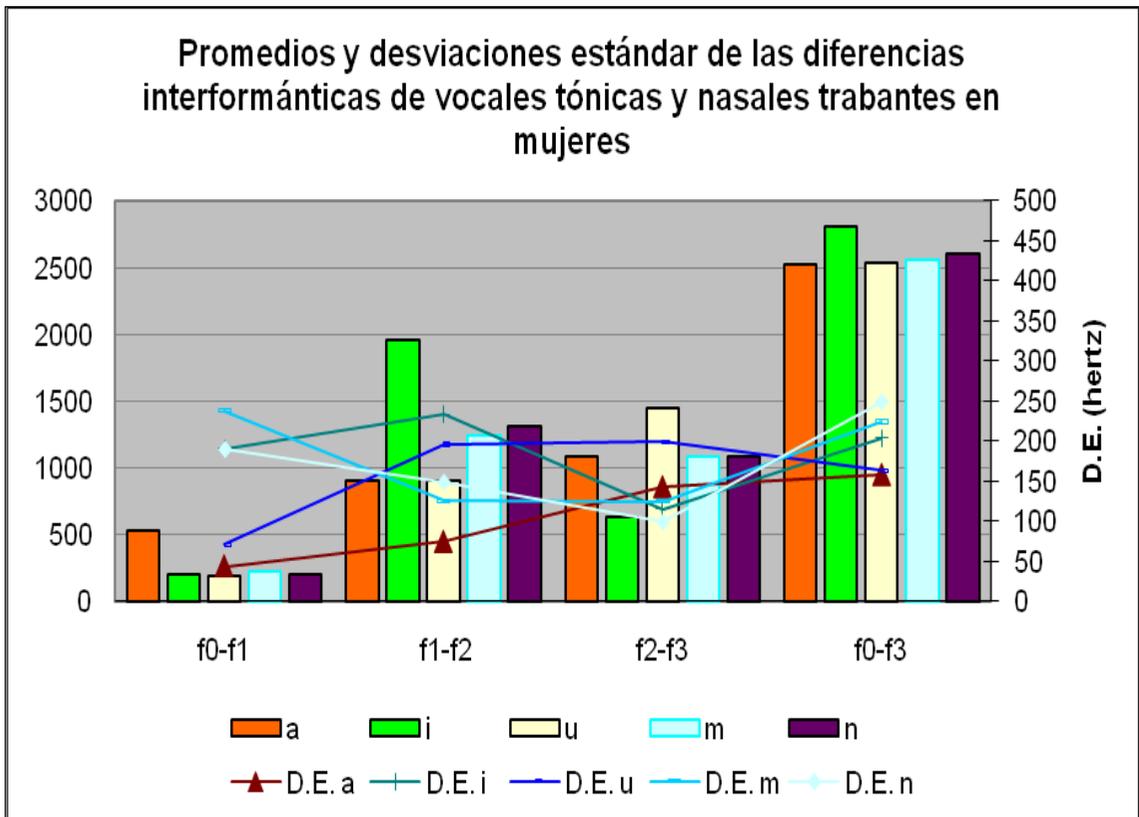


Gráfico 9. Promedios y desviaciones estándar de las diferencias interformánticas de vocales tónicas y nasales trabantes en mujeres.

En la vocal /a/ tónica, el rango de F0-F1 se evidencia la mayor diferencia entre los promedios de los formantes, mientras que en los otros fonemas los valores son similares. En cuanto a la diferencia F1-F2 y F0-F3 se destaca una mayor distancia entre formantes, predominante de la vocal tónica /i/ y además en la diferencia entre F0 y F3 existe una similitud entre los valores de los demás fonemas, y una mayor variabilidad entre la diferencia F2-F3 en la vocal tónica /u/.

Los valores de desviaciones estándar se encuentran en un rango entre 249,005 y 44,194 Hz, siendo más baja la D.E. de la vocal /a/ tónica en la diferencia F0-F1 y más alta las D.E. de la consonante nasal /n/ trabante en la diferencia F0-F3.

