

Seminario Internacional de Acústica 2003

Aplicaciones del análisis acústico en los estudios de la voz humana

Matías Zañartu Salas
Unidad de Acústica - Escuela de Fonoaudiología
Universidad Mayor
Av. Libertador Bdo. O'Higgins 2013 - Santiago - Chile
Fono: (56) 2 4205715 - Fax: (56) 2 4205775
mzanartu@email.umayor.cl

Resumen

El análisis espectral tiene una serie de aplicaciones relativas al procesamiento de los sonidos de la voz humana, las cuales se presentan en softwares dedicados que incluyen una serie de herramientas que han sido especialmente diseñadas para facilitar el análisis e interpretación de los resultados. En este trabajo se exponen las principales aplicaciones en fonética acústica y en identificación de patologías de la voz realizadas con estas herramientas, que actualmente se desarrollan en el Laboratorio de Acústica de la Escuela de Fonoaudiología de la Universidad Mayor, las cuales son llevadas a cabo por un equipo multidisciplinario de especialistas en el tema, conformado por Ingenieros Acústicos, Fonoaudiólogos especialistas tanto en Lingüística, Audiología como Trastornos de la voz, y Médicos otorrinolaringólogos, entre otros. El trabajo presenta además una breve reseña de los criterios acústicos asociados al estudio de la voz humana, de los modelos de voz normal y patológica actualmente aceptados internacionalmente, de las variables fonéticas de mayor interés, y de las características básicas del funcionamiento de los programas de análisis dedicados para el estudio de la voz de la familia de softwares de Kay Elemetrics Corp.

1. Introducción

1.1. Conceptos básicos sobre la generación de la voz humana

Tradicionalmente necesitamos de tres grupos o sistemas de órganos para que, coordinadamente podamos producir lenguaje hablado. Estos grupos son: Sistema Respiratorio, Sistema Fonatorio y Sistema Articulatorio. Mediante estos sistemas se puede definir un modelo acústico de generación de voz humana, el cual considera las siguientes etapas:

- Generación del sonido madre: Mediante vibración de las cuerdas vocales (se distinguen dentro de esta categoría al menos tres patrones vibratorios diferentes), mediante flujo de aire y mediante golpes de presión acumulada.
- Modulación del tracto vocal: Se puede considerar al tracto vocal como tubo de sección variable, controlado según la articulación determinada por los músculos de la faringe, el cuerpo y punta de la lengua, el velo del paladar y los labios. Al igual que en un tubo simple, el fenómeno de resonancias en el tracto aumentan la amplitud de un grupo de frecuencias alrededor de una determinada banda de frecuencia. A cada resonancia se le denomina formante.
- Radiación de salida: La salida del sonido tiene asociada una impedancia de radiación que influye sobre el sonido variando la composición espectral del sonido y los niveles de presión sonora con respecto al ángulo de salida.

