

Universidad Central “Marta Abreu” de Las Villas

Facultad de Ingeniería Eléctrica

**Centro de Estudios de Electrónica y Tecnología de la
Información**



TRABAJO DE DIPLOMA

Medición Acústica Multidimensional de Jadeo

Autor: Dianet González Santibañez

Tutor: Dr. Carlos Ariel Ferrer Riesgo

Santa Clara

2014

"Año 56 de la Revolución"

Universidad Central “Marta Abreu” de Las Villas

Facultad de Ingeniería Eléctrica

**Centro de Estudios de Electrónica y Tecnología de la
Información**



TRABAJO DE DIPLOMA

Medición Acústica Multidimensional de Jadeo

Autor: Dianet González Santibañez

E-mail: dgsantibanez@uclv.edu.cu

Tutor: Dr. Carlos Ariel Ferrer Riesgo

E-mail: cferrer@uclv.edu.cu

Santa Clara

2014

"Año 56 de la Revolución"



Hago constar que el presente trabajo de diploma fue realizado en la Universidad Central “Marta Abreu” de Las Villas como parte de la culminación de estudios de la especialidad de Ingeniería Biomédica, autorizando a que el mismo sea utilizado por la Institución, para los fines que estime conveniente, tanto de forma parcial como total y que además no podrá ser presentado en eventos, ni publicados sin autorización de la Universidad.

Firma del Autor

Los abajo firmantes certificamos que el presente trabajo ha sido realizado según acuerdo de la dirección de nuestro centro y el mismo cumple con los requisitos que debe tener un trabajo de esta envergadura referido a la temática señalada.

Firma del Autor

Firma del Jefe de Departamento
donde se defiende el trabajo

Firma del Responsable de
Información Científico-Técnica

PENSAMIENTO

*Si hay algo que puedes hacer o sueñas que eres capaz de hacer, encáralo. El coraje posee genio, magia y poder en sí mismo. Sólo comprométete y tu mente crecerá impetuosa. Empieza y la tarea será terminada.
La más larga caminata comienza con un paso.*

Proverbio hindú

DEDICATORIA

*A las mujeres más maravillosas del mundo: Moraima, Maribel y Arminda. Por
suerte para mí, mis MADRES.*

A mis PAPÁS: Delvito, Ariel, y mi abuelo "Santibañez".

*A José Carlos, por todos estos años juntos, y hacer mis días más fáciles, y
felices*

AGRADECIMIENTOS

Para todos aquellos que han inspirado en mí el deseo de superación profesional y personal, mi gratitud eterna. Quisiera corresponder especialmente:

☺ *A mis abuelos Armida y Eduardo, que han estado cuidando todos mis pasos, desde que empecé a caminar, por el derroche de apoyo y supervisión, por inspirarme a dar lo mejor de mí siempre.*

☺ *A mi madre, no existe nadie que se halla sacrificado tanto como tú en todos estos años de estudio. Eres un ídolo para mí, creo que no me alcanzará la vida para retribuir todo lo que has hecho por mí.*

☺ *A mi tía Maribel y mi tío Ariel que a pesar de no estar aquí en estos últimos años, nunca me han faltado sus palabras de aliento para continuar. Por ser mis padres, de toda la vida, y por comprender mejor que nadie mis inquietudes y valorar mis méritos. Muchas gracias.*

☺ *A Delvito por ser mi figura paterna, por los consejos y conocimientos que me ha brindado siempre, por hacer tan feliz a mi mamá y ser su eterno compañero.*

☺ *A José Carlos, por complementarme a la perfección, por aceptarme como soy, y tolerar mis errores. Por seguirme hasta el final del mundo y ser tan especial. Por hacerme parte de una familia que aprecio y valoro muchísimo y considero mía también. A todos muchas gracias.*

☺ *A “La Amiga”. Elianny tú sabes que lo que nosotras hemos vivido juntas dudo que otro par de amigas lo haya vivido. Ha sido un privilegio conocerte, y poder compartir tantas experiencias contigo, pero de todo sin dudas lo mejor es el cariño que nos profesamos.*

☺ *A Karla, por estar conmigo también desde primer año, y espero sigamos juntas. Nos vemos en la vida laboral*

☺ *A Yeilis y demás compañeros del aula, de la universidad en general y fuera de esta, todos son personas maravillosas.*

☺ *A mi tutor, la verdad que hemos pasado un “poquitico” de trabajo, pero sin usted esto no hubiese sido posible. No me arrepiento de la decisión que tomé un día, lo considero la persona más original y excepcional del mundo. Para usted mi más sincero respeto y admiración.*

☺ *Al resto de los profesores, que durante todos estos años de estudio han moldeado la profesional que soy hoy. Son todos ejemplares.*

*Sin dudas este largo camino, no hubiese sido posible recorrerlo sin el apoyo de ustedes.
.....Llégue a todos mi gratitud eterna, de corazón.....*

Yo

TAREA TÉCNICA

- Búsqueda bibliográfica sobre los antecedentes del tema, sintetizando la información relevante y analizando críticamente la misma.
- Análisis de los índices empleados para medir jadeo.
- Análisis de las metodologías empleadas para establecer la correspondencia entre el índice y la valoración subjetiva.
- Selección de índices representativos de las características del jadeo.
- Diseño de una combinación de los índices.
- Obtención de una base de datos representativa de valoraciones de jadeo.
- Evaluación del índice.

Firma del Autor

Firma del Tutor

RESUMEN

La voz es el instrumento por excelencia, mediante el cual los seres humanos establecemos nuestras ideas y expresamos lo que somos. Para la comunidad científica, el diagnóstico y tratamiento de las afecciones que padece el habla es de gran interés. En la práctica clínica se emplea frecuentemente la voz del paciente como material diagnóstico y de seguimiento por ser de fácil obtención y constituir una técnica no invasiva. Entre los parámetros que más frecuentemente se utilizan para evaluar el estado del paciente se encuentran los de calidad vocal, generalmente referentes a los cinco rasgos propuestos por Hirano: aspereza, jadeo, ronquera, astenicidad y esfuerzo. Estos pueden ser evaluados mediante el empleo tanto de valoraciones subjetivas como de mediciones objetivas. Las medidas objetivas permiten encontrar detalles acústicos que son imperceptibles al sentido del oído, además de reducir la dependencia de la experiencia del especialista.

Este trabajo se dirige a desarrollar una medida acústica para el jadeo, que se corresponda con la percepción del mismo, partiendo de la valoración crítica de las medidas existentes hasta el momento. La medida fue evaluada sobre dos Bases de Datos, valoradas subjetivamente. Los resultados muestran que el algoritmo fue sensible, específico y consistente para ambas muestras.

TABLA DE CONTENIDOS

PENSAMIENTO	i
DEDICATORIA	ii
TAREA TÉCNICA.....	v
RESUMEN	vi
INTRODUCCIÓN.....	1
Organización del informe	2
CAPÍTULO 1. MEDICIÓN DE ALTERACIONES DEL HABLA	4
1.1 Habla Normal	4
1.1.1 Modelo de Producción del Habla.....	5
1.1.1.1 Modelo de producción del Habla desde el punto de Vista Anatómico	5
1.1.1.2 Modelo de Producción de Habla desde el punto de vista Ingenieril. Modelo Fuente-Filtro.....	6
1.1.2 Características Acústicas de los Sonidos	7
1.1.2.1 Sonoros	8
1.1.2.2 Sordos	9
1.2 Habla Patológica.....	9
1.2.1 Alteraciones por nivel de Análisis	10
1.3 Medición de Calidad Vocal	11
1.3.1 Mediciones Subjetivas	11

1.3.2	Mediciones Objetivas	13
1.3.3	Problemas de Correspondencia.....	14
1.4	Medición de Jadeo	15
1.4.1	Definición Perceptual	15
1.4.2	Medidas Empleadas	17
1.4.2.1	Perturbación de la periodicidad	17
1.4.2.1.1	Jitter & Shimmer	17
1.4.2.1.2	Ruido Aditivo.....	17
1.4.2.1.3	Disperiodicidad Total.....	18
1.4.2.2	Medidas Glotales	18
1.4.2.2.1	Duraciones relativas de las fases glotales	18
1.4.2.2.2	GNE.....	18
1.4.2.3	Conformación espectral.....	19
1.4.2.3.1	Medidas de Amplitud Relativa.....	19
1.4.2.3.2	Razones entre Bandas de Energía	19
1.4.2.3.3	Medidas Espectrales Globales.....	20
1.5	Análisis crítico y aspectos metodológicos sobre la medición de jadeo	20
1.5.1	Análisis crítico de algunos resultados.....	20
1.5.2	Medidas Acústicas a emplear	21
1.5.2.1	Características deseables de los resultados de las medidas	23
1.5.3	Base de Datos.....	23
1.6	Conclusiones parciales	24
CAPÍTULO 2. MATERIALES Y MÉTODOS.....		25
2.1	Base de Datos	25

2.1.1	Youry Maryn's. Descripción	25
2.1.2	Kay Elemetrics. Descripción	26
2.2	Medidas acústicas seleccionadas	26
2.2.1	Empleadas anteriormente.....	26
2.2.1.1	CPP y CPPs	26
2.2.1.2	Jitter	30
2.2.2	Orientadas a los dos factores	30
2.2.2.1	HNR.....	31
2.2.2.2	Prominencia del Primer Armónico.....	31
2.2.2.3	Combinación.....	33
2.3	Descripción del experimento	33
2.4	Conclusiones parciales	34
CAPÍTULO 3. RESULTADOS Y DISCUSIÓN		35
3.1	Distribución de las Valoraciones Subjetivas en las Bases de Datos.....	35
3.1.1	Youry Maryn's.....	35
3.1.1.1	Distribución	35
3.1.1.2	Correlación entre jueces	37
3.1.2	Kay Elemetrics.....	37
3.1.2.1	Distribución	37
3.1.2.2	Correlación entre jueces	39
3.2	Regresiones de las medidas con las Valoraciones G, R, B.....	39
3.2.1	Youry Maryn's.....	39
3.2.2	Kay Elemetrics.....	41
3.3	Conclusiones del capítulo.....	43

CONCLUSIONES Y RECOMENDACIONES	44
Conclusiones	44
Recomendaciones	44
REFERENCIAS BIBLIOGRÁFICAS	46
ANEXOS	51
Anexo I Programa de obtención de CPP y CPPs	51

INTRODUCCIÓN

Sin dudas una de las habilidades más importantes con que cuenta el ser humano es el habla. A través del lenguaje se materializan los pensamientos de una persona, sirviéndose del modelo o sistema que facilita la lengua. Con las palabras no solo transmitimos un mensaje, también estados de ánimos, emocionales y procedencia social. Debido al gran número de sistemas y órganos que intervienen en este proceso, existen múltiples afecciones que provocan alteraciones de la misma. Por estas razones la señal de voz, es el centro de estudio para muchos investigadores en la actualidad.

Las alteraciones patológicas del habla se evalúan tradicionalmente de manera subjetiva por los especialistas médicos. Los rasgos subjetivos a evaluar se pueden dividir en tres grandes categorías, referentes a las unidades del habla involucradas: Calidad Vocal (fonemas sostenidos), Articulación (sílabas) o Prosodia (palabras y unidades de mayor duración). Las alteraciones de Calidad Vocal, generalmente han sido valoradas perceptivamente mediante el empleo de los rasgos: aspereza, jadeo, ronquera, astenicidad y esfuerzo. Estos cinco rasgos conforman la escala GRBAS propuesta por (Hirano, 1981), para la evaluación clínica de las disfonías.

Con el transcurso del tiempo se han realizado esfuerzos por efectuar de manera objetiva esta evaluación clínica, reduciendo la dependencia de la experiencia del evaluador. A este fin se han desarrollado mediciones sobre la señal acústica orientadas a brindar una medición objetiva equivalente, de alguna manera, a la de un parámetro subjetivo dado. Estas medidas intentan por una parte medir las características físicas que se le atribuyen a una voz con presencia del rasgo dado, y por otra corresponder proporcionalmente en magnitud a la percepción del rasgo.

Existen múltiples medidas objetivas de Calidad Vocal, más de 500 referencias en (Buder, 2000) (perturbaciones de la periodicidad, parámetros de la forma de onda del flujo glotal, etc.) que se han propuesto como índices de severidad para varios rasgos. Sin embargo estas medidas no han resultado lo suficientemente sensibles, o específicas al rasgo que se está evaluando, o consistentes entre diferentes estudios. Las causas pueden ser múltiples, pero en este trabajo se explora la posibilidad de que la percepción del jadeo en particular sea de carácter multicausal, y responda a varias características de la señal acústica y no sólo a una de ellas.

Por lo tanto el problema de investigación que se impone es la no existencia hasta el presente de mediciones objetivas que respondan con claridad a la definición y percepción del jadeo.

Esta situación puede ser resuelta con el cumplimiento de los siguientes objetivos:

Objetivo General:

- Desarrollar un índice objetivo de jadeo que se corresponda con la definición y percepción del mismo.

Objetivos Específicos:

- Determinar las características distintivas del jadeo.
- Proponer un algoritmo de medición que reúna las características seleccionadas.
- Evaluar el algoritmo en un conjunto de sujetos valorados subjetivamente.

Impacto posible:

Luego de realizado este proyecto se persigue científicamente mejorar la comprensión del fenómeno de la percepción del jadeo; socialmente tiene implicaciones en la mejora de la calidad en la atención médica y económicamente contribuye a reducir los errores y la duración de los tratamientos.

Organización del informe

Este informe se distribuye de la siguiente manera:

En el Capítulo 1 se definen cuestiones inherentes al proceso de producción del habla tanto anatómicas como desde el punto de vista ingenieril. Posteriormente se comentan las alteraciones patológicas del habla, y los tipos de medidas que se han empleado para cuantificar el nivel patológico en los pacientes. En el caso particular del jadeo, se tratan diferentes definiciones dadas por varios autores, así como las medidas objetivas que se han empleado. Por último quedan sentadas las bases para la elección de las medidas, que se relacionen con la percepción del jadeo.

En el Capítulo 2 se describen las Bases de Datos que se emplearán en el proyecto. Se selecciona un total de 3 medidas (CPP, CPPS, Jitter) atendiendo a los resultados analizados en el Capítulo 1. Luego se sugieren las medidas HNR y FHP, dada la relación teórica que guarda con las características perceptuales del jadeo, y posteriormente la definición matemática de cada una. Por último se describe el método seguido en este proyecto.