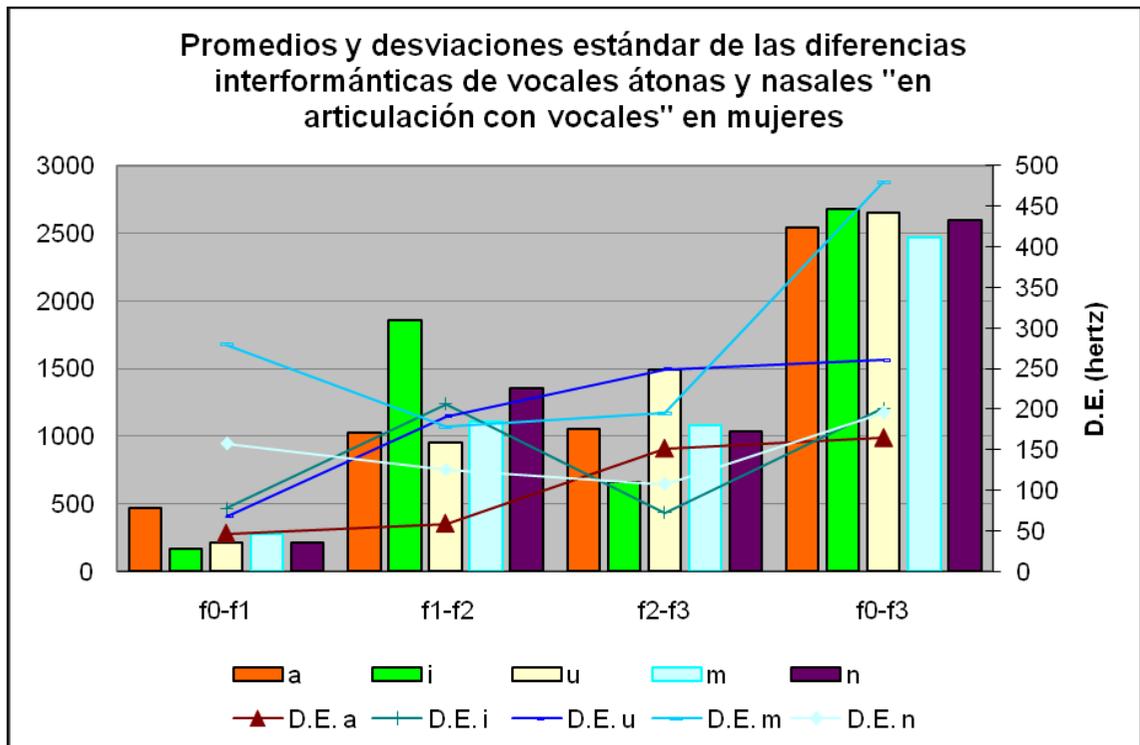


Gráfico 10. Promedios y desviaciones estándar de las diferencias interformánticas de vocales átonas y nasales en articulación con vocales en mujeres.

Los valores de promedio y sus distribuciones en cada rango interformántico se encuentran de manera similar a los valores en la tabla 9, presentando diferencias en el rango de F0-F3, en donde la vocal atona /u/ se aproxima a los valores de /i/ en el mismo rango. Las desviaciones estándar de /m/ e /i/ en la tabla 10 presentan la mayor diferencia respecto a la tabla anterior especialmente para el rango de F0-F3 en la consonante /m/ y el rango de F0-F1 de la vocal /i/.

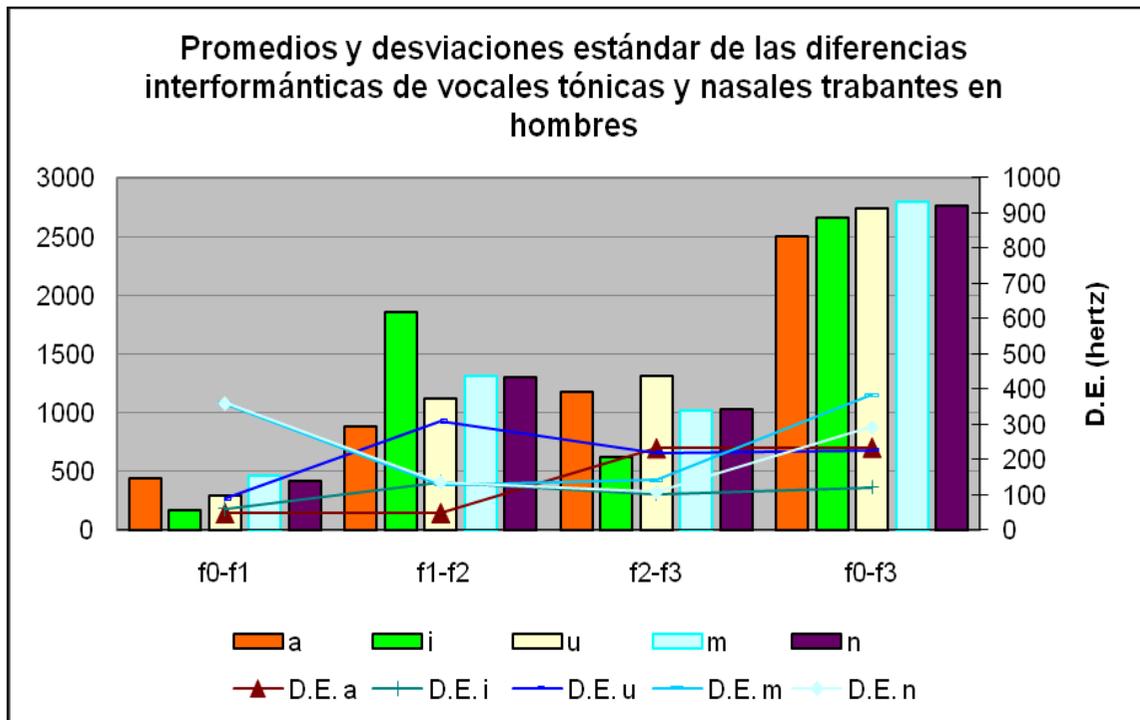


Gráfico 11. Promedios y desviaciones estándar de las diferencias interformánticas de vocales tónicas y nasales trabantes en hombres.