Figura Nº1. Modelo acústico básico de la generación de la voz

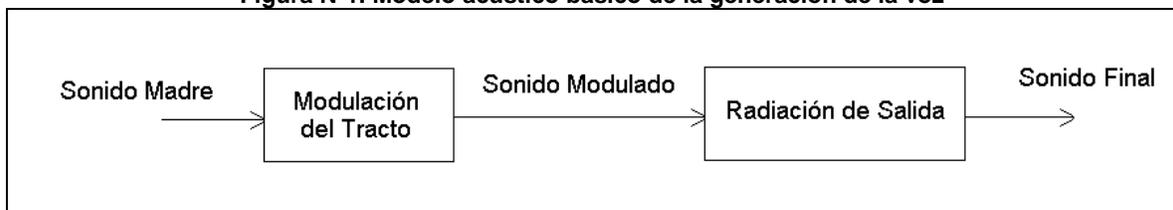
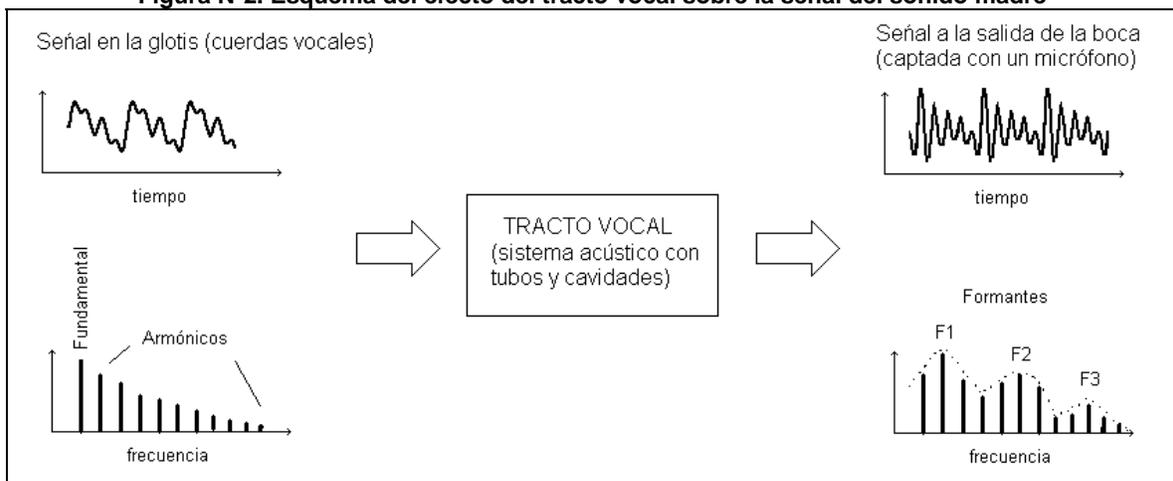


Figura Nº2. Esquema del efecto del tracto vocal sobre la señal del sonido madre



1.2. Fonemas: Vocales y Consonantes

El fonema es la mínima unidad lingüística capaz de producir cambios de significados. Estos fonemas pueden ser agrupados en dos grandes entidades: Fonemas Vocálicos y Fonemas Consonánticos. Ambos tipos de fonemas pueden clasificarse según el estado vibratorio de las cuerdas vocales, en fonemas fonados (ó sonoros) y áfonos (ó sordos). Los Fonemas Vocálicos son siempre fonados, a diferencia de los consonánticos que pueden ser o no fonados. Sin embargo, la principal diferencia entre los fonemas vocálicos y consonánticos es el tamaño de estas constricciones asociadas a las articulaciones en el tracto, siendo en las consonantes estas obstrucciones mucho más estrechas.

Las vocales se clasifican según la apertura del tracto vocal (vocales abiertas y cerradas) y según el grado de elevación del dorso de la lengua (vocales anteriores, centrales y posteriores). Por otro lado las consonantes se clasifican según el punto de articulación (lugar de contacto de los articuladores que participan en la producción de un fonema específico), modo de articulación (forma como se interrumpe el flujo aéreo espiratorio), función de las cuerdas vocales (presencia o ausencia de este al momento de producir un fonema), y posición del velo del paladar (capacidad de controlar el paso de aire hacia la cavidad nasal u oral).

1.3. Prosodia: Acentuación y Entonación

La prosodia evalúa las variaciones del tono de los fonemas sonoros en base al seguimiento de su frecuencia fundamental. Se distinguen dentro de ella dos patrones lingüísticos de interés: la acentuación y la entonación. La acentuación es un rasgo que permite poner en relieve un fonema para diferenciarlo de otras unidades del mismo nivel dentro de un nivel morfológico (palabra). El concepto de entonación es similar al de acentuación sólo que representa expresión a nivel de oraciones, siendo además mucho más clara la variación en la frecuencia fundamental.