El Capítulo 3 contiene los valores de correlación entre los jueces para estas alteraciones, además de las tablas de resultados de las medidas para Grado, Aspereza y Jadeo, donde numéricamente puede comprobarse si las medidas son sensibles, específicas y consistentes.

Por último se presentan las conclusiones finales de la tesis y las recomendaciones de trabajo futuro.

El cuerpo de la tesis consta 50 páginas, con 8 tablas, 11 figuras y 6 ecuaciones contenidas en el texto, además de 1 Anexo.

CAPÍTULO 1. MEDICIÓN DE ALTERACIONES DEL HABLA

El presente capítulo introduce las características fundamentales del habla desde su importante papel como medio de comunicación para los humanos hasta su modelo de producción ya sea visto anatómicamente o con una perspectiva ingenieril. En su desarrollo se brinda una panorámica de las alteraciones que puede sufrir el habla, y el empleo de técnicas de medición subjetivas y objetivas, la correspondencia entre ambas y por qué puede fallar esta. Por último se plantea el problema de la medición objetiva del jadeo como rasgo de la escala GRBAS, algunas consideraciones sobre trabajos reportados y las técnicas a emplear para su medición a partir de la relación teórica con la definición perceptual.

1.1 Habla Normal

En los manuales de anatomía se habla de la voz humana como una función secundaria dentro del aparato respiratorio. Nuestra necesidad de respirar es constante y vital pero ello no quita importancia al papel jugado por el aparato respiratorio/fonador en la comunicación oral lograda por los humanos, así como el uso individual que cada persona hace del modelo general de la lengua, conocida también como habla (Bernal et al., 2000).

Dada la importancia del lenguaje hablado, como portador de un mensaje en el cual queda implícito información sobre los estados de ánimo, emocionales así como la procedencia social de una persona, resulta comprensible el gran interés de la comunidad científica internacional en el campo de las aplicaciones médicas del procesamiento de voz, lo cual redundará en un sinnúmero de investigaciones dedicadas a su estudio.

1.1.1 Modelo de Producción del Habla

Se puede concebir el habla desde el punto de vista del trayecto del mensaje, comenzando con la concepción en el cerebro hasta la emisión de la señal acústica, mediante impulsos nerviosos que llegan a cada uno de los órganos involucrados, lo cual se corresponde con la visión lingüística, Figura 1.1. o desde el punto de vista ingenieril, modelando los distintos elementos del proceso (Fant, 1960).

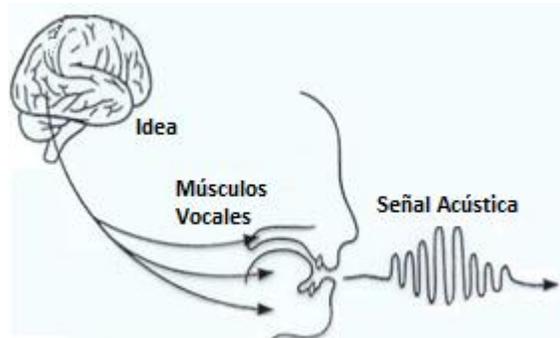


Figura 1.1. Modelo de Producción de Voz

1.1.1.1 Modelo de producción del Habla desde el punto de Vista Anatómico

El proceso de producción del habla atraviesa por varias etapas, que pueden ser ubicadas gráficamente en la figura 1.2.



Figura 1.2. Aparato Fonador

El proceso se inicia con la espiración del aire que es librado por los pulmones al ser presionados por el diafragma. Este flujo de aire no es audible, pues no produce sonidos.

Existen tres maneras de generar sonidos a partir de ese flujo. Dos de ellas se generan partiendo de la creación de turbulencias en el flujo, y se diferencian en cuanto a la manera de crear dicha turbulencia. Una manera es la formación de una constricción o estrechamiento en el tracto vocal, dando lugar a los sonidos “fricativos”, que se caracterizan por poder tener una duración más o menos larga, a voluntad del locutor. La segunda manera es la oclusión total del tracto en alguna ubicación del mismo, y la repentina liberación de la misma. El sonido generado es de muy corta duración, pero generalmente de mayor intensidad que las fricativas. La tercera manera de generar sonidos surge de una vibración cuasiperiódica a diferencia del ruido turbulento aperiódico de las anteriores. En esta, el flujo de aire al pasar a través de las cuerdas vocales las hace vibrar conformando los llamados sonidos sonoros. La frecuencia, a la que vibran las cuerdas es determinada, por la tensión de las mismas; a esta frecuencia se le conoce como frecuencia fundamental; cuando el tono es grave indica que la frecuencia es baja y viceversa.

Los sonidos fricativos y explosivos se conocen como sonidos “sordos”, mientras que los sonidos generados a partir de la vibración de las cuerdas vocales (periódicos) se conocen como sonidos “sonoros”.

Para todos estos según como se encuentren articulados los órganos, a partir del punto de generación de sonido, se formará una caja de resonancia distinta, la cual potenciará un conjunto de frecuencias y atenuará el resto. Aunque articulemos de forma similar los distintos fonemas, según la distancia, forma dureza, etc. de los órganos, aparecen características especiales de cada individuo, lo cual se define como timbre. Finalmente sale al exterior la voz.

1.1.1.2 Modelo de Producción de Habla desde el punto de vista Ingenieril. Modelo Fuente-Filtro.

La onda del habla es la respuesta del sistema de filtro del tracto vocal a una o más fuentes de sonido. La teoría fuente-filtro de la producción del habla se ejemplifica por

el diagrama de bloques de la Fig.1.3 que contiene un número de secciones interconectadas de filtro, cada una representa una parte de la cavidad vocal (Fant, 1960). En esta figura se puede apreciar la fuente de sonido de las cuerdas vocales, además del acoplamiento de las cavidades nasales con el resto del tracto vocal en el límite entre la cavidad de la faringe y la cavidad bucal.

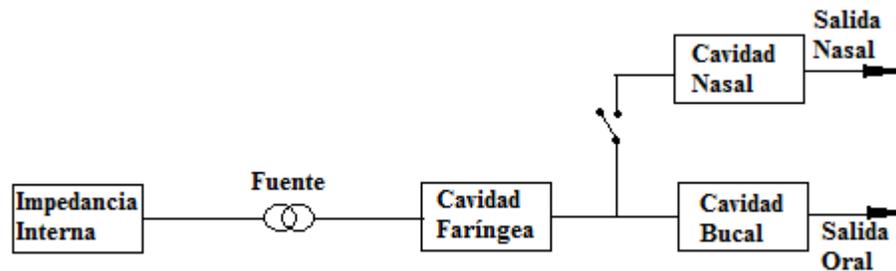


Figura 1.3. Producción del habla como un proceso de filtrado.

La función del filtro, en la terminología técnica se refiere a menudo como la función de transferencia ($H(f)$), que a su vez está definida como la relación entre un punto en el campo de sonido en la parte frontal del altavoz ($S(f)$) y la presión (o velocidad de volumen) de la fuente de sonido ($G(f)$).

Función de transferencia del filtro:

$$H(f) = \frac{S(f)}{G(f)} \quad (1.1)$$

La fuente de sonido es conmutable entre un ruido aleatorio (asumido generalmente como blanco y gaussiano) y una señal periódica (simplificada al máximo en forma de tren de impulsos de Dirac) en dependencia de si el sonido es sordo o sonoro.

1.1.2 Características Acústicas de los Sonidos

El estudio de la fonética acústica se vio mejorado en gran medida en el siglo 19 por la invención del fonógrafo de Edison. El fonógrafo permitió a la señal de voz ser grabada y luego procesada y analizada. La expresión verbal podría ser construida por la reproducción de la misma señal de voz desde el fonógrafo varias veces, cada vez filtrada con un filtro de paso de banda diferente. En una serie de artículos por Ludimar Hermann publicado en Pflügers Archiv en las dos últimas décadas del siglo 19 se investigó las propiedades espectrales de las vocales y consonantes con el

fonógrafo de Edison, y fue en estos documentos que el término “formante” fue introducido por primera vez (Hermann, 1890).

Los sonidos generados se clasifican de acuerdo a la periodicidad de la señal emitida en sonoros o sordos, como se mencionó en la sección 1.1.1.1.

1.1.2.1 Sonoros

Los sonidos sonoros son periódicos, se generan cuando se produce una constricción en las cuerdas vocales. Estas poseen una frecuencia de oscilación (conocida como Frecuencia Fundamental F_0 , o su recíproco Período Fundamental T_0) en dependencia de la tensión de los músculos aledaños su longitud y su masa, que permite que la generación del sonido vocálico pueda mantenerse mientras exista la presión de aire en los pulmones.

Una captura de la señal sonora en el tiempo evidencia la periodicidad antes descrita de los sonidos sonoros (Figura 1.4. Izquierda):

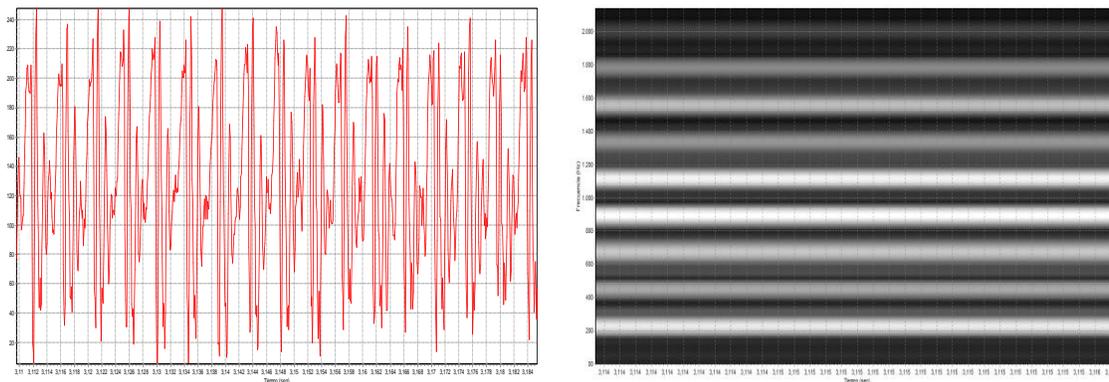


Figura 1.4. Características Acústicas de los sonidos sonoros para la vocal ‘a’. A la Izquierda en el tiempo (Oscilograma), a la Derecha en la frecuencia (Espectrograma).

Los sonidos sonoros son típicos de las vocales (y algunas consonantes como “m”, “n”, “b”, etc.) en los cuales se observa una estructura de formantes definida por zonas estables a lo largo del tiempo donde existe una alta concentración de energía fruto de la emisión del flujo de aire por el conducto bucal sin apenas resistencia. Una forma de observar gráficamente lo antes descrito es mediante la técnica del espectrograma, en el cual se muestra la distribución de amplitudes de cada frecuencia (Figura 1.4. Derecha).

1.1.2.2 Sordos

Los sonidos sordos (aperiódicos), son únicos de las consonantes, y se generan al hacer pasar el aire por constricciones en el tracto vocal (consonantes fricativas: “f”, “s”, “j”, otras) o al liberar repentinamente una oclusión en el mismo (consonantes explosivas sordas: “p”, “t”, “k”).

Un sonido aperiódico al ser descompuesto en sus armónicos básicos resulta que es poco estable a lo largo del tiempo, comportamiento que se puede observar claramente en la Figura 1.5. Izquierda:

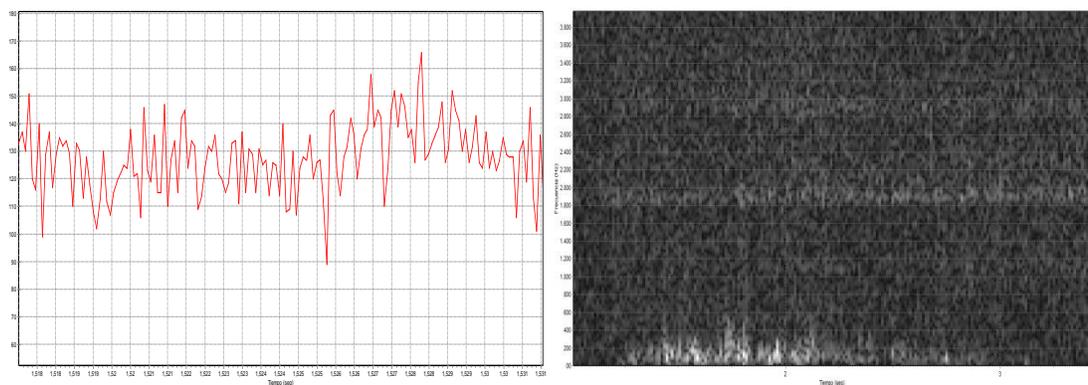


Figura 1.5. Características Acústicas de los sonidos sordos para la consonante ‘s’. A la Izquierda en el tiempo (Oscilograma), a la Derecha en la frecuencia (Espectrograma).

Además la energía está muy dispersa en el conjunto de armónicos posibles. El espectrograma de un sonido sordo evidencia una dispersión de la energía en la franja de altas frecuencias (Figura 1.5. Derecha).

1.2 Habla Patológica

Dada la multiplicidad de sistemas y órganos involucrados en la producción del habla existe un gran número de enfermedades que provocan alteraciones en la misma. Un ejemplo son los Trastornos Motores del Lenguaje, lo que equivale al quebrantamiento sobre el control de los movimientos musculares del aparato fonador como consecuencia de una lesión del sistema nervioso central o periférico.