La vocal tónica /i/ en sujetos hombres muestra las diferencias más significativas respecto a los otros fonemas analizados, esto se muestra en el rango de F0-F1 como un valor menor al promedio de los fonemas analizados, en el rango F1-F2 el valor está sobre la media teniendo un valor mayor respecto a los otros fonemas, y en el rango de F2-F3 sucede un fenómeno parecido al rango F0-F1 donde el valor es menor al promedio de los fonemas analizados.

Respecto a la desviación estándar, los datos que se escapan del promedio son D.E. de /m/ y /n/ en el rango de F0-F1.

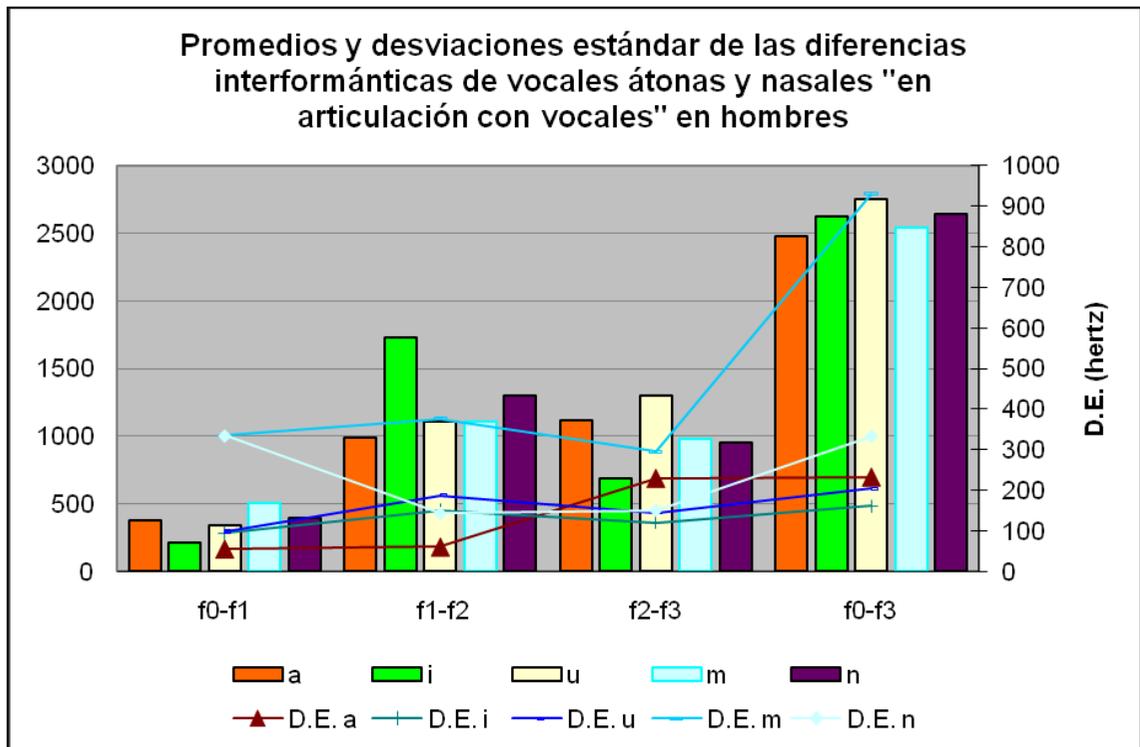


Gráfico 12. Promedios y desviaciones estándar de las diferencias interformánticas de vocales átonas y nasales en articulación con vocales en hombres.

En el gráfico 12 se pueden observar los valores de las diferencias interformánticas en las vocales átonas y consonantes nasales en articulación con vocales, sobresaliendo las variaciones que presentan los valores de /n/ en el rango de F1-F2 donde aumentan respecto a los valores del gráfico 11, y los valores de desviación estándar se mantienen en valores similares a excepción de la D.E. de /m/, que se escapa del promedio presentando un aumento en todos los rangos de diferencias interformánticas estando por sobre las demás desviaciones estándar a excepción de la D.E. de /n/ en el rango de F0-F1 donde están en valores similares.

9. Conclusiones y Discusiones

Se describe en la literatura que existen diferencias entre un alófono vocálico átono y un alófono vocálico tónico, considerándose que este último tiene una mayor acumulación de energía, por lo que su duración es mucho mayor en la palabra, es más audible y sus formantes son mucho más visibles en el espectrograma, es decir, la zona donde se ubica el formante debería apreciarse más contundente como una zona mucho más oscura que incluso podría facilitar al investigador para individualizarla en la palabra. Sin embargo a la luz de este trabajo se establece que esto no es una verdad absoluta, ya que si bien, los alófonos tónicos son más perceptibles acústica y visualmente (en el espectrograma), en algunos fonemas como la /a/ no se aprecian grandes diferencias entre las categorías tónicas y átonas, ni en su promedio, ni en las desviaciones estándar por lo que se entiende que ambas se distribuyen de manera muy similar. A futuro sería de utilidad establecer qué parámetros son realmente distintivos entre una variante alofónica tónica y átona de un mismo fonema en el español hablado de Chile.

Dentro del presente seminario se presentó una muestra pequeña, básicamente por una dificultad a la hora de analizar los datos de cada persona, ya que eran aproximadamente 2088 datos por individuo, si eso se extrapola a los 20 sujetos analizados, nos resultan 41760 datos brutos, esto corresponde a una gran base de datos que es muy difícil de sintetizar y manipular.