1.4. Computerized Speech Lab (CSL)

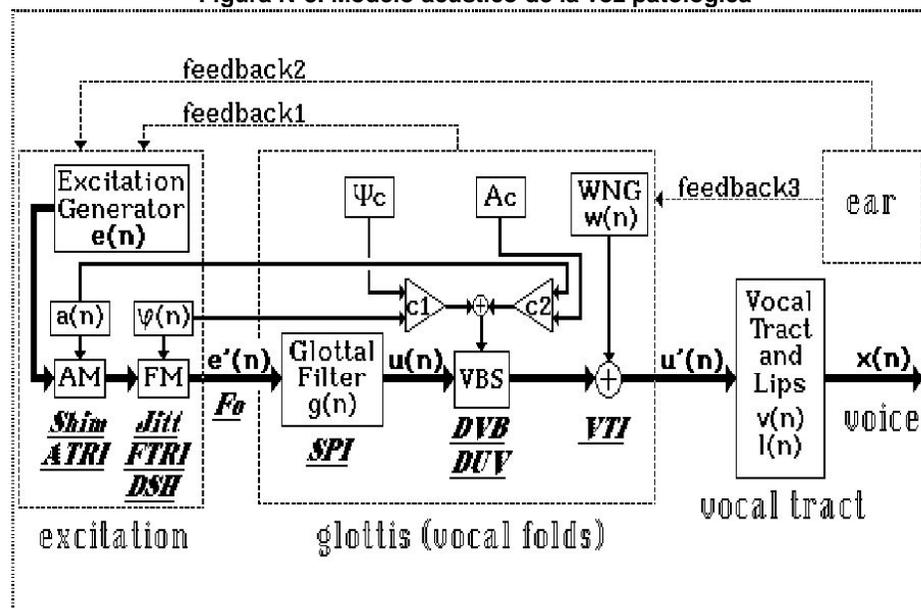
Software de análisis espectral enfocado especialmente para el análisis de la voz humana. Tiene las mismas herramientas de FFT que otros softwares de análisis espectral tales como espectro lineal, espectrograma, espectro 3D, más una serie de análisis especialmente diseñados para el estudio de la voz humana, tales como LPC (calcula la envolvente del espectro en base a una predicción lineal), historia de formantes, (identifica los formantes a lo largo del tiempo), Pitch (seguimiento de la fundamental a lo largo del tiempo), Cepsturm (calcula el logaritmo del espectro), Palatogramas (representa los movimientos lengua-paladar), estadísticas dedicadas, entre otros.

1.5. Multi-Dimensional Voice Program (MDVP)

El Multi-Dimensional Voice Program (MDVP) es un software que trabaja en conjunto con el Computerized Speech Lab (CSL). El MDVP permite la adquisición, análisis y cálculo de más de 33 parámetros de la voz a partir de una vocalización sostenida de un fonema sonoro, los cuales están disponibles como un archivo numérico, en ventanas de análisis o pueden ser mostrados gráficamente comparándolos con una base de datos. Esta comparación gráfica permite una rápida visualización de todos los parámetros del paciente. Este programa no evalúa posibles trastornos resonanciales a nivel del tracto vocal.

Algunos de los parámetros que se extraen en el programa han sido largamente estudiados y se ha logrado establecer sus relaciones con ciertas patologías de la voz (Ej. Jitter, Shimer). Sin embargo, una gran cantidad de estos parámetros son nuevos y aún están en estudio, por lo que se consideran sólo referenciales hasta no contar con más información que los respalde. El origen de todos los parámetros contenidos en este programa se basan en el modelo de voz patológica presentado en la Figura N°3, el cual presenta la misma estructura del modelo simple pero con mayor detalle en el origen del sonido a nivel de las cuerdas vocales. Se observa además en la Figura que no se consideran en este modelo parámetros asociados al tracto vocal ni a la radiación de salida en este modelo

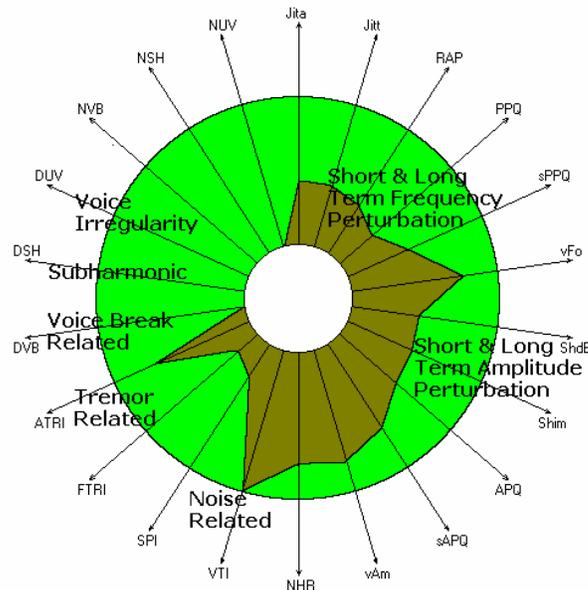
Figura N°3. Modelo acústico de la voz patológica