1.2.1 Alteraciones por nivel de Análisis

Las alteraciones del habla pueden clasificarse en tres tipos, de acuerdo con la escala de tiempo en que se perciben:

- Alteraciones de la Calidad Vocal: Son perturbaciones que se perciben en sonidos aislados, generalmente vocales. Son característicos de este grupo las perturbaciones de la periodicidad de la señal de excitación generada en la glotis (la fuente en el modelo fuente-filtro) y la hiper/hipo-nasalidad, correspondiente a un inadecuado acople de la cavidad nasal al tracto vocal, que modifica su estructura de formantes (Klatt and Klatt, 1990, Shrivastav and Sapienza, 2003). Se reportan múltiples parámetros de la señal acústica que resultan anómalos a nivel de sonido individual (*jitter*, *shimmer*, ruido aditivo, medidas espectrales) (Buder, 2000). La medición se aprecia en sonidos (fonemas) sostenidos, y entre los rasgos más empleados en la literatura se pueden mencionar la aspereza, el jadeo, la ronquera, la astenicidad y el esfuerzo.
- Alteraciones de la Articulación: Estas perturbaciones se aprecian en las transiciones entre sonidos, por lo que requieren que la persona pronuncie como mínimo sílabas. La causa está en un inadecuado control de los músculos que regulan la conformación del tracto vocal, y algunos rasgos representativos son las consonantes imprecisas, los fonemas prolongados o repetidos, y las vocales distorsionadas. La característica acústica más reportada es la alteración de las trayectorias temporales de los formante (Kent, 2000). El nivel de complejidad de las mediciones acústicas es considerablemente superior al caso de Calidad Vocal, pues se requiere detectar la zona de interés dentro de la grabación (primer problema, no trivial en habla patológica) para luego obtener las curvas de interés (segundo problema) y compararlas (tercer problema) con patrones normales (cuarto problema).
- Alteraciones de la Prosodia: Se perciben en unidades del habla de mayor duración, como frases u oraciones y también tienen causa en el control muscular. Como ejemplos pueden mencionarse la acentuación igual y excesiva, las ráfagas cortas del habla, los silencios inapropiados y las monotonías de intensidad y tono. La complejidad en la obtención de medidas de anomalía es mucho mayor que en el caso

de Articulación, pues se mantienen los cuatro problemas mencionados, pero con un nivel de dificultad superior. Resulta particularmente complicado el establecimiento de los patrones de normalidad, con rangos muy amplios en cualquiera de los parámetros a considerar.

Existen gran cantidad de trabajos reportados sobre medidas de calidad vocal, a pesar de las dificultades inherentes al procesamiento digital y la variabilidad de la señal, que se amplifica por los desórdenes del habla; pero por otro lado hay una gran carencia en cuanto a medidas objetivas relacionadas con la articulación y la prosodia debido a la dificultad que acarrea analizar unidades más complejas del lenguaje (Schoentgen, 2006).

1.3 Medición de Calidad Vocal

La calidad vocal es una interacción entre una señal acústica de voz y un oyente, la señal acústica por sí misma no posee calidad, la evoca en el oyente (Kreiman and Gerratt, 2000). Por esta razón, las medidas acústicas son propias de lo que escuchan los oyentes (Gerratt and Kreirnan, 1995, Kreiman and Gerratt, 1996). Como estas medidas pueden ser altamente variables, ha sido de gran importancia la validación de una escala de calidad vocal, puesto que designa los protocolos de medida para los modelos de interpretación.

Dado que la calidad vocal es una respuesta perceptual para una señal acústica, al evaluar protocolos para la calidad vocal se encuentran diferencias acerca de ambas: la señal y el oyente.

1.3.1 Mediciones Subjetivas

La categorización del tipo de voz ha sido tradicionalmente lograda a través de la evaluación perceptiva, y hasta la fecha, muchos consideran la evaluación perceptual de la voz el método clave mediante el cual las disfonías se identifican y se realiza el seguimiento en cuanto a los avances en la terapia.

Cuando la calidad vocal es medida por medio de promedios en escalas para aspectos particulares de calidad, esto implica que la impresión global que el oyente recibe puede estar descompuesta en varios aspectos perceptivamente distintos siendo esto

propio de condiciones diversas como jadeo y la aspereza. Aunque la categorización perceptual del tipo de calidad de voz puede parecer obvia en ciertos casos, la categorización perceptual-auditivo puede ser difícil en varias situaciones (Shaheen et al., 2005):

- Cuando el paciente tiene una disfonía relativamente leve.
- Cuando el tipo de disfonía puede ser mixta o inconsistente.
- Cuando el examinador tiene una limitada experiencia en la categorización del tipo de calidad de voz.
- Cuando es hecha por diferentes jueces con variados entrenamientos y métodos para realizar la evaluación perceptual.
- Cuando se trata de realizar un seguimiento objetivo y los cambios son relativamente sutiles en el tipo de calidad de voz a través del tiempo.

En este tipo de valoración el instrumento de medición empleado es el oído humano. El sentido del oído está especializado en integrar las unidades del habla (de fonemas a sílabas, a palabras, a frases, a oraciones), para favorecer la extracción de información lingüística. Esto hace que muchas veces el escucha no quede con una clara conciencia de los detalles acústicos que se combinan para formar una determinada percepción, y son generalmente estos detalles los que deben ser evaluados en la práctica clínica (Baken, 1999). El oído es confundido muy fácilmente, por ejemplo: la percepción de la frecuencia (tono) fundamental de la voz puede ser más influenciada por la conformación espectral del tracto que por la frecuencia real del armónico fundamental (Baken, 1999, Ferrer and Hernández-Díaz, 2001).

Para la evaluación subjetiva se requiere establecer algunos aspectos de procedimiento:

1. Definición del término empleado para denominar el rasgo que se desea evaluar. En diferentes estudios se reportan muy diversas cantidades rasgos a los que se le atribuye valor diagnóstico (Darley et al., 1969a, Hirano, 1981, Enderby, 1983, Gelfer, 1988), incluso con definiciones diferentes para el mismo rasgo o término. A esto se suma la influencia de la experiencia del

especialista en la interpretación del término y la evaluación del rasgo en la práctica.

2. Determinación de la escala de medición empleada, con dos tipos principales: continuas y discretas. En las primeras se emite un número cualquiera en un intervalo dado, mientras en la segunda el número se selecciona de un conjunto finito de opciones (generalmente los enteros). No se reporta un resultado concluyente sobre la superioridad de una u otra alternativa, aunque las escalas discretas de 4 y 7 puntos son las más empleadas (Darley et al., 1969a, Kreiman et al., 1993, Hirano, 1981). En particular, la escala propuesta por Hirano es muy empleada por ser muy sencilla (5 rasgos, escala discreta de 4 puntos) y sin embargo brindar resultados satisfactorios en la práctica clínica. Esta escala se conoce por las siglas GRBAS a partir de las iniciales de los rasgos propuestos por Hirano (Grade: nivel general de anomalía, Roughness: aspereza, Breathiness: jadeo, Asthenicity: astenicidad y Strain: esfuerzo).

Los rangos perceptuales se usan también para validar medidas objetivas de voz, ya sean acústicas o instrumentales de otro tipo (Krom, 1995, Fritzell et al., 1986, Hillenbrand et al., 1994b, Martin et al., 1995, Sodersten et al., 1995, Ferrer et al., 2005b).

1.3.2 Mediciones Objetivas

La respuesta al problema de la confiabilidad de las medidas subjetivas se ha basado en la creación de medidas objetivas que surgen a partir del análisis acústico. Este se puede clasificar como: análisis en el dominio del tiempo, en el dominio de la frecuencia, y en el dominio tiempo-frecuencia de la señal de voz.

Las mediciones de parámetros de la señal acústica, son una tendencia internacional, en el caso particular de las aplicaciones médicas se encuentran ampliamente abordadas en la literatura a partir de los trabajos de (Lieberman, 1963) aunque, al igual que las valoraciones subjetivas, no están exentas de dificultades. Una tendencia marcada en las mediciones acústicas es la de intentar que los parámetros obtenidos se correspondan con rasgos con valor diagnóstico descritos en las valoraciones subjetivas tradicionales. En esta dirección se desea que el índice sea sensible

(altamente correlacionado) a la percepción del rasgo, que sea específico (no se correlacione significativamente con otros rasgos) y además consistente (ante diferentes muestras y paneles de jueces) (Shrivastav and Sapienza, 2003).

1.3.3 Problemas de Correspondencia

Las medidas objetivas de Calidad Vocal no han logrado imponerse definitivamente en la práctica clínica como un complemento a las valoraciones subjetivas del especialista, a pesar de su proliferación (más de 500 referencias en (Buder, 2000)). Esto se debe, en parte, a que los terapeutas del lenguaje no están muy expuestos a las ciencias exactas en su entorno profesional (Baken, 1999) pero mayormente debido a que la correspondencia entre las mediciones acústicas y las valoraciones subjetivas no ha resultado consistente entre diferentes experimentos (Kreiman and Gerrat, 2000). Los casos más tratados de aspereza, jadeo y ronquera han producido correlaciones desde insignificantes (cerca de cero) hasta muy altas (0.94) (ver tablas resumen en (Kreiman and Gerrat, 2000)). En ese trabajo se plantean como posibles explicaciones a la ausencia de correspondencia, las siguientes:

1. Incorrecta suposición teórica en cuanto a la correspondencia entre la valoración subjetiva y la medida objetiva. En este caso la valoración subjetiva es correcta, pero la medida objetiva no está realmente correlacionada con ella.
2. La suposición es correcta, pero existen fallos en la obtención de la medida objetiva. En este caso el problema es achacable al instrumento y al proceso de medición.
3. La relación existe y las medidas objetivas son correctas, pero hay problemas en la determinación de la correspondencia de estas últimas con las valoraciones subjetivas (errónea selección de la muestra o las técnicas estadísticas, efectos aleatorios, etc.).
4. El problema está en las valoraciones subjetivas. Aquí cabe la inconsistencia de la coincidencia de los jueces, tanto entre ellos como de uno en específico, y la naturaleza multidimensional de las percepciones de los rasgos (por ejemplo, el efecto de enmascaramiento de un rasgo por la presencia de otro).

Es imposible determinar a través de estudios asociativos (entre las valoraciones subjetivas y las mediciones acústicas) cuál de las razones anteriores provoca la falta de correspondencia entre ambas. En (Kreiman and Gerrat, 2000) concluyen que la 4ta razón está definitivamente presente, lo cual estimula a los defensores de las mediciones objetivas. En las mediciones subjetivas de Calidad Vocal se consideran excelentes valores de correlación por encima de 0.7, y aceptables por encima de 0.5 (según la compilación de (Kreiman and Gerrat, 2000)).

1.4 Medición de Jadeo

La percepción del jadeo en la voz es indicada por los cambios en el espectro vocal, estos cambios están relacionados con alteraciones en la intensidad del ruido de aspiración y la pendiente espectral de la energía armónica (Shrivastav and Sapienza, 2003). Aunque se han propuesto varias medidas acústicas para cuantificar la percepción de la voz jadeante, la mayoría ha dejado de dar una correlación consistente y con altas puntuaciones.

1.4.1 Definición Perceptual

Con respecto a este término, jadeo, son varios los autores que han aportado un concepto en la medida de la forma particular en que cada uno lo percibe. A continuación se proponen definiciones tradicionales de Jadeo formuladas por varios autores:

“La cualidad jadeo, resulta cuando las cuerdas vocales fallan en aproximarse completamente cuando vibran, y un flujo estable de aire atraviesa audiblemente la glotis y las cavidades resonadoras... La debilidad de la voz y el jadeo están estrechamente vinculados, ya que es imposible producir tonos extremadamente intensos con calidad jadeante”(Fairbanks, 1940).

“Escape de aire audible, dando como resultado una fonación débil, relacionado con una incapacidad funcional para la abducción de las cuerdas vocales con firmeza” (Bassich and Ludlow, 1986) p.133.

“Ruido audible creado en la glotis, probablemente a raíz del cierre insuficiente de la glotis, las cuerdas vocales vibran, pero un poco retenidas” (Askenfelt and Hammarberg, 1986, Hammarberg and Gauffin, 1995) p.291.

“El jadeo, se cree que es generado por el ruido de turbulencia producido en la glotis... Este supuesto parece razonable si se piensa en un susurro que se deriva del ruido turbulento en la glotis y da una impresión parecida al jadeo” (Fukazawa et al., 1988) p.1189.

“Escape audible de aire a través de la glotis debido al insuficiente cierre glotal, el grado de severidad del jadeo es inversamente proporcional a la duración de la fase de la glotis cerrada” (Eskenazi et al., 1990) p.301.

“La fonación jadeante se caracteriza por una fuente glotal con (1) un aumento del cociente de apertura, lo que resulta en un crecimiento de la amplitud relativa de la componente fundamental en el espectro, y (2) una tendencia a que los armónicos más altos puedan ser sustituidos por el ruido de aspiración” (Klatt and Klatt, 1990) p.825.

“Escape audible de aire a través de la glotis, debido al cierre insuficiente” (Sodersten and Lindestad, 1990) (Zwirner et al., 1993) p.166.

“Impresión de ruido glotal, fuga de aire y turbulencia durante la fonación” (Krom, 1995) p.798.

“La combinación de tono vocal y el flujo de aire turbulento procedente de la glotis parcialmente abierta” (Martin et al., 1995) p.767.

“Contiene el sonido de la respiración (espiración) durante la fonación; acústicamente, el jadeo... tiene la mayor parte de su energía en la fundamental, además presenta un componente importante de ruido debido a la turbulencia en o cerca de la glotis” (Titze, 1995) p.336.

Puede apreciarse un consenso en las definiciones perceptuales citadas anteriormente. En cuanto a su causa fisiológica, el jadeo se ha relacionado con un insuficiente cierre, y por consiguiente se produce un escape de aire excesivo. La característica acústica más mencionada, resulta ser, el ruido audible de aspiración cuando el aire pasa a través de la glotis.

1.4.2 Medidas Empleadas

A continuación se brinda una panorámica de las diferentes medidas que se han empleado en la medición objetiva del jadeo, ya sea en el dominio del tiempo (secciones 1.4.2.1 y 1.4.2.2) o en el de la frecuencia (sección 1.4.2.3).

1.4.2.1 Perturbación de la periodicidad

Entre las medidas del dominio del tiempo, las más empleadas son las relacionadas con las perturbaciones de la periodicidad.

1.4.2.1.1 Jitter & Shimmer

Las medidas de perturbación de la periodicidad (como las distintas maneras de medir Jitter y Shimmer) se han empleado para predecir jadeo en un gran número de estudios. En trabajos realizados por (*Shrivastav and Sapienza, 2003*), (*Speyer et al., 2003*) los coeficientes r^2 para Jitter, resultaron valores excelentes 0.74 y 0.67 respectivamente. El porcentaje de este en (*Eskenazi et al., 1990*) fue aceptable $r^2=0.55$, en (*Martin et al., 1995*) fue débil 0.003. Para (*Prosek et al., 1987*) el Cociente de Perturbación del Pitch fue moderado con un $r^2=0.38$.