Se hace sumamente necesario generar una base de datos mucho más grande que la obtenida aquí, que hace que los patrones sean más generalizables y utilizables para la identificación de individuos mediante la generación de un perfil fonético acústico; es decir, en el futuro se podrían almacenar los datos fonético acústicos de distintos sujetos indubitados para posteriormente analizarlos y compararlos con datos de algún sujeto dubitado. Esta es una aplicación empírica de la fonética forense, un ejemplo de ello es la existencia de asociaciones profesionales como la "IAFP, International Association for Forensic Phonetics", que publica, desde 1991, la revista "*The International Journal of Speech, Language and the Law*" (London: Equinox Publishing), que celebra congresos anuales y mantiene una lista electrónica de distribución.

Con respecto al análisis de los resultados, se pudieron observar ciertos datos importantes que discutiremos a continuación:

Si bien el F0 es un parámetro muy estable tanto en el análisis interlocutor como en el intralocutor, no presta grandes antecedentes para el objetivo del presente seminario, por cuanto no nos permite realizar individuación de un sujeto, es decir, no permite distinguirlo de los demás hablantes de una lengua ni nos entrega datos de valor para caracterizar la voz de un sujeto desconocido, ya que el promedio de este valor se mezcla mucho con el de los hablantes que lo rodean y que tienen una lengua común.

En los análisis de las mujeres encontramos que el parámetro más estable entre sujetos (interlocutor) es la F1 de la vocal /a/ tónica en base a la poca variación del promedio y una baja desviación estándar. Por otro lado, un dato relevante de mencionar es el F0 del fonema /a/ como un dato muy estable entre todos los sujetos de la muestra.

De los gráficos de las mujeres se desprende:

En el fonema /a/: si comparamos las formantes F0, F1, F2, F3 y las respectivas diferencias interformánticas del alófono [a] átono v/s [á] tónico dentro de una palabra, observamos que los promedios y las desviaciones estándar son muy similares, por ende, concluimos que ambos alófonos, tanto tónico como átono podrían ser analizados de manera conjunta en un sujeto sin que esto altere el resultado final.

En los análisis de los hombres encontramos que el parámetro más estable entre sujetos (interlocutor) es la F1 de la vocal /a/ tónica en base a la poca variación del promedio y una baja desviación estándar. Por otro lado, un dato relevante de mencionar es el F0 del fonema /a/ como un dato muy estable entre todos los sujetos de la muestra.

De los gráficos de los hombres se desprende:

En el fonema /a/: si comparamos las formantes F0, F1, F2, F3 y las respectivas diferencias interformánticas del alófono [a] átono v/s [á] tónico dentro de una palabra, observamos que los promedios y las desviaciones estándar son muy similares, por ende, concluimos que ambos alófonos, tanto tónico como átono podrían ser analizados de manera conjunta en un sujeto sin que esto altere el resultado final.

En un análisis general que incluye a hombres y mujeres podemos mencionar: El dato que menos varía entre sujetos (dato interlocutor) es la F0 del fonema /a/.

Para el análisis de tantos datos, se obtuvieron muestras de 42 personas, de las cuales a cada evaluador se le asignó un grupo de individuos a analizar. Por asuntos de recursos humanos, tiempo, disponibilidad, etc. cada integrante del seminario analizó a un promedio aprox. de 4 personas, sumando un total de 20 sujetos, poco menos de la mitad del total de las muestras recaudadas. Aun así, el manejo de esta información es muy complejo, teniendo incluso problemas con el planteamiento del ANOVA, por lo que finalmente se optó por un análisis meramente descriptivo de los datos, pero no por ello menos válido, ya que se logran observar y se extrapolar de los gráficos y tablas; tendencias, promedios, dispersión de los datos, etc. Qué también ayudan a establecer parámetros detallados del comportamiento del hablante promedio, de sus características individuales y muestrales.

Algunas dificultades se presentaron a la hora de analizar las muestras con el programa Praat, entre estas podemos mencionar:

- a) Algunos fonemas de las muestras no presentaban F0, pero se sabe que esto es imposible, no existe un sonido que no tenga una frecuencia fundamental, ya que es la frecuencia que presenta la fuente del sonido, en este caso, la vibración de las cuerdas vocales, entonces ¿Cómo explicamos esto? Suponemos que el programa era incapaz de graficar el dato o de establecerlo, por ende entregaba la palabra <undefined>, que daba cuenta que el dato “no estaba presente”, así, se nos planteaba el problema de registrar o no algo en las tablas Excel y finalmente se decidió dejar esa casilla en blanco
- b) Algunas personas omiten ciertos alófonos (ejemplo: /a/ en /noventitres/), lo reemplazaban por otro sonido del habla o alófono que no era el que se estaba analizando (ejemplo: algún sujeto dijo /pienola/ en lugar de /pianola/), o dicen el alófono en una ventana de tiempo tan pequeña que resulta imperceptible, por lo que estos datos también se dejaron en blanco en las tablas y se anotó un dato cualitativo que señala la razón por la cual no se pusieron (ver anexo 3).

- c) El dato era un ruido y no un sonido del habla, en algunos casos el alófono estaba en un registro frito, por ejemplo: en /Neruda/ la /a/ generalmente tenía un patrón rítmico descendiente y un registro que no parecía humano: era un ruido. No se considero a la hora del registro en la plantilla Excel.
- d) Otro problema del programa Praat era que en ciertas ocasiones daba un resultado F0 mayor a F1, lo que sabemos también es físicamente imposible, la F1 es una representación de los armónicos que son múltiplos de la F0, por ende la F1 debe ser siempre mayor a la F0, en estos casos también se dejaron en blanco las cuadrículas de la tabla Excel.
- e) Sucedió ocasionalmente que un formante específico (por ejemplo F1) de un fonema (ejemplo /a/) era muy distinto del resto de formantes (F1 del resto de las /a/ dichas por el sujeto), a esos datos se les llamó “aberrantes” y de todas maneras se consideraron como datos válidos, y aunque aumenta la desviación estándar de la muestra del sujeto, nunca afectó de manera considerable, ya que solo era parte de la dispersión, además cada vez que ocurrió esto y se dejó el dato, fue porque perceptivo-acústicamente ese fonema era indiscutiblemente el que se estaba analizando.

Por otro lado, también se puede establecer que hay muy pocas diferencias interevaluador. Al analizar a las distintas personas se encontró una baja variación en casi todos los datos, reduciendo así el factor humano en los análisis, ya que si se hubiese presentado mucha variación interevaluador, se habría notado en las tablas de datos generales o en un aumento en la desviación estándar del promedio que habría dado cuenta de datos más dispersos. Empíricamente se observa que los promedios de los parámetros son muy similares y a su vez presentan bajas desviaciones estándares hechas de los distintos sujetos del estudio que a su vez fue hecho por distintos evaluadores. En conclusión existe una alta confiabilidad inter-evaluador. Y se deja abierto a los futuros trabajos del área a utilizar el análisis subjetivo de los evaluadores, ya que si se entrena el oído para el análisis perceptivo acústico y el uso del software Praat (u otro software) como en el presente seminario, se podrá llegar a un resultado con alta confiabilidad inter-evaluador.

Desde un principio esta investigación intentó utilizar como modelo una experiencia similar realizada en España (el proyecto VILE) por un equipo multidisciplinario que tardó más de 5 años y que incluso actualmente aún está en funcionamiento, por lo que resulta bastante utópico compararse con un trabajo de tamaña magnitud, y en este sentido, podemos señalar algunas debilidades de este seminario:

- 1) Los recursos fueron extremadamente limitados, tanto recursos temporales, económicos y humanos.
- 2) La base de datos generada en España ya estaba hecha (Marrero, 2003), por lo que se sumó un trabajo más para el cual el equipo de trabajo no dio a basto.
- 3) El corpus español Ahumada-Gaudí utilizado en el proyecto VILE, es bastante exclusivo, en ese sentido no pudimos acceder a sus metodologías, análisis, etc. Por lo que debimos plantear métodos similares pero de ninguna manera comparables con la experiencia española, es así como desde un principio se nos planteó la dificultad del instrumento a utilizar. “El abuelo” es un intento de realizar una herramienta que cumpliera con los criterios utilizados en España que de cierta manera es muy limitado pero es efectivo.

9. Bibliografía

- Adrian, J.A., Casado, J.C. (2002). *La evaluación clínica de la voz. Fundamentos médicos y logopédicos*. Málaga: Aljibe.
- Benito, A. (2005). *Gramática práctica*. 9ª Edición. Madrid: EDAF, S.A.
- Battaner, E., Carbó, C., Gil, J., Llisterri, J., Machuca, M. J., Madrigal, N., Marrero, V., de la Mota, C., Riera, M., Ríos, A. (2007, 24-26 Octubre) *VILE: Estudio acústico de la variación inter e intralocutor en español*, in *Actas do 3o Congreso Internacional de Fonética Experimental*. Santiago de Compostela, España
- Bernal, J., Bobadilla, J., & Gómez, P. (2000). *Reconocimiento de voz y fonética acústica*. Mexico D.F.: Alfaomega.
- Bonastre, J., Bimbot, F., Boë, L., Campbell, J., Reynolds, D. & Magrin-Chagnolleau, I. (2003, 1-4 septiembre). *Person Authentication by Voice: A Need for Caution*. Ponencia presentada en EUROSPEECH 2003 - INTERSPEECH 2003. Proceedings of the 8th European Conference on Speech Communication and Technology. Ginebra, Suiza.
- *Caracterización acústica de las vocales del español*. Domingo Román Montes de Oca. Recuperado el 24 de Abril en http://www.domingo-roman.net/vocales_esp_caract_acustica.html.
- *Caracterización de los suprasegmentos: La prosodia*. Alexander Iribar. Recuperado 27 abril 2010 de <http://paginaspersonales.deusto.es/airibar/Fonetica/Apuntes/07.html>
- Cohen & Massaro. (1993). Modeling Coarticulation in Synthetic Visual Speech. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.49.6809&rep=rep1&type=pdf> . Recuperada el 25 de mayo del 2010.
- *El análisis fonético del habla y sus aplicaciones*. Joaquim Llisterri, Universitat Autònoma de Barcelona. Recuperado el 27 Abril 2010 en http://liceu.uab.es/~joaquim/phonetics/Valencia_04/Guion_bibliografia.pdf
- Eriksson, A. (2005, 4-8 septiembre). *Tutorial on forensic speech science. Part I: Forensic phonetics*. Ponencia presentada en Interspeech 2005 - Eurospeech. Proceedings of the 9th European Conference on Speech Communication and Technology. Lisboa, Portugal. 2005.
- Farías, P. (2007) *Ejercicios para restaurar la función vocal*. 1ª edición. Buenos

Aires: Librería AKADIA Editorial.