Los resultados para Shimmer son similares. En (*Feijoo and Hernández, 1990*) fue excelente predictor de jadeo $r^2= 0.91$. En (*Prosek et al., 1987*) resultó ser fuerte predictor con $r^2=0.60$. El valor medio de Shimmer en (*Martin et al., 1995*) fue moderado $r^2=0.42$ y en (*Kreiman et al., 1990*) fue muy bajo $r^2=0.07$. En (*Arends et al., 1990*) resultó bajo $r^2=0.34$.

1.4.2.1.2 Ruido Aditivo

La relación señal ruido, también se ha evaluado como predictora de jadeo. Para (*Shrivastav and Sapienza, 2003*) el SNR (Relación Señal-Ruido) correlacionó fuertemente con el jadeo $r^2=0.852$. En (*Krom, 1994*) se realizó un estudio para bandas de frecuencia seleccionadas de una grabación de vocal sostenida, para el cual los valores de HNR (Relación Armónico-Ruido) resultaron moderados entre 0.46 y 0.69. (*Speyer et al., 2003*) obtuvo resultados bajos $r^2=0.38$. Para (*Kreiman et al., 1990*) el

valor de r^2 fue bajo también 0.26 y en (Martin et al., 1995) similar $r^2=0.25$. Menos alentadores aún fueron para (Heman-Ackah et al., 2002) $r^2=0.12$.

1.4.2.1.3 Disperiodicidad Total

En el caso de las medidas globales los resultados son más estimulantes. Hasta el momento los mejores resultados se derivan de la Prominencia del Pico Cepstral (CPP), introducido por (Hillenbrand and Houde, 1996) con un valor excelente de $r^2=0.96$, otra medida empleada en este estudio fue RPK que obtuvo un $r^2=0.84$. El CPP se ha empleado además en (*Shrivastav and Sapienza, 2003*) con los siguientes resultados, individual 0.75, combinado con SNR 0.78 y esta misma combinación pero añadiendo el Shimmer obtuvo un $r^2=0.81$. Debido a estos excelentes resultados se decidió evaluar el índice de jadeo en este proyecto basado en su supuesta robustez.

1.4.2.2 Medidas Glotales

Otra vertiente frecuente en el dominio del tiempo es efectuar mediciones sobre la señal de excitación en la glotis.

1.4.2.2.1 Duraciones relativas de las fases glotales

Estas medidas han sido empleadas como predictoras de jadeo en (Alku and Vilkman, 1996b), donde se proponen las medidas de cociente AQ (Cociente Amplitud, Máximo de la señal de flujo / Mínimo de la derivada de la misma), SQ (Cociente de Velocidad) y OQ (Cociente de Apertura) entre otros, resultando en una relación monótona entre voces de jadeante - a normal - a esforzada. Todos estos cocientes aumentan en el caso particular de las voces jadeantes. Un resultado equivalente fue obtenido por (Bickley, 1982) y (Ladefoged, 1983) donde el OQ va en incremento a medida que aumenta la calidad vocal del jadeo. Para (Childers and Lee, 1991), el jadeo obtiene el valor extremo de OQ con respecto a otras patologías.

1.4.2.2.2 GNE

Estas siglas denotan una medida de correlación entre varias bandas de frecuencia de la señal glotal (GNE, Razón de Excitación Glotal-Ruido). Primeramente fue empleada por (*Frohlich et al., 1998*). En ese estudio, solo el GNE diferenció con

éxito entre un grupo de pacientes con cáncer, normales y distintos tipos de voces afónicas. El GNE se empleó junto a un grupo de medidas con resultados excelentes de 0.92, también en (Stranik et al., 2014).

1.4.2.3 Conformación espectral

En el dominio de la frecuencia se han efectuado múltiples medidas relacionadas con el jadeo.

1.4.2.3.1 Medidas de Amplitud Relativa

Las relaciones de amplitud es una de las tantas medidas que se ha empleado para predecir jadeo. El criterio de agrupación para esta serie de medidas es la relación entre puntos del espectro. En (Krom, 1994) se buscó la relación entre el primer y segundo armónico (H1-H2), dando un $r^2=0.40$. (Hillenbrand et al., 1994a) utilizó la amplitud del primer armónico, y este se correlacionó con un $r^2=0.66$. (Shrivastav and Sapienza, 2003) emplea varias relaciones H1-H2 con $r^2=0.553$, H1-A1 (Primer Armónico-Amplitud del Primer Formante) $r^2=0.575$ y H1-A3 (Primer Armónico-Amplitud del Tercer Formante) $r^2=0.417$, todos moderados.

1.4.2.3.2 Razones entre Bandas de Energía

La razón entre bandas de energía es empleada por (Krom, 1994), en este estudio se divide la grabación en tramas: la primera trama de 200 ms a partir del inicio de la vocal, 200 ms a mediados de la vocal, 500 ms después del inicio y por último uno de 1000 ms que abarcan tanto el inicio como los fragmentos posteriores. Esta actividad se efectúa con el fin de conocer la influencia que puede tener estudiar con espacios de tiempo distintos en las vocales sostenidas. Este método es utilizado también por (Stranik et al., 2014), donde combinado con otras medidas resultaron excelentes predictores de jadeo, numéricamente $r^2=0.92$.

Para estas medidas es importante señalar las ventajas que poseen con respecto a las medidas referentes a las relaciones entre amplitudes. Para estas últimas es imprescindible determinar la ubicación exacta de los formantes y armónicos mientras que las medidas entre bandas de energía no poseen esta limitante.

1.4.2.3.3 Medidas Espectrales Globales

En algunos trabajos se han empleado parámetros absolutos del espectro (valor de F_0 , intensidad global SPL (Nivel de Presión de Sonido), representación cepstral en MEL o Barks, etc. (escalas). En (Godino-Llorente *et al.*, 2010) MFCC (Coeficientes cepstrales en escala MEL) resultó medianamente bueno con un $r^2=0.57$. Para (Prosek *et al.*, 1987) F_0 predice jadeo con un resultado de $r^2=0.75$. Resultados análogos se alcanzaron para la medida SPL en (Klich, 1982).

1.5 Análisis crítico y aspectos metodológicos sobre la medición de jadeo

Este epígrafe parte de una investigación crítica sobre algunos resultados de las medidas citadas en el epígrafe inmediato superior. Con base en estos análisis se plantean dos medidas por estar teóricamente relacionadas con las definiciones perceptuales del jadeo. Por último se comentan las características idóneas que deben cumplir estas, así como las condiciones que deben tener las Bases de Datos para poder comprobarlas.

1.5.1 Análisis crítico de algunos resultados

A pesar de su popularidad, las medidas tradicionales de HNR, jitter y shimmer han demostrado ser indicadores inestables de disfonía en muchos estudios.

Aunque varias investigaciones han revelado asociaciones razonables entre las medidas de perturbación acústica y las categorías de calidad vocal, algunos investigadores han cuestionado la necesidad, validez y utilidad clínica de las medidas específicas de perturbación, especialmente, cuando se aplica a las voces con trastornos moderados o graves. Una de las razones para esta falta de previsión puede ser que lo que se percibe como la disfonía puede no ser lógicamente asociado con cualquier medida de perturbación. Otra posibilidad es que debido a que estas mediciones se basan en la capacidad de determinar la frecuencia fundamental (F_0), pequeños errores en la determinación de F_0 pueden dar lugar a errores importantes en la medición de las perturbaciones y HNR. Por lo tanto, debido a la dificultad en la determinación de F_0 , las mediciones precisas de periodicidad en muestras de voces disfónicas que son prácticamente periódicas son a menudo difíciles de obtener.

En el caso de las medidas cepstrales, en particular los CPP de Hillenbrand, hasta el momento son las medidas que mejores valores de correlación guarda con el jadeo, pero esto es objetable. La mayoría de los estudios que han empleado esta medida, parten de una selección inicial de las muestras jadeantes, como se puede observar en (Shrivastav and Sapienza, 2003), donde se restringieron la muestra a solo voces jadeantes, incluso el propio autor en (Hillenbrand and Houde, 1996) y (Hillenbrand et al., 1994b) seleccionó primeramente las muestras. En cambio para estudios con muestras más diversas (Heman-Ackah et al., 2002, Maryn et al., 2007a), las correlaciones fueron inferiores a 0.5. Al CPP se la señala, también la carencia de la especificidad requerida para ser un buen indicador de cualquier dimensión individual y ser el más adecuado como indicador del grado en general (Ferrer et al., 2007). La falta de especificidad del CPP es notorio en (Awan and Roy, 2005), donde destacó como el mejor en la separación de voces normales a partir de grabaciones patológicas, pero era incapaz de realizar la separación entre las voces jadeantes, áspera y ronca. En estudios de (Stephen and Lowit, 2014) para clasificación de pacientes con desórdenes motores del lenguaje, las medidas cepstrales CPP y CPPs devolvieron altos niveles de correlación con grado en general, jadeo y astenicidad.

1.5.2 Medidas Acústicas a emplear

Para evitar la primera causa de falta de correspondencia objetivo-subjetivo; las mediciones acústicas de un rasgo en específico deben estar teóricamente relacionadas con las características que se le atribuyen a esa cualidad en particular. Con esto se quiere decir que para la selección de una medida, debe existir un basamento teórico, que previamente justifique y asegure el éxito, en la elección de una medida y no otra. Las medidas acústicas de jadeo deben reflejar los rasgos que se han incluido en las definiciones perceptuales del jadeo.

Para algunos autores, como se plasmó en el epígrafe 1.4.1, el jadeo es el resultado del ruido de aspiración producido en la glotis, causado por el cierre incompleto de las cuerdas glotales, derivado de este concepto emergen dos medidas:

1. En concordancia con la manifestación acústica de ruido audible de aspiración cuando el aire pasa a través de la glotis, la medida que debe ser adoptada tiene que

vincularse a este planteamiento. El ruido representa una desviación del patrón de periodicidad debido a una aleatoriedad. Este concepto nos induce a descartar las otras dos medidas de perturbación de periodicidad, *Jitter* y *Shimmer*, por estar relacionadas específicamente con alteraciones del pulso.

2. Vinculada a la causa fisiológica, cierre incompleto de las cuerdas vocales, la medida que surge en primera instancia es: Prominencia del Primer Armónico (FHP).

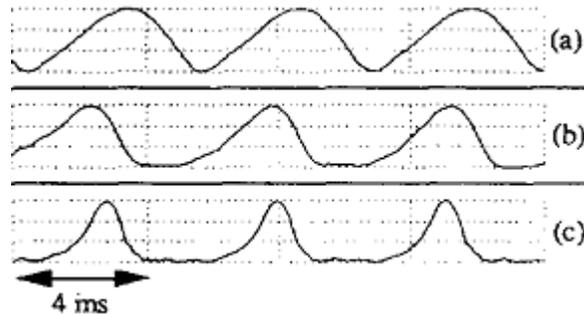


Figura 1.6. Formas de onda a) Voz jadeante, b) Voz Normal, c) Fonación forzada.

Tomada de(Alku and Vilkman, 1996a).

Como puede ser observado en la figura anterior, el comportamiento de una voz jadeante evidenciado en a), es prácticamente sinusoidal, esta estructura armónica puede ser prevista por la medida FHP.

Por lo anterior dicho debe tomarse en cuenta además, el carácter propio del jadeo de tener dos formas de manifestarse. Debido a esta condición la forma de medir el jadeo debe mostrar estas dos causas, esta es la razón por la cual este trabajo defiende el principio de que la medida de jadeo debe ser multifactorial, porque el jadeo por definición es multifactorial. Puede darse el caso igualmente, de que un factor anule al otro, esto puede entenderse para el caso de la voz susurrada, donde no existe vibración alguna de las cuerdas vocales solo ruido, donde no se observa la periodicidad, y viceversa, donde no existe ruido y la sinusoide expresa también jadeo, pero en general tiene dos componentes fundamentales.

En el caso de la aspereza definida por (Hirano, 1981), (Dejonckere et al., 1993, Martin et al., 1995) entre otros, como vibración irregular de las cuerdas vocales que se corresponde con una fluctuación irregular en la frecuencia fundamental y/o la

amplitud de la fuente glotal del sonido, si puede ser analizada por las perturbaciones de la periodicidad, antes mencionadas en cuanto a duración y amplitud de los pulsos.

Otra consideración que debe ser analizada, es la tendencia a emplear la mayor cantidad de medidas para determinar cuál de ellas, individual o combinadas con otras resulta mejor, sin poseer relación alguna con el rasgo a medir. Este tipo de proceder fue seguido por (Stranik et al., 2014), donde el autor escoge de un total de 92 medidas solo a 4, por ser las de mejores valores de correlación en la predicción del jadeo. Esta puede ser considerada una de las causas de inconsistencia para las medidas objetivas.

1.5.2.1 Características deseables de los resultados de las medidas

Una medida objetiva de la dimensión de calidad vocal en particular (es decir, la aspereza, jadeo etc) debe correlacionarse con la medida perceptual, además cumplir con varias características (Shrivastav and Sapienza, 2003):

- Sensible: La sensibilidad se expresa habitualmente en términos de la correlación de la medida objetiva con las calificaciones de percepción de la dimensión particular.
- Específica: no es sensible a otras dimensiones, es decir no se correlaciona con otras medidas.
- Consistente: capacidad de producir resultados similares a través de diferentes estudios.

1.5.3 Base de Datos

En muchos estudios existen limitaciones en cuanto a la posibilidad de evaluar las características deseables de las medidas a partir de la base de datos existente:

-Con respecto a la sensibilidad: es importante valorar la inclusión o exclusión de sujetos sanos o restricciones a solo sujetos patológicos. En el caso de que la base de datos se restrinja a sólo sujetos jadeantes, la evaluación de la sensibilidad tiende a una sobrevaloración, puesto que medidas que responden prácticamente a todas las patologías, como el Cepstrum, resultaría un excelente predictor para el jadeo, lo cual puede ser considerado como falso debido a la falta de especificidad, referenciada en el epígrafe 1.5.1.

-Con respecto a la especificidad: debe tenerse en cuenta la inclusión o exclusión de voces con varios rasgos (aspereza, jadeo, astenicidad, etc). Una prueba de especificidad, pudiera ser, examinar sensibilidad en una base sólo de jadeo, y la misma medida en una de sólo aspereza por ejemplo y comprobar ambos resultados, o efectuar la medición para todas las voces.