- Frias, X. Introducción a la fonética y fonología del español. *Ianua. Revista Philologica Romanica*. Suplemento 04. p. 3-4, 2001. Recuperado de <http://www.romaniaminor.net/ianua/sup/sup04.pdf>
- García Tapia, R., y Cobeta, I. (1996). *Diagnóstico y tratamiento de los trastornos de la voz*. Barcelona: Lebón
- Gil, J. (2007) *Fonética para profesores de español: de la teoría a la práctica*. Madrid: Arco/Libros (Manuales de formación de profesores de español 2/L).
- Hualde, J; Olarrea, A & Escobar, A. (2001). *Introducción a la lingüística hispánica*. Madrid: Cambridge.
- Instituto Nacional de Estadísticas. (2009). *1.2 estadísticas demográficas*. Compendio estadístico. Estado de Chile
- Jackson-Menaldi, M. (1992). *La voz patológica*. Buenos Aires: Ed. Panamericana.
- Llisterri, J, et al.(2005) "corpus orales para el desarrollo de las tecnologías del habla en español" . *oralia*. Analisis del discurso oral. 8. 289-325
- Llisterri, J. (n.d.). Coarticulación.
http://liceu.uab.es/~joaquim/phonetics/fon_produccio/Coarticulacion.pdf .
recuperada el 25 de mayo del 2010.
- *Los elementos suprasegmentales*. Joaquim Llisterri, Universitat Autònoma de Barcelona
http://liceu.uab.cat/~joaquim/phonetics/fon_prosod/suprasegmentales.html
- Marrero, V., Gil, J., Battaner, E. (2003, 3-9 Agosto) *Inter-Speaker Variation in Spanish. An Experimental and Acoustic Preliminary Approach*, in *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, España.
- Morrison, M., Rammage, L. (1996). *Tratamiento de los trastornos de la voz*. Editorial Masson, S.A.
- Nolan, F. (2001, 12-15 diciembre). *Speaker identification evidence: its forms, limitations and roles*. Ponencia presentada en *Proceedings of the conference "Law and Language: Prospect and Retrospect"*. Levi, Finlandia.
- Ortega, J., Díaz, C., García, R., Gonzáles, J., Lucena, J., Marrero, V. & Sánchez, J. (1998, 12– 15 de mayo).

Ahumada: a large speech corpus in spanish for speaker identification and verification. Ponencia presentada en el Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998: IEEE International Conference on. Seattle, USA.

- Perez, H. (2003) . Frecuencia de fonemas. “Fonética segmental de español público de Chile: un estudio de variación estilística de la pronunciación en los noticieros de la televisión chilena”, tesis doctoral . Universidad de Concepción
- Rabiner, L. , Schafer R. (1978) *Digital processing of speech signal*. Michigan: Editorial Prentice-Hall.
- Román, D. Estudios de fonética experimental 7(17), 2008. Recuperado de http://dialnet.unirioja.es/servlet/listaarticulos?tipo_busqueda=ANUALIDAD&revista_busqueda=8634&clave_busqueda=2008
- Rose, P. (2002). *Forensic Speaker Identification*. 1ª edición. Londres: Taylor and Francis.
- Romero, L. & Villanueva, P. (2007). *Eufonía*. Santiago de Chile: Ediciones Escuela de Fonoaudiología.
- Ruiz, M. y J. Soto-Barba. 2005. Timbre vocálico en hablantes de español como segunda lengua, *Onomázein*, XI(1), 57-65.
- Sabido, M. (2002). *El tono: andanzas teóricas, aventuras prácticas, el entretenimiento con beneficio social*. Mexico: UNAM.
- Shaheen & Nelson. (2009). Outcomes Measurement in Voice Disorders: Application of an Acoustic Index of Dysphonia Severity. *Journal of Speech, Language, and Hearing Research*, 52, 482-499.
- *Tutorial de Fonética: La prosodia*. Universidad de los Andes, Merida Venezuela. Recuperado el 27 Abril 2010 de http://ceidis.ula.ve/cursos/humanidades/fonetica/tutorial_de_linguistica/prosodia.html
- Villayandre, M. (2009). *Fonética y Fonología, tema 3. Fonética Acústica*. España, recuperado 27 de abril de 2010 de <http://www3.unileon.es/dp/dfh/Milka/FyF/F3.htm>
- Zañartu, M. (2003) *Aplicaciones del análisis acústico en los estudios de voz humana*. Universidad Perez rosales, Chile. Recuperado 27 de Abril 2010 de <http://web.ics.purdue.edu/~mzanartu/Documents/Paper%20-%20Análisis%20voz%20humana%202003.pdf>

11. Anexos

El abuelo

Le voy a contar sobre mí abuelo. Bueno, el tiene cerca de noventa y tres años de edad y aún piensa tan lúcidamente como siempre. Se viste solo. Casi siempre se pone su vieja chaqueta negra que tiene varios botones menos, además de un gorro parecido al que usaba Pablo Neruda, de un color amarillento como las hojas del álamo en el otoño.

Mi abuelo se llama Luis. Una larga barba cuelga de su cara, inspirando a aquellos que lo observan un profundo sentimiento de respeto. Cuando habla, su voz parece un poco quebrada y temblorosa, y su discurso elocuente refleja la sabiduría de sus años vividos.

Dos veces al día él disfruta tocando hábilmente su preciada pianola que en los años cincuenta le compró a un inmigrante español.

Todos los días, por la mañana el abuelo da un corto paseo por el parque y por el barrio donde ha vivido toda su vida, y durante las tardes se divierte jugando brisca con sus queridos amigos Gustavo y Emilio. Excepto en el invierno cuando la lluvia o el frío se lo impiden.

11. 2. **Tablas de datos /a/:** señalan el contexto de los alófonos de /a/ a analizar en los 4 párrafos. El número a la izquierda de cada tabla corresponde al número del alófono a analizar y que coincide con el colocado en las hojas de registro que están en la página siguiente.

Alófonos tónicos de /a/					
N \ P	1	2	3	4	
1	contar	llama	tocando	mañana	
2	años	larga	hábilmente	da un	
3	edad	barba	Preciada	parque	
4	tan	cara	años	barrio	
5	casi	inspirando	inmigrante	durante	
6	varios	a aquella		tardes	
7	además	cuando		jugando	
8	usaba	habla		Gustavo	
9	pablo	quebrada		cuando	
10	álamo	años			
Total	10	10	5	9	34

Alófonos átonos de /a/					
N \ P	1	2	3	4	
1	voy a contar	abuelo	al	días	
2	mi abuelo	llama	día	mañana	
3	cerca de	una	disfruta	mañana	
4	noventa	larga	preciada	mañana	
5	aún	barba	pianola	abuelo	
6	piensa	cuelga	pianola	paseo	
7	lúcidamente	cara	cincuenta	ha	
8	vieja	aquellos	compro a un	toda su	
9	chaqueta	observan	español	vida	
10	chaqueta	habla		las tardes	
11	negra	parece		brisca	
12	además	quebrada		amigos	
13	parecido	temblorosa		la lluvia	
14	al	refleja		lluvia	
15	usaba	la			
16	Neruda	sabiduría			
17	amarillento	sabiduría			
18	amarillento				
19	las hojas				
20	hojas				
21	álamo				
Total	21	17	9	14	61

11.3. Hojas de registro de alófonos de /a/: la primera (Fig. 11.3.a.) corresponde a la hoja de análisis de las tónicas, y la segunda (fig. 11.3.b) y la tercera (fig. 11.3.c), a las hojas de los alófono de /a/ átonos. Estas últimas están separadas en 2 para fines prácticos, debido a su gran extensión.

Hoja de registro para analizar alófonos de /a/ tónica																																				
Párrafo	1										2										3					4									x	s
N°	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	1	2	3	4	5	6	7	8	9		
f0																																				
f1																																				
f2																																				
f3																																				
f0-f1																																				
f1-f2																																				
f2-f3																																				
f0-f3																																				

Fig. 11.2.a.

Hoja de registro para analizar /a/ átonas párrafos 1 y 2																																								
Párrafo	1											2																x	s											
N°	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17		
f0																																								
f1																																								
f2																																								
f3																																								
f0-f1																																								
f1-f2																																								
f2-f3																																								
f0-f3																																								

Fig. 11.2.b.

Hoja de registro para analizar /a/ átonas en párrafos 3 y 4																								
Párrafo	3									4							x	s						
N°	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9	10	11	12	#	#	
f0																								
f1																								
f2																								
f3																								
f0-f1																								
f1-f2																								
f2-f3																								
f0-f3																								

Fig. 11.3.c.

11.4. Tablas de datos /i/: señalan el contexto de los alófonos de /i/ a analizar en los 4 párrafos. El número a la izquierda de cada tabla corresponde al número del alófono a analizar y que coincide con el colocado en la hoja de registro (de acuerdo a cada párrafo).

Alófonos tónicos para /i/					
N \ P	1	2	3	4	
1	mí abuelo	Luís	día	días	
2	viste	sabiduría		vivido	
3	parecido	vividos		vida	
4				brisca	
5				queridos	
6				amigos	
7				Emilio	
8				frío	
9				impiden	
Total	3	3	1	9	17

Alófonos átonos para /i/					
N \ P	1	2	3	4	
1	voi a	mi	disfruta	parque i por	
2	tiene	inspirando	hábilmente	barrio	
3	setenta i tres	inspirando	preciada	vivido	
4	edad i aun	sentimiento	pianola	vida i durante	
5	piensa	sentimiento	cincuenta	divierte	
6	lucidamente	i temblorosa	inmigrante	divierte	
7	siempre	i su discurso	inmigrante	Gustavo i Emilio	
8	casi siempre	discurso		Emilio	
9	siempre	sabiduría		invierno	
10	vieja	vividos		invierno	
11	tiene			lluvia	
12	varios			impiden	
13	amarillento				
total	13	10	7	12	41

11.5. Hojas de registro de alófonos de /i/

Hoja de registro para la /i/ tónica																		
Párrafo	1			2			3	4						x	S			
N°	1	2	3	1	2	3	1	1	2	3	4	5	6	7	8	9		
f0																		
f1																		
f2																		
f3																		
f0-f1																		
f1-f2																		
f2-f3																		
f0-f3																		

Fig. 11.5.a.