-Con respecto a la consistencia: se valora esta, comprobando la combinación de los dos aspectos anteriores en varios estudios.

Otras consideraciones sobre la Base de Datos están relacionadas con la manera en que se efectúan las valoraciones subjetivas: cantidad de jueces, escala de valoración empleada y ejercicio sobre el que se emiten las valoraciones. En (Kent, 1996) se describen múltiples aspectos de estas cuestiones que influyen en las valoraciones subjetivas.

1.6 Conclusiones parciales

A partir de lo descrito en este capítulo, se puede apreciar el consenso en cuanto a las características acústicas del jadeo. Se observan además múltiples aspectos criticables en los trabajos de medición objetiva del jadeo, relacionados con las medidas que se emplean, sus combinaciones, y las bases de datos utilizadas.

CAPÍTULO 2. MATERIALES Y MÉTODOS

En este capítulo se detallan las bases de datos empleadas en el trabajo y las medidas objetivas seleccionadas. Se describe además la manera escogida de establecer la correspondencia entre valoraciones subjetivas y las medidas objetivas.

2.1 Base de Datos

Las medidas subjetivas de este proyecto fueron tomadas de dos bases de datos la de Youry Maryn's (Maryn et al., 2007b) y la Kay Elemetrics (Kay-Elementrics, 1992). Estas bases de datos han sido empleadas anteriormente en varios trabajos sobre las disfonías y la escala GRBAS (Shrivastav and Sapienza, 2003, Bhuta et al., 2004, Godino-Llorente et al., 2010, Alpana et al., 2012). A continuación se hace una descripción detallada de las mismas.

2.1.1 Youry Maryn's. Descripción

La Base de Datos de Youry Maryn's, posee un total de 229 voces de pacientes recogidas sin seguir ningún tipo de regla específica. Fueron adquiridas con una frecuencia de muestreo de 44,1 kHz y 16 bits de resolución. Estas grabaciones consisten en la lectura de un párrafo estándar (números) y la fonación sostenida de la vocal "a", por un tiempo de 3 segundos. En la Base de Datos están incluidas las valoraciones de las dimensiones GRB de la escala GRBAS, realizadas por 5 jueces, terapeutas del habla con más de 5 años de experiencia, que evaluaron el nivel de alteración de estos rasgos de acuerdo a valores numéricos (0-3).

2.1.2 Kay Elemetrics. Descripción

La Base de Datos Kay Elemetrics (Kay-Elemetrics, 1992), cuenta con grabaciones de pacientes que presentan múltiples tipos de enfermedades, además de un conjunto de sujetos sanos. En ella se recogen 657 grabaciones de voces patológicas, realizando el ejercicio de vocal “a” sostenida. Las grabaciones de las vocales patológicas se encuentran digitalizadas a 50 KHz, y las vocales de sujetos sanos y las lecturas de párrafo a 25 KHz, todo ello con 16 bits de resolución. El nivel patológico se valoró con el empleo de la escala GRBAS, por un total de 3 jueces, los cuales valoraron la base en dos ocasiones con un tiempo intermedio de más de 2 semanas entre valoración.

2.2 Medidas acústicas seleccionadas

De acuerdo con lo planteado anteriormente en la sección 1.5, se decidió limitar la cantidad de medidas a evaluar al mínimo. Con este fin, se dispuso emplear una medida para cada factor del jadeo (ruido e insuficiente cierre glotal) y comparar sólo con los mejores resultados reportados (CPP y su variante suavizada CPPs) así como una medida de aspereza.

2.2.1 Empleadas anteriormente

De este conjunto de medidas, CPP surgió como el mejor predictor. Por esta razón, se decidió evaluar otros índices, con su supuesta robustez (de acuerdo con el análisis de las deficiencias planteadas en el capítulo 1. Dada la importancia que reporta esta medida se decide exponer, a continuación la forma de obtención del mismo.

Otra de las medidas empleadas resultó ser el *Jitter*, reportada como medida de aspereza, con el propósito de comprobar influencias de este rasgo sobre el jadeo.

2.2.1.1 CPP y CPPs

El método de obtención del CPP (Hillenbrand and Houde, 1996), es como sigue. Primeramente hay que partir del conocimiento del cepstrum como concepto, este es un espectro de potencia logarítmico de un espectro de potencia logarítmico. Para señales periódicas, el primer espectro de potencia muestra la energía a frecuencias

armónicamente relacionadas, y el segundo espectro (es decir, el espectro del espectro) mostrará un fuerte componente correspondiente a la regularidad de los picos armónicos. El tiempo ("quefrecy") en el pico cepstral corresponde al período fundamental de la señal. Una señal cuyo espectro muestra una estructura armónica bien definida mostrará un pico cepstral muy prominente. CPP es una medida de la amplitud del pico cepstral correspondiente al período fundamental, normalizada para la amplitud global de la señal (véase de (Krom, 1993), para una medida relacionada). Un primer pico cepstral se encuentra entre el mínimo y el máximo período fundamental esperado (3,3 a 16,7 ms, o 60 a 300 Hz, en los trabajos originales de Hillenbrand). El quefrecy en el pico cepstral generalmente corresponde al período fundamental, y la amplitud del pico refleja tanto el nivel de organización armónica como la amplitud global de la señal. Para normalizar la amplitud global, se calcula una línea de regresión lineal en relación a la magnitud quefrecy cepstral. La medida CPP es la diferencia en amplitud (en dB) entre el pico y el valor correspondiente en la línea de regresión situado debajo del mismo (véase la Figura 2.1). La medida CPP se calcula a partir de las señales filtradas, mediante un filtrado paso banda entre 2,5 y 3,5 kHz, y filtrado paso alto a 2,5 kHz. Las condiciones de filtrado, paso alto y paso banda se incluyen, debido a que el ruido de aspiración tiende a ser mayor para las medias y altas frecuencias (por ejemplo, (Klatt and Klatt, 1990)). El filtrado fue logrado con un filtro digital de cuarto orden Butterworth. La medida se realizó cada 10 ms utilizando una ventana de análisis de 1.024 puntos (41 ms). La medida de CPP, se promedia para todos los marcos de análisis. Debido a que el CPP es una medida del grado de organización armónica, es de esperar que las señales de voces jadeantes tiendan a mostrar valores más pequeños de CPP que las señales no jadeantes. No se corrigen los errores de seguimiento de tono, los cuales son muy comunes para señales jadeantes. Además, no se eliminan las regiones sin voz, en las señales de habla fluida.

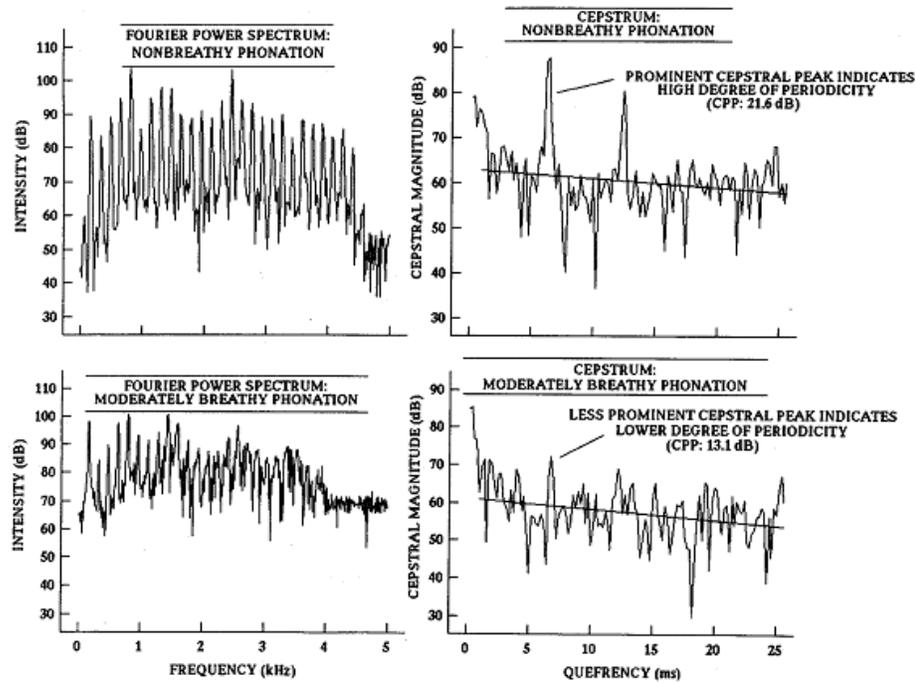


Figura 2.1. Espectro de Potencia de Fourier y representación cepstral para señales no jadeantes y moderadamente jadeantes. La señal no jadeante recibe una valoración de 0.8 para una escala de 0-10, la moderadamente jadeante recibe un 4.0. La regresión lineal relacionando quefrequency con magnitud cepstral, es usada en la medición de CPP para normalizar el pico cepstral en toda la señal. CPP es la diferencia de amplitud entre el pico cepstral y el valor de la regresión lineal que se encuentra directamente inferior al pico.

Otra medida desarrollada por Hillenbrand fue el CPP- suavizado (CPPs). Luego de experimentar varias veces con la medida descrita anteriormente, se descubrió que una modificación bastante simple al algoritmo del CPP produce una notable mejora en la precisión de la predicción. Se adiciona un paso al procesamiento, este implica suavizar el cepstrum individual antes de la extracción del pico cepstral y calcular la prominencia máxima. Para el CPPs, se calcula el pico cepstral cada 2 ms, en lugar de cada 10 ms, y el suavizado se lleva a cabo en dos pasos (véase la Figura 2.2). En primer lugar, como se muestra en el panel central de la figura 2.2, el cepstrum se promedia en el tiempo, es decir, cada cepstrum sin filtrar se sustituye por la media de un número de tramas cepstrales a la izquierda de la trama actual y un número de tramas cepstrales a la derecha de la trama actual. Por ejemplo, para una ventana de suavizado de cinco tramas, la salida suavizada para un marco dado estaría constituida por la media de la trama actual con los dos marcos anteriores y los dos subsiguientes.

En el siguiente paso de suavizado, que se muestra en el panel inferior de la Figura 2.2, se calcula un promedio de la magnitud cepstral a través del espacio quefreny. Por ejemplo, para una ventana con un promedio de cinco cuadros, cada una de las magnitudes cepstrales se sustituye por el promedio de la actual con las dos adyacentes de menor quefreny y las dos adyacentes de mayor quefreny. La experimentación con una variedad de configuraciones de ventanas de suavizado mostró que tanto ventanas muy grandes de suavizado a través del tiempo y relativamente pequeñas a través de quefreny funcionan mejor para vocales sostenidas. Se utilizó pues, una ventana de tiempo de suavizado de 150- marcos (300 ms), seguido de un suavizado de 3 puntos en quefreny, esto en el caso de vocal sostenida. Debido a que la frecuencia fundamental (F_0) varía con el tiempo en habla fluida, se necesitan ventanas de tiempo suavizadas mucho más pequeñas. Con este fin se utiliza una ventana de tiempo de suavizado de 10 marcos (20 ms), seguido de 10 bin quefreny.

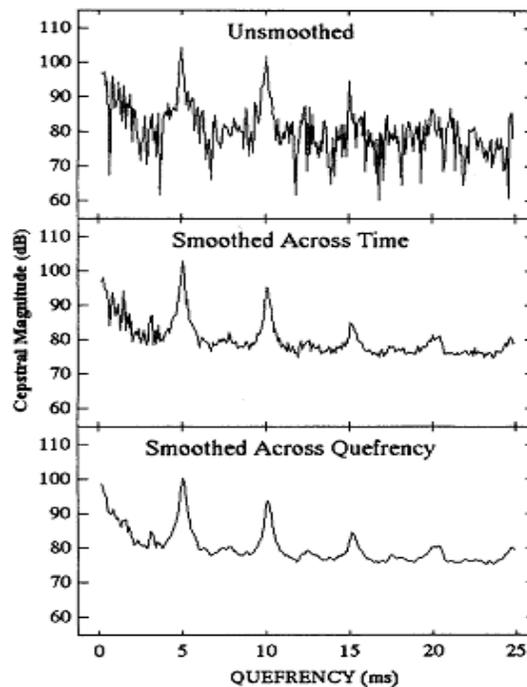


Figura 2.2. Muestra el método de suavizado usado en CPPs. La primera figura es la original, no suavizada; la central muestra el cepstrum después de suavizado a través del tiempo; en la base el cepstrum luego de suavizado en el dominio quefreny

El cálculo de la medida CPP se hizo a través del software desarrollado por Hillenbrand, tomado de (Hillenbrand, accessed May 21, 2012), al cual se le especifica primeramente si la base de datos corresponde a habla fluida o vocal sostenida con los comandos SP o VO respectivamente, estas son las abreviaturas indicadas en el modo de uso con letras mayúsculas. Estas opciones se pueden comprobar en el mensaje del uso del programa. El valor por defecto para cada opción se da entre corchetes al lado de la descripción.

Se elaboró un Script en Matlab que ejecuta cpp.exe para todos los ficheros .wav contenidos en una carpeta. Para ello se emplea la instrucción [a,b]=dos('cpp.exe params'). La variable de salida “a”, recibe cualquier explicación sobre la ocurrencia de un posible error, y en “b” la salida a la pantalla de la ejecución del programa, “params” es SP o VO como se declaró con anterioridad.

Después de ejecutar el programa para ‘CPP’, obtenemos los valores medios de Frecuencia Fundamental y CPP, en caso de emplear el ‘CPPs’, el programa también informa la media del CPPs (ver Anexo 1).

2.2.1.2 Jitter

Como medida de Jitter se empleó el *Jitter Percent* (Buder, 2000) definido como:

$$\frac{\sum_{i=2}^N \frac{|T_i - T_{i-1}|}{\frac{(T_i + T_{i-1})}{2}}}{N} \quad (2.1)$$

Donde T_i es la duración del pulso glotal i -ésimo y N es representa el número total de pulsos.

2.2.2 Orientadas a los dos factores

Luego del análisis bibliográfico desarrollado en el capítulo 1, quedaron sentadas las bases para la selección de las medidas que guardaban una relación teórica con la definición perceptual del jadeo, centrando el esfuerzo en decidir qué medir y no en la aplicación de varias alternativas. Estas medidas son el FHP y HNR, descritas a continuación.