Hoja de registro para la /i/ átonas																																												
Párrafo	1													2							3			4									x	s										
N°	1	2	3	4	5	6	7	8	9	10	11	12	13	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	1	2	3	4	5	6	7	8	9	10	11	12		
f0																																												
f1																																												
f2																																												
f3																																												
f0-f1																																												
f1-f2																																												
f2-f3																																												
f0-f3																																												

Fig. 11.5.b.

11.6. Tablas de datos /u/: señalan el contexto de los alófonos de /i/ a analizar en los 4 párrafos. El número a la izquierda de cada tabla corresponde al número del alófono a analizar y que coincide con el colocado en la hoja de registro (de acuerdo a cada párrafo).

Alófonos tónicos de /u/					
N \ P	1	2	3	4	
1	aun	profundo	disfruta	lluvia	
2	lucidamente	discurso			
3	Neruda				
Total	3	3	1	9	16

Alófonos átonos de /u/					
N \ P	1	2	3	4	
1	abuelo	abuelo	...mente su preciada	abuelo	
2	bueno	Luis	cincuenta	da un corto	
3	su voz	una	compro a un	su vida	
4	de un gorro	cuelga		durante	
5	que usaba	su cara		jugando	
6	de un color	un profundo		con sus	
7		cuando		Gustavo	
8		su voz		cuando	
9		un poco			
10		su discurso			
11		elocuente			
12		sabiduría			
13		sus años			
total	6	13	3	8	30

11. 8. Tablas de datos /m/: los alófonos marcados con mayúscula son los trabantes, y los otros son los que articulan con vocal.

Alófonos trabantes de /m/					
N \ P	1	2	3	4	
1	sieMpre	Mi abuelo	hábilMente	iMpiden	
2	sieMpre	teMblorosa	coMpro		
3			inMigrante		
total	2	2	3	1	8

Alófonos de /m/ que coarticulan con vocal					
N \ P	1	2	3	4	
1	sobre mi abuelo	llama		la mañana	
2	lucidamente	sentimiento		amigos	
3	como siempre			Emilio	
4	menos				
5	además				
6	amarillento				
7	como las hojas				
8	álamo				
total	8	2	0	3	13

11.9. Hojas de registro /m/

Hoja de registro para la /m/ trabante										
Párrafo	1		2		3			4	x	s
N°	1	2	1	2	1	2	3	1		
f0										
f1										
f2										
f3										
f0-f1										
f1-f2										
f2-f3										
f0-f3										

Fig. 11.9.a.

Hoja de registro para la /m/ coarticulación vocálica																
Párrafo	1								2		3	4			x	S
N°	1	2	3	4	5	6	7	8	1	2	0	1	2	3		
f0																
f1																
f2																
f3																
f0-f1																
f1-f2																
f2-f3																
f0-f3																

Fig. 11.9.b.

11.10. Tablas de datos de los alófonos de /n/: Los alófonos marcados con mayúscula son los trabantes, y los otros los que articulan con vocal.

Alófonos trabantes de /n/						
N	P	1	2	3	4	
		1	coNtar	iNspirando	tocaNdo	uN corto
2		noveNta	inspiraNdo	habilmeNte	doNde	
3		auN	uN profundo	eN los	duraNte	
4		pieNsa	profuNdo	ciNcuenta	jugaNdo	
5		taN	seNtimiento	cincueNta	iNvierno	
6		lucidameNte	sentimieNto	iNmigrante	invierNo	
7		uN gorro	kuaNdo	inmigraNte	cuaNdo	
8		uN color	uN poco		impideN	
9		amarilleNto	elokueNte			
total		9	9	7	8	33

Alófonos de /n/ que coarticulan con vocal						
N	P	1	2	3	4	
		1	bueno	una larga	pianola	mañana
2		tiene	observan	a un inmigrante	con sus	
3		noventa			excepto en el	
4		pone				
5		negra				
6		tiene				
7		botones				
8		menos				
9		Pablo Neruda				
10		en el otoño				
total		10	2	2	3	17

11.11. Hojas de registro /n/

Hoja de registro para la /n/ trabante																																			
Párrafo	1									2					3				4					x	s										
N°	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	1	2	3	4	5	6	7	8		
f0																																			
f1																																			
f2																																			
f3																																			
f0-f1																																			
f1-f2																																			
f2-f3																																			
f0-f3																																			

Fig.11.11.a.

Hoja de registro para la /n/ en coarticulación vocálica																			
Párrafo	1										2	3	4	x	s				
N°	1	2	3	4	5	6	7	8	9	10	1	2	1	2	1	2	3		
f0																			
f1																			
f2																			
f3																			
f0-f1																			
f1-f2																			
f2-f3																			
f0-f3																			

Fig. 11.11.b.