Estimación F_0 : Un requisito previo para ambas medidas es la determinación de un estimado de frecuencia fundamental: $F_{0E} = 1/T_{0E}$, siendo T_{0E} la estimación de período fundamental. Para este propósito, se utiliza el detector clásico de tono AMDF (ROSS et al., 1974), la segmentación de la señal en tramas de 40 ms sin solapamiento, y la detección de la F_{0E} en el intervalo de 90 a 400 Hz. El T_{0E} para la grabación se elige como la mediana de los valores obtenidos para cada trama.

2.2.2.1 HNR

En el dominio espectral, existen medidas teóricamente incapaces de separar el ruido de otras aperiodicidades como los son jitter y shimmer como se planteó en el capítulo 1, por lo que se prefirió emplear una medida en el dominio del tiempo. En particular, se empleó el promedio de Yumoto (Yumoto et al., 1982), con la corrección de (Ferrer et al., 2005a) que permite el uso de tan pocos pulsos como se desee:

$$HNR = \frac{N-1}{N} HNR_{Yum} - \frac{1}{N} \quad (2.2)$$

En (2.2), N representa el número de impulsos en el conjunto, y HNR_{Yum} es la relación de armónicos-a-ruido. El uso de un menor número de pulsos reduce la influencia de la variabilidad en la estimación de la forma de onda de HNR. En voces patológicas, es común encontrar variaciones graduales de la forma de onda del pulso, los requisitos iniciales HNR_{Yum} son segmentos constantes de 30-50 pulsos, por lo que la limitación es un problema resuelto en (2.2). En este trabajo se empleó la menor cantidad posible de pulsos ($N=2$).

2.2.2.2 Prominencia del Primer Armónico

Como indicador de cierre glotal incompleto se decidió emplear una medida espectral de prominencia del primer armónico. Se descartó la variante de efectuar mediciones sobre la señal glotal, para evitar las complicaciones de obtener un adecuado filtrado inverso.

Como las medidas de relación de energía por bandas tienen mayor robustez que las orientadas a puntos individuales del espectro, se decidió calcular la prominencia del primer armónico incluyendo toda una banda de frecuencias, y no en la frecuencia del

punto de la FFT correspondiente a F_0 . La prominencia se calcula además con respecto a una banda de frecuencias y no con respecto a una frecuencia específica (típicamente ubicada en H2, F1 ó F2). La banda seleccionada se ubica teniendo en cuenta el ejercicio sobre el que se realiza la medición (vocal “a” sostenida), e intenta contener la energía correspondiente a ambos formantes F1 y F2. La medida se denota como FHP y se define como:

$$FHP = 10 * \log_{10} \left(\frac{F_0 R}{F_{12} R} + 1 \right) = 10 * \log_{10} \left(\frac{F_0 R + F_{12} R}{F_{12} R} \right) \quad (2.3)$$

donde $F_0 R$ y $F_{12} R$ son las energías en la región correspondiente al primer armónico, y el primer y segundo formantes, respectivamente. Las expresiones de ambas energías son:

$$F_0 R = \sum_{f=F_0-\Delta F_0/2}^{F_0+\Delta F_0/2} X(f) \quad (2.4)$$

y

$$F_{12} R = \sum_{f=F_c-\Delta F_c/2}^{F_c+\Delta F_c/2} X(f) \quad (2.5)$$

donde $X(f)$ es el espectro de potencia de la grabación, obtenida por medio de un periodograma de Welch, utilizando una ventana de Hanning en segmentos de 40 milisegundos, con 50% de superposición. F_0 es la T_{0E} resultante del detector de tono AMDF, y F_c es la frecuencia central del primer y segundo formante de la vocal (“a”), asumida de manera fija a 1KHz. Ambas sumas incluyen la energía en las bandas de frecuencia centradas en F_0 y F_c , es decir ΔF_0 y ΔF_c , establecido en 100 Hz y 1 kHz, respectivamente. En la Figura 2.3 se muestra una representación gráfica de estos parámetros que intervienen en el cálculo de la FHP.

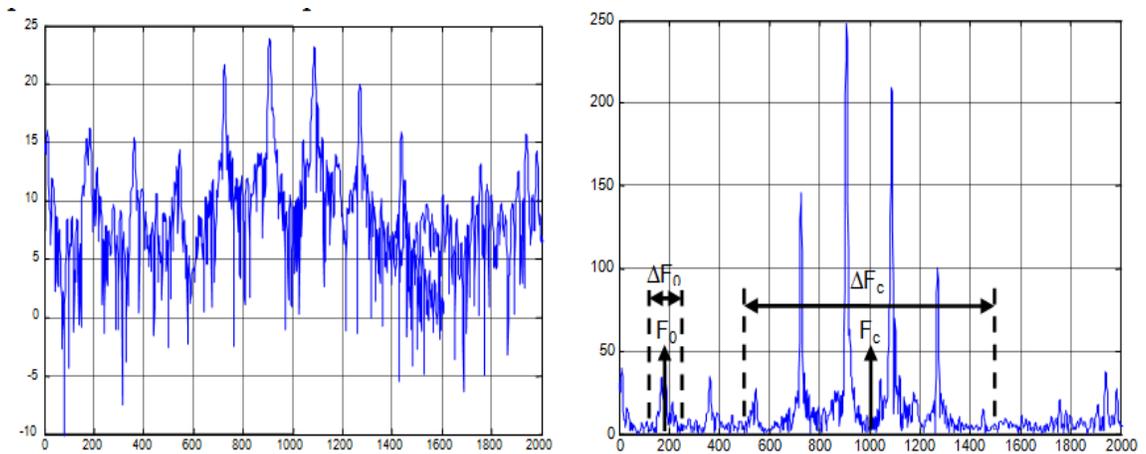


Figura 2.3 Espectro de amplitud (a la izquierda) y el espectro de amplitud lineal (derecha) de una "a" estándar con una representación de los parámetros que intervienen en el cálculo de la FHP.

2.2.2.3 Combinación

Aunque no es una medida en sí misma, la combinación de varias medidas también se ha estudiado como un indicador de jadeo ((Awan and Roy, 2005), (Castillo-Guerra and Ruiz, 2009), (Maryn et al., 2007a), (Stranik et al., 2014)). En este trabajo la combinación se realiza basada en el carácter multifactorial de la percepción del jadeo explicado en el epígrafe 1.5.2.1. En cuanto a la forma en que las medidas individuales se combinan, la regresión lineal y el análisis de componentes principales, son las técnicas más reportadas. En este caso empleamos la regresión lineal, en ausencia de un criterio mejor fundamentado de combinación.

2.3 Descripción del experimento

Se calcularon los valores de regresión para todas las medidas seleccionadas por separado y para todas sus combinaciones posibles, comparando la efectividad de la combinación de HNR y FHP con el resto de las medidas. Esto se efectuó para los dos grupos de valoraciones subjetivas, correspondientes a las Bases de Datos de Youry Maryn's y de la Kay Elemetrics.

2.4 Conclusiones parciales

Se describió el contenido de las Bases de Datos, y se justificó la selección de las medidas empleadas, así como la manera en que se evalúa la correspondencia Valoraciones Subjetivas – Mediciones Objetivas.

CAPÍTULO 3. RESULTADOS Y DISCUSIÓN

En este capítulo se muestran los resultados obtenidos tanto del análisis de las valoraciones subjetivas de las Bases de Datos como de la correspondencia de estas con las mediciones objetivas seleccionadas. Por último se comentan las implicaciones de los resultados.

3.1 Distribución de las Valoraciones Subjetivas en las Bases de Datos

Como primer criterio a emitir en la posible correspondencia entre las Medidas Objetivas y las Valoraciones Subjetivas, se efectúa un análisis de la distribución de estas últimas teniendo en cuenta los cuestiones propuestas en la sección 1.5.3

3.1.1 Youry Maryn's

3.1.1.1 Distribución

En aras de evaluar la utilidad de la Base de Datos en cuanto a su capacidad para valorar las características deseadas de las medidas (sensible y específica), se analizó la distribución de las valoraciones por los distintos rasgos (Grado, Aspereza y Jadeo).

En la (Figura 3.1) se muestra la distribución de los valores medios de los cinco jueces para Grado, Aspereza y Jadeo de la Base de Datos de Youry Maryn's.

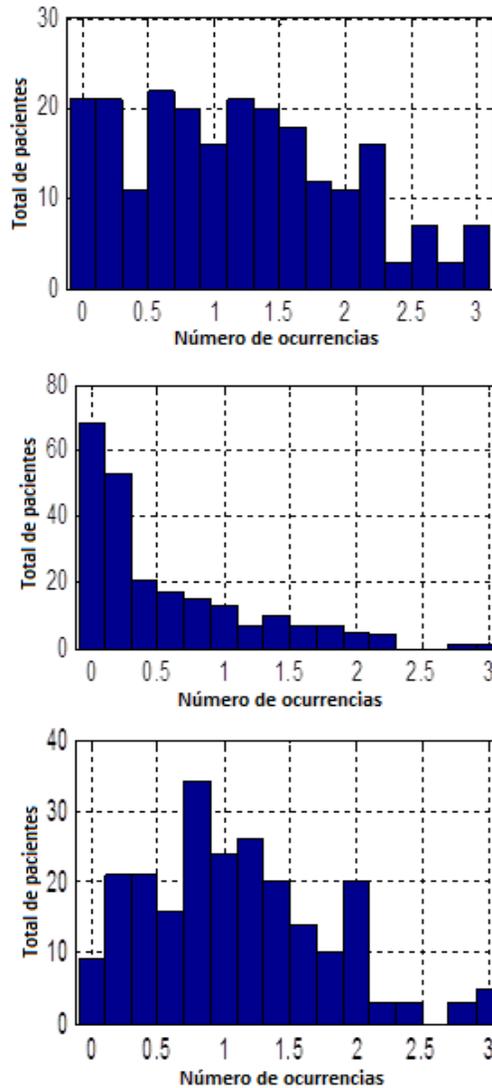


Figura 3.1 Histogramas de ocurrencias de las medias de los cinco jueces en las 229 grabaciones. De arriba a abajo: Grado, Aspereza y Jadeo.

En esta figura se puede observar que la aspereza no se distribuye uniformemente a lo largo de la escala, los valores altos de ocurrencia son escasos, se muestra cargada a la parte baja de las valoraciones, es decir se muestra muy concentrada entre 0 y 1, mientras que el grado y el jadeo se distribuyen de manera más uniforme. Se puede afirmar entonces que esta Base de Datos es más adecuada para medir jadeo, que para evaluar aspereza. Por otra parte, en cuanto al jadeo puede observarse que, incluso sin sujetos sanos hay un número de sujetos con valores de jadeo 0, y la región más poblada de la escala se encuentra alrededor de $B=1$. En el extremo superior de la escala la población con jadeo es menor. Teniendo en cuenta esta distribución, es de

esperar que las correlaciones se vean afectadas por el aumento de la varianza de las valoraciones subjetivas en la región media de la escala (Kreiman and Gerrat, 2000).

3.1.1.2 Correlación entre jueces

Las correlaciones entre las medias de los jueces en las tres dimensiones son 0,55 para G-R, 0,81 en el caso de G-B y 0,13 en R-B, reafirmando el hecho de que en esta base de datos el jadeo es más fuerte que la aspereza, pues influye más en la valoración general de G.

Los valores medios de correlación entre jueces son de 73% para Grado, 63% para la Aspereza y 63% para el Jadeo considerados valores moderados por (Dancey and Reidy, 2004).

Tabla 3.1 Correlación entre los 5 jueces para los tres rasgos.

	Grado				Aspereza				Jadeo			
	J2	J3	J4	J5	J2	J3	J4	J5	J2	J3	J4	J5
J1	0.76	0.68	0.77	0.66	0.74	0.59	0.65	0.71	0.67	0.56	0.72	0.53
J2	-	0.78	0.79	0.79	-	0.68	0.62	0.69	-	0.64	0.75	0.58
J3	-	-	0.68	0.73	-	-	0.52	0.56	-	-	0.64	0.60
J4	-	-	-	0.70	-	-	-	0.56	-	-	-	0.62

3.1.2 Kay Elemetrics

3.1.2.1 Distribución

En el caso de la Base de Datos Kay Elemetrics, la distribución de los valores medios de las seis valoraciones para Grado, Aspereza y Jadeo se muestra en la Figura 3.2.

Con respecto a la distribución de las valoraciones en esta Base de Datos, sobresale el gran número de pacientes con una valoración de 3 en Grado a pesar de no corresponderse esto con las distribuciones en Aspereza y Jadeo, donde la barra correspondiente a la máxima valoración se queda por debajo de 50 pacientes en ambos casos. Contrariamente sucede para los valores 0, en Grado este total queda por debajo de los 10 pacientes prácticamente.

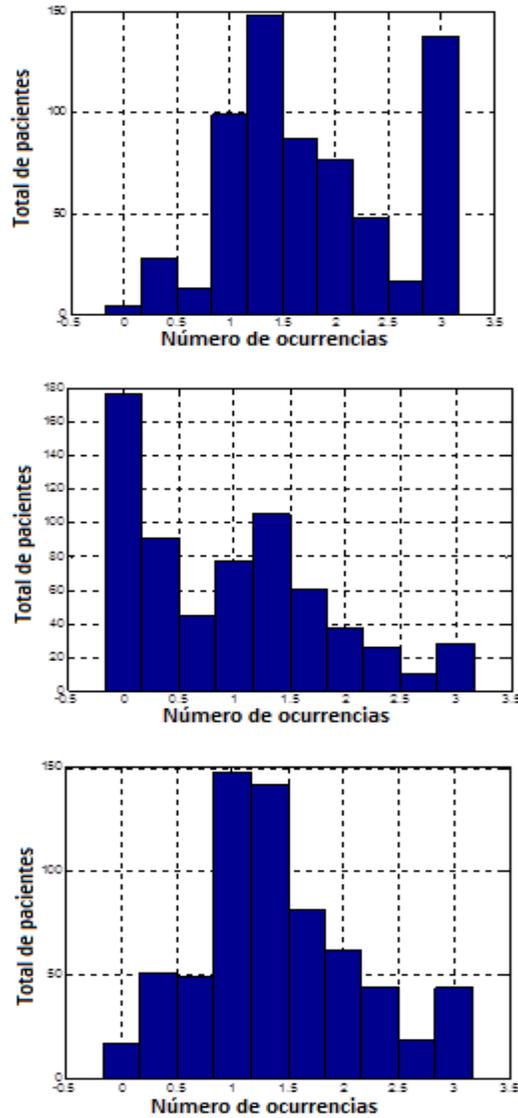


Figura 3.2 Histogramas de ocurrencias de las medias de Grado, Aspereza y Jadeo en ese orden de arriba hacia abajo para las 657 grabaciones.

Para la Aspereza, puede observarse que existe un gran número de sujetos con valoración con 0, esta distribución de valoraciones se recarga a los valores menores de la escala, de manera semejante a lo ocurrido en la Base de Datos de Youry Maryn's, aunque no tan notorio.

En el caso particular del Jadeo, existe una concentración en el centro de la escala, valores entre 1 y 2.

3.1.2.2 Correlación entre jueces

Tabla 3.2 Correlación entre los 3 jueces incluyendo la primera y segunda valoración para los tres rasgos.

	Grado					Aspereza					Jadeo				
	J2	J3	J1*	J2*	J3*	J2	J3	J1*	J2*	J3*	J2	J3	J1*	J2*	J3*
J1	0.79	0.80	0.76	0.73	0.74	0.79	0.77	0.73	0.73	0.74	0.68	0.73	0.67	0.63	0.66
J2	-	0.77	0.73	0.75	0.71	-	0.77	0.73	0.75	0.71	-	0.67	0.63	0.69	0.63
J3	-	-	0.69	0.69	0.73	-	-	0.70	0.69	0.77	-	-	0.68	0.65	0.69
J1*	-	-	-	0.73	0.73	-	-	-	0.73	0.77	-	-	-	0.66	0.75
J2*	-	-	-	-	0.71	-	-	-	-	0.71	-	-	-	-	0.64

El promedio de las correlaciones Intra-juez para Grado fue de 75%, mientras el Inter-juez fue de 74%. En el caso del Jadeo el promedio de las correlaciones Intra-juez fue de 75%, y el Inter-juez 74%. Para la Aspereza, el Intra-juez 69% y el Inter-juez 67%. En los tres casos la variabilidad de las valoraciones entre los jueces es comparable a la variabilidad con respecto a los demás, por lo que se puede afirmar que poseían un estándar semejante de las definiciones de los rasgos.

3.2 Regresiones de las medidas con las Valoraciones G, R, B

Los resultados de las regresiones lineales de las medidas y sus combinaciones en función de los parámetros G, R y B se muestran a continuación para la Base de Datos. Las regresiones se establecieron mediante la función *Regress* del software Matlab.

$[r1, r2, r3, r4, stats] = \mathbf{regress}(y,x)$

Donde en “y” se define el promedio de las valoraciones subjetivas de todos los jueces para cada paciente en un rasgo (G, R ó B) y “x” es una matriz de mediciones objetivas donde cada columna es una medición y cada fila un paciente. En “stats” se obtiene el porcentaje de varianza de “y” explicado por el modelo a partir de “x”.

3.2.1 Youry Maryn’s

A continuación se muestran los resultados para la Base de Datos de Youry Maryn’s. Las medidas se han abreviado como “S” el CPPs, “C” el CPP, “H” la medida de HNR, “J” Jitter y “F” corresponde a “FHP”.

Tabla 3.3 Resultado de las regresiones lineales para Jadeo en Youry Maryn's

	Individual	Parejas	Mejora (%)	Tríos	Todas
1^{er}	S (0.54)	(C+H) (0.58)	8.6	(C+H+S) (0.59)	(0.60)
2^{do}	C (0.54)	(C+J) (0.58)	7.5	(C+H+J) (0.59)	
3^{ro}	H (0.43)	(S+C) (0.57)	6.2	(C+H+F) (0.59)	
4^{to}	J (0.25)	(S+H) (0.56)	2.6	(S+H+F) (0.58)	
5^{to}	F (0.04)	(H+F) (0.55)	28.9	(C+J+S) (0.58)	
6^{to}		(C+F) (0.55)	1.9	(S+C+F) (0.58)	
7^{mo}		(S+F) (0.55)	0.73	(C+J+F) (0.58)	
8^{vo}		(S+J) (0.54)	0.02	(H+S+J) (0.56)	
9^{no}		(H+J) (0.43)	0.6	(H+F+J) (0.56)	
10^{mo}		(J+F) (0.35)	29.1	(S+F+J) (0.55)	

En la tabla 3.3 puede apreciarse que hay una coincidencia con lo reportado en la literatura. CPP y CPPs son los mejores predictores individuales de Jadeo. Sin embargo la unión de HNR y FHP los supera ligeramente (0.55 por 0.54) teniendo esta última unión una ganancia de un 28.9% en cuanto a varianza explicada con respecto a sus componentes individuales. Se puede afirmar que H+F tiene una sensibilidad tan buena o mejor que las medidas cepstrales individuales, pero con mejor correspondencia teórica, como fue explicada en capítulos anteriores.

Tabla 3.4 Resultado de las regresiones lineales para Aspreza en Youry Maryn's

	Individual	Parejas	Mejora (%)	Tríos	Todas
1^{er}	H (0.13)	(C+F) (0.16)	220	(C+F+H) (0.17)	(0.17)
2^{do}	J (0.08)	(S+F) (0.15)	87.5	(C+F+J) (0.16)	
3^{ro}	S (0.08)	(H+F) (0.15)	15.38	(C+F+S) (0.16)	
4^{to}	C (0.05)	(H+J) (0.13)	1.45	(S+F+H) (0.16)	
5^{to}	F (0.05)	(H+C) (0.13)	0.54	(S+F+J) (0.15)	
6^{to}		(H+S) (0.13)	0	(H+F+J) (0.15)	
7^{mo}		(J+F) (0.11)	28.78	(H+J+C) (0.13)	
8^{vo}		(J+S) (0.10)	16.55	(H+J+S) (0.13)	
9^{no}		(J+C) (0.10)	14.39	(H+C+S) (0.13)	
10^{mo}		(S+C) (0.08)	2.15	(J+S+C) (0.10)	

Para la Aspreza no se obtienen resultados relevantes, ni siquiera la unión de todas las medidas logra explicar más allá de 1/6 de la varianza de la misma. No obstante es

destacable la unión de CPP y FHP, que logra una mejoría notable sobre cualquiera de ellos, de manera individual.

Tabla 3.5 Resultado de las regresiones lineales para Grado en Youry Maryn's

	Individual	Parejas	Mejora (%)	Tríos	Todas
1^{er}	S (0.43)	(S+H) (0.47)	10	(S+F+C) (0.48)	(0.49)
2^{do}	H (0.41)	(H+C) (0.46)	10.91	(C+F+J) (0.47)	
3^{ro}	C (0.35)	(S+F) (0.46)	6.85	(S+H+F) (0.47)	
4^{to}	J (0.29)	(S+J) (0.44)	80.96	(H+C+J) (0.47)	
5^{to}	F (0.002)	(C+F) (0.44)	27.23	(S+H+J) (0.47)	
6^{to}		(C+J) (0.44)	27.18	(H+C+S) (0.47)	
7^{mo}		(S+C) (0.43)	0.38	(H+C+F) (0.47)	
8^{vo}		(H+J) (0.42)	2.95	(S+F+J) (0.46)	
9^{no}		(H+F) (0.42)	2.26	(S+J+C) (0.45)	
10^{mo}		(J+F) (0.29)	1.92	(H+J+F) (0.43)	

Casi la mitad de la varianza del Grado puede explicarse a partir de las cinco medidas, sin embargo la unión de H+F sólo mejora en un 2% a la individual.

3.2.2 Kay Elemetrics

Tabla 3.6 Resultado de las regresiones lineales para Jadeo en Kay Elemetrics

	Individual	Parejas	Mejora (%)	Tríos	Todas
1^{er}	C (0.66)	(C+H) (0.70)	6.65	(C+H+F) (0.73)	(0.73)
2^{do}	S (0.61)	(H+F) (0.69)	21.94	(H+F+S) (0.71)	
3^{ro}	H (0.56)	(C+S) (0.68)	0.70	(C+H+S) (0.70)	
4^{to}	J (0.24)	(C+J) (0.67)	2.25	(C+H+J) (0.70)	
5^{to}	F (0.07)	(C+F) (0.66)	0.05	(H+F+J) (0.69)	
6^{to}		(S+H) (0.64)	6.31	(C+S+J) (0.68)	
7^{mo}		(S+F) (0.62)	2.91	(C+S+F) (0.68)	
8^{vo}		(S+J) (0.61)	0.56	(C+J+F) (0.68)	
9^{no}		(H+J) (0.56)	0	(S+H+J) (0.65)	
10^{mo}		(J+F) (0.36)	49.09	(S+F+J) (0.62)	

En esta base de datos los resultados para el Jadeo en general, son superiores a los anteriores. Las medidas H+F son más sensibles que las cepstrales solas, además de mejorar en un 20% de la varianza explicada a sus componentes individuales.

Tabla 3.7 Resultado de las regresiones lineales para Aspereza en Kay Elemetrics

	Individual	Parejas	Mejora (%)	Tríos	Todas
1^{er}	H (0.10)	(S+F) (0.14)	55.57	(S+F+C) (0.15)	(0.16)
2^{do}	S (0.09)	(H+F) (0.12)	17.96	(S+J+F) (0.15)	
3^{ro}	J (0.05)	(S+C) (0.12)	28.06	(S+F+H) (0.14)	
4^{to}	C (0.04)	(H+C) (0.11)	5.79	(S+C+H) (0.14)	
5^{to}	F (0.03)	(H+S) (0.11)	5	(H+F+J) (0.12)	
6^{to}		(H+J) (0.10)	0.89	(H+F+C) (0.12)	
7^{mo}		(C+F) (0.09)	159.6	(S+C+J) (0.12)	
8^{vo}		(S+J) (0.09)	1.09	(H+C+J) (0.11)	
9^{no}		(J+F) (0.07)	35.15	(H+S+J) (0.11)	
10^{mo}		(J+C) (0.06)	16.17	(C+F+J) (0.10)	

Con respecto a la aspereza, ocurre una situación análoga a la anterior, no se obtiene ningún resultado relevante, ni siquiera la combinación de todas las medidas, lo que no explica más allá de 1/6 de la varianza de la misma. No obstante se destaca la unión de CPP y FHP, que logra una mejoría notable sobre cualquiera de las medidas de manera individual

Tabla 3.8 Resultado de las regresiones lineales para Grado en Kay Elemetrics

	Individual	Parejas	Mejora (%)	Tríos	Todas
1^{er}	S (0.63)	(S+H) (0.67)	6.41	(S+H+J) (0.67)	(0.68)
2^{do}	H (0.59)	(S+F) (0.63)	0.51	(S+H+F) (0.67)	
3^{ro}	C (0.50)	(S+C) (0.63)	0.03	(S+H+C) (0.67)	
4^{to}	J (0.29)	(S+J) (0.63)	0	(S+F+C) (0.64)	
5^{to}	F (0.006)	(H+C) (0.63)	4.78	(S+F+J) (0.64)	
6^{to}		(H+F) (0.62)	4.78	(S+C+J) (0.63)	
7^{mo}		(H+J) (0.59)	0.54	(H+C+J) (0.63)	
8^{vo}		(C+J) (0.55)	11	(H+C+F) (0.63)	
9^{no}		(C+F) (0.52)	4.18	(H+F+J) (0.62)	
10^{mo}		(J+F) (0.32)	9.32	(C+J+F) (0.56)	

En la tabla de resultados de Grado, los valores en general también son superiores a los alcanzados en la anterior base de datos. La unión de H+F no supera al mejor resultado de las medidas cepstrales, y sólo mejora en un 5% con respecto a sus componentes individuales.

3.3 Conclusiones del capítulo

Los resultados obtenidos sustentan el empleo de la combinación lineal de HNR y FHP como indicadores de jadeo. Esta combinación resultó ser más sensible al jadeo que las mejores medidas individuales reportadas en la literatura (CPP y CPPs). En cuanto a la cualidad deseable de ser específica, no pudo emitirse un criterio en el rasgo de aspereza, con valores de predicción muy pequeños, pero sí en cuanto a grado, donde quedó por debajo de CPP y/o CPPs. Como evidencia de consistencia, puede mencionarse la obtención de resultados semejantes para dos amplias bases de datos diferentes.

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

Este trabajo ha buscado solución para la ausencia de medidas objetivas de jadeo que se correspondan con las características percibidas. Estos aportes se centran alrededor de los objetivos específicos planteados en la introducción. Como conclusión:

- Se determinó como características distintivas del jadeo la presencia de ruido de aspiración y el cierre incompleto de la glotis.
- Se propuso un algoritmo de medición que se corresponde teóricamente con la percepción del jadeo, al combinar las medidas acústicas, cada una de ellas relacionada con uno de los rasgos distintivos determinados.
- El algoritmo propuesto se evaluó para dos Bases de Datos con un conjunto de sujetos valorados subjetivamente y se correlacionó satisfactoriamente con estas valoraciones, mostrando ser sensible, específico y consistente, las tres características deseables en una medida objetiva

Recomendaciones

A partir de los resultados descritos en este proyecto, queda abierta la posibilidad de reevaluar la correspondencia entre las medidas acústicas obtenidas, empleando los métodos aquí propuestos, y las valoraciones subjetivas de diferentes rasgos con valor diagnóstico.

Como sugerencias inmediatas para trabajos futuros se proponen:

-
- Probar con otras maneras de medir los dos factores propuestos, estableciendo variantes para el HNR y el FHP.
 - Realizar un trabajo semejante para la aspereza. En este trabajo no se encontró una correlación aceptable de ninguna medida o combinación de estas con la percepción de la aspereza.

REFERENCIAS BIBLIOGRÁFICAS

- ALKU, P. & VILKMAN, E. 1996a. Amplitude domain quotient for characterization of the glottal volumetric velocity waveform estimated by inverse filtering. *Speech Commun.*, 18, 131-138.
- ALKU, P. & VILKMAN, E. 1996b. A comparison of glottal voice source quantification parameters in breathy, normal and pressed phonation of female and male speakers. *Folia Phoniatr. Logop.*, 48, 240-254.
- ALPANA, A., SCHOENTGENA, J., MARYNC, Y., GRENEZA, F. & MURPHYD, P. 2012. Assessment of disordered voice via the first harmonic. *Speech Communication*, 54, 655–663.
- ARENDS, N., POVEL, D., OS, E. V. & SPETH, L. 1990. Predicting voice quality of deaf speakers on the basis of glottal characteristics. *Journal of Speech and Hearing Research*, 33, 116-122.
- ASKENFELT, A. G. & HAMMARBERG, B. 1986. Speech wave form perturbation analysis: A perceptual-acoustical comparison of seven measures. *Journal of Speech and Hearing Research*, 29, 50-64.
- AWAN, S. N. & ROY, N. 2005. Acoustic prediction of voice type in women with functional dysphonia. *J Voice*, 268–282.
- BAKEN, R. J. 1999. *Clinical Measurement of Speech and Voice*, San Diego.
- BASSICH, C. & LUDLOW, C. L. 1986. The use of perceptual methods by new clinicians for assessing voice quality. *Journal of Speech and Hearing Disorders*, 51, 125-133.
- BERNAL, J., BOBADILLA, J. & GOMEZ, P. 2000. *Reconocimiento de Voz y Fonética Acústica*, Madrid, RA-MA Editorial.
- BHUTA, T., PATRICK, L. & GARNETT, J. D. 2004. Perceptual Evaluation of Voice Quality and its Correlation with Acoustic Measurements. *Journal of Voice*, 18, 299-304.
- BICKLEY, C. 1982. Acoustic Analysis and Perception of Breathless Vowels. *Speech Commun.*, 1, 71-82.
- BUDER, E. H. 2000 Acoustic analysis of vocal quality: a tabulation of algorithms 1902-1990. In: KENT, R. D. & BALL, M. J. (eds.) *Voice Quality Measurement*. San Diego: Singular.

- CASTILLO-GUERRA, E. & RUIZ, A. 2009. Automatic modeling of acoustic perception of breathiness in pathological voices. *IEEE Transactions on Biomedical Engineering*.
- CHILDERS, D. & LEE, C. 1991. Vocal quality factors: Analysis, synthesis, and perception. *J. Acoust. Soc. Am*, 90, 2394–2410.
- DANCEY, C. P. & REIDY, J. 2004. *Statistics Without Maths for Psychology: Using SPSS for Windows*, Prentice Hall.
- DARLEY, F. L., ARONSON, A. E. & BROWN, J. R. 1969a. Differential diagnostic patterns of dysarthria. *Journal of Speech & Hearing Research*, 12, 246-269.
- DEJONCKERE, P. H., OBBENS, C., MOOR, G. M. D. & WIENEKE, G. H. 1993. Perceptual evaluation of dysphonia: reliability and relevance. *Folia Phoniatrica*, 45, 76-83.
- ENDERBY, P. 1983. *Frenchay Dysarthria Assessment.*, California, College Hill.
- ESKENAZI, L., CHILDERS, D. G. & HICKS, D. M. 1990. Acoustic correlates of vocal quality *Journal of Speech and Hearing Research*, 33, 298-306.
- FAIRBANKS, G. 1940. *Voice and articulation drillbook*, New York, Harper and Brothers.
- FANT, G. 1960. *Acoustic Theory of Speech Production*, The Hague, Mouton
- FEIJOO, S. & HERNÁNDEZ, C. 1990. Short-term stability measures for the evaluation of vocal quality. *Journal of Speech and Hearing Research*, 33.
- FERRER, C., BODT, M. S. D., MARYN, Y., HEYNING, P. V. D. & DÍAZ, M. E. H. Year. Properties of the cepstral peak prominence and its usefulness in vocal quality measurements. *In: 5th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications, Maveba 2007*, 2007 Firenze, Italia.
- FERRER, C. A., GONZÁLEZ, E. & HERNÁNDEZ-DÍAZ, M. E. 2005a. Correcting the use of ensemble averages in the calculation of harmonics to noise ratios in voice signals. *J. Acoust. Soc. Am*, 118, 605-608.
- FERRER, C. A. & HERNÁNDEZ-DÍAZ, M. E. 2001. Obtención de un índice objetivo de razón lenta. *Memorias del VII Simposio de Comunicación Social Santiago 2001*. Santiago de Cuba: Editorial Oriente.
- FERRER, C. A., TORO, A. D., GONZÁLEZ, E. & HERNÁNDEZ-DÍAZ, M. E. 2005b. Using Neural Networks in the estimation of consonant imprecision ratings. *Research in Computer Science*, 15, 179-186.
- FRITZELL, B., HAMMARBERG, B., GAUFFIN, J., KARLSSON, I. & SUNDBERG, J. 1986. Breathiness and insufficient vocal fold closure. *Journal of Phonetics*, 14, 549-553.
- FROHLICH, M., MICHAELIS, D. & STRUBE, H. W. 1998. Acoustic “Breathiness measures” in the description of pathologic voices. *IEEE*, 937-940.
- FUKAZAWA, T., EI-ASSUOOTY, A. & HONJO, I. 1988. A new index for evaluation of the turbulent noise in pathological voice. *Journal of the Acoustical Society of America*, 83, 1189-1193.

- GELFER, M. P. 1988. Perceptual attributes of voice: Development and use of rating scales. *Journal of Voice.*, 2, 320-326.
- GERRATT, B. R. & KREIRNAN, J. 1995. The utility of acoustic voice measures. . In: D.WONG (ed.) *Proceedings of the workshop on standardization in acoustic voice analysis*. Denver: National Center for Voice and Speech.
- GODINO-LLORENTE, J. I., OSMA-RUIZ, V., SAENZ-LECHON, N., GOMEZ-VILDA, P., BLANCO-VELASCO, M. & CRUZ-ROLDAN, F. 2010. The effectiveness of the glottal-to-noise excitation ratio for the screening of voice disorders. *J Voice*, 24, 47-56.
- HAMMARBERG, B. & GAUFFIN, J. 1995. *Perceptual and acoustic characteristics of quality differences in pathological voices as related to physiological aspects.*, San Diego, Singular.
- HEMAN-ACKAH, Y. D., MICHAEL, D. D. & GODING, G. S. 2002. The Relationship Between Cepstral Peak Prominence and Selected Parameters of Dysphonia. *Journal of Voice*, 16, No. 1, 20-27.
- HERMANN, L. 1890. Ueber das Verhalten der Vocale am neuen Edison'schen Phonographen. *Archiv für die gesamte Physiologie des Menschen und der Tiere*, 47, 42-44.
- HILLENBRAND, J. accessed May 21, 2012. cps.exe (software). <http://homepages.wmich.edu/~hillenbr>
- HILLENBRAND, J., CLEVELAND, R. A. & ERICKSON, R. L. 1994a. Acoustic correlates of breathy vocal quality *Journal of Speech and Hearing Research*, 37, 769-778.
- HILLENBRAND, J., CLEVELAND, R. A. & ERICKSON, R. L. 1994b. Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research*, 37, 769-778.
- HILLENBRAND, J. & HOUDE, R. A. 1996. Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *J. Speech Hea. . Res.*, 39, 311-321.
- HIRANO, M. 1981. Clinical examination of voice. *New York: Springer-Verlag*.
- KAY-ELEMENTRICS 1992. Voice disorders database. In: INFIRMARY, V. A. S. L. M. E. A. E. (ed.).
- KENT, R. D. 1996. Hearing and Believing: Some Limits to the Auditory-Perceptual Assessment of Speech and Voice Disorders. *American Journal of Speech-Language Pathology*, 5.
- KENT, R. D. 2000. Research on speech motor control and its disorders: a review and prospective. *Journal of Communication Disorders.*, 33, 391-428.
- KLATT, D. H. & KLATT, L. C. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, 87, 820-857.

- KLICH, R. 1982. Relationships of vowel characteristics to listener ratings of breathiness *Journal of Speech and Hearing Research*, 25, 574-580.
- KREIMAN, J. & GERRAT, B. 2000. Measuring Voice Quality. In: KENT, R. D. & BALL, M. J. (eds.) *Voice Quality Measurement*. San Diego: Singular.
- KREIMAN, J., GERRAT, B. R., KEMPSTER, G. B., ERMAN, A. & BERKE, G. S. 1993. Perceptual evaluation of voice quality: Review, Tutorial, and a Framework for future research. *JSHR*, 36, 21-40.
- KREIMAN, J. & GERRATT, B. R. 1996. The perceptual structure of pathologic voice quality. *J Acoust Soc Am.*, 100, 1787-1795.
- KREIMAN, L., GERRATT, B. R. & PRECODA, K. 1990. Listener experience and perception of voice quality. *Journal of Speech and Hearing Research*, 33, 103-115.
- KROM, G. A. D. 1993. Cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals. *J. Speech Hear. Res.*, 36, 254-266.
- KROM, G. D. 1994. Spectral Correlates of breathiness and roughness for different types of vowel fragments. *Research Institute for Language and Speech*. 3512 JK Utrecht, the Netherlands: University of Utrecht Trans 10.
- KROM, G. D. 1995. Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. *Journal of Speech and Hearing Research*, 38, 794-811.
- LADEFOGED, P. 1983. The Linguistic Used of Different Phonation Types. In: BLESS, D. & ABBS, J. (eds.) *Vocal Fold Physiology: Contemporary Research and Clinical Issues*. College Hill, San Diego.
- LIEBERMAN, P. 1963. Some acoustic measures of fundamental periodicity of normal and pathological larynges. *JASA*, 35, 597-603.
- MARTIN, D., FITCH, L. & WOLFE, V. 1995. Pathologic voice type and the acoustic prediction of severity. *Journal of Speech and Hearing Research*, 38, 765-771.
- MARYN, Y., CORTHALS, P., BODT, M. D. & CAUWENBERGHE, P. V. 2007a. Cepstral peak prominence as a measure for overall voice quality in vowel as well as in continuous speech segments. *7th Pan European Voice Conference PEVOC'07*.
- MARYN, Y., CORTHALS, P., BODT, M. D. & CAUWENBERGHE, P. V. 2007b. Cepstral peak prominence as a measure for overall voice quality in vowel as well as in continuous speech segments. *7th Pan European Voice Conference PEVOC'07*.
- PROSEK, R. A., MONTGOMERY, A. A., WALDEN, B. E. & HAWKINS, D. B. 1987. An evaluation of residue features as correlates of voice disorders. *Journal of Communication Disorder*, 20 105-117.
- ROSS, M. J., SHAFFER, H. L., COHEN, P., FREUDBERG, R. & MANLEY, H. J. 1974. Average magnitude difference function pitch extractor. *IEEE Transaction on Acoustic Speech Signal Processing*, 5, 353-362.
- SCHOENTGEN, J. 2006. Vocal Cues of Disordered Voiced. *Acta Acustica united with Acustica*, 92, 667.

- SHAHEEN, N., AWAN & NELSON, R. 2005. Bloomsburg, Pennsylvania, and Salt Lake City, Utah. *Journal of Voice*, 19 No. 2, 268–282.
- SHRIVASTAV, R. & SAPIENZA, C. M. 2003. Objective measures of breathy voice quality obtained using an auditory model. *J. Acoust. Soc. Am.*, 114, 2217- 2224.
- SODERSTEN, M., HERTEGARD, S. & HAMMARBERG, B. 1995. Glottal closure, transglottal air flow, and voice quality in healthy middle-aged women. *Journal of Voice*, 9, 182-197.
- SODERSTEN, M. & LINDESTAD, P. 1990. Glottal closure and perceived breathiness during phonation in normally speaking subjects. *Journal of Speech and Hearing Research*, 33, 601-611.
- SPEYER, R., WIENEKE, G. H. & DEJONCKERE, P. H. 2003. Documentation of Progress in Voice Therapy: Perceptual, Acoustic, and Laryngostroboscopic Findings Pretherapy and Posttherapy. *Journal of Voice*, 18, 325–340.
- STEPHEN, J. & LOWIT, A. 2014. Cepstral Analysis of Hypokinetic and Ataxic Voices: Correlations With Perceptual and Other Acoustic Measures. *Journal of Voice*.
- STRANIK, A., CMEJILA, R. & VOKRAL, J. 2014. Acoustic Parameters for Classification of Breathiness in Continuous Speech According to the GRBAS Scale. *Journal of Voice*.
- TITZE, R. 1995. Workshop on acoustic voice analysis: Summary statement. *CO: National Center for Voice and Speech*. Denver.
- YUMOTO, E., GOULD, W. J. & BAER, T. 1982. The harmonic-to-noise ratio as an index of the degree of hoarseness. *J. Acoust. Soc. Am.*, 71, 1544-1550.
- ZWIRNER, P., MURRY, T. & WOODSON, G. E. 1993. Perceptual acoustic relationships in spasmodic dysphonia. *Journal of Voice*, 7, 165-171.

ANEXOS

Anexo I Programa de obtención de CPP y CPPs

```
function [cpp]=CppOfADir(vowel,detect);
%'vowel' chooses if default configs for vowel/RS are used, 'detect' to
%use/skip the voiced/unvoiced decision within the signal. cpp is a two
%column vector, first column is raw, second is smoothed.
[fname pname]=uigetfile('*.wav','Select a file within target directory');
dirlist=dir(pname);
cpp=[];
for k=3:length(dirlist),
    [p,n,e]=fileparts(dirlist(k).name);
    n=str2num(n);
    if detect,
        if vowel
            [a,b]=dos(['cpps.exe ' [pname dirlist(k).name] ' vo']);
        else
            [a,b]=dos(['cpps.exe ' [pname dirlist(k).name] ' sp']);
        end
    else
        if vowel
            [a,b]=dos(['cpps.exe ' [pname dirlist(k).name] ' vowel
allvoiced']);
        else
            [a,b]=dos(['cpps.exe ' [pname dirlist(k).name] ' ']);
        end
    end
    %aquí ya en b hay una cadena enorme con los resultados...
    x=regexp(b,'CPP\w*');%esto busca los inicios de CPP y CPPs
    y=regexp(b,'\n');%esto los cambios de linea
    cppi=find(y>x(1));
    cppsi=find(y>x(2));
    cpp(n,1)=str2num(b( (y(cppi(1))-6) : (y(cppi(1))-1) ));
    cpp(n,2)=str2num(b( (y(cpsi(1))-8) : (y(cpsi(1))-3) ));
end;
```