Los resultados son obtenidos a partir de una serie de procedimientos automáticos y comparados con una base de datos interna del programa que cuenta con valores referenciales de voces normales y patológicas. Esta base de datos es bastante pobre ya que cuenta con un total de sólo 68 personas evaluadas en EEUU (normales y patológicas). Sin embargo el programa permite la modificación de los valores de referencia a partir de datos del usuario. Esto permite adaptar los umbrales, a medida que se cuente con mayor información. De esta forma, podemos resumir el análisis realizado por el MDVP en seis grupos de análisis, de acuerdo al nivel de perturbación:

- Variaciones de Pitch y de Amplitud: Problemas asociados a excitación a nivel fonatorio, dado por un aumento de volumen en las cuerdas o problemas respiratorios.
- Quiebre e irregularidades de la Voz: Asociados a la interrupción de un sonido sostenido.
- Componentes Sub-Armónicos: Evalúa la cantidad de componentes de frecuencia menores a la fundamental. Estos efectos son típicos para las voces diplofónicas,
- Escape de aire y ruido: Posibles problemas de escape de aire (hiatus) y deficiente fonación.
- Temblores de la voz: Evalúa variaciones de frecuencia y amplitud de la señal asociadas a trastornos a nivel respiratorio que influye en la excitación en las cuerdas.

Figura N°4 Reporte de parámetros en forma de diagrama radial



No necesariamente existe una correlación directa entre todos los parámetros y los eventos fisiológicos que ocurren en las patologías de la voz, de modo que algunos parámetros son más fáciles de interpretar que otros. Aunque varios de los parámetros nos entreguen una idea del tipo de desorden, como es el caso de valores de Jitter y Shimmer elevados que indican la presencia de una masa anómala a nivel de las cuerdas vocales, sus valores no son necesariamente indicadores de las características propias de la patología, como el tamaño y ubicación de la masa. Por otro lado resulta muy difícil diferenciar auditivamente algunos de los parámetros tales como Jitter y Shimmer.

Finalmente se considera al MDVP como un método que permite respaldar de forma complementaria un diagnóstico, pero debe ser contrastado con otros métodos de análisis.

2. Aplicaciones en Lingüística: Fonética Acústica

Podemos clasificar las aplicaciones del análisis acústico en el estudio de lenguajes en base a los siguientes tipos análisis:

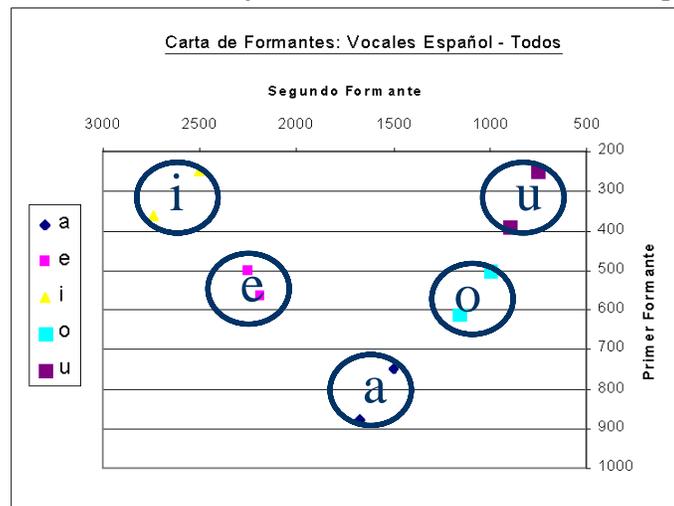
2.1. Estudio de fonemas individuales y de coarticulación.

Se realiza mediante análisis espectral de cada fonema y de la coarticulación entre ellos. La coarticulación se asocia a la transición de los formantes que se produce desde la configuración de un fonema hacia otro. El movimiento que realizan los formantes aporta información sobre el modo y lugar de articulación. Se evalúa cada fonema a partir de frases que contengan información fonéticamente balanceada.

2.2. Estudio de la articulación vocálica

En base al análisis de los formantes en fonemas vocálicos, se puede obtener una relación directa con su clasificación lingüística. Esto se realiza mediante la carta de formantes, análisis que permite obtener la configuración de los elementos articulatorios que se producen en las distintas vocales, en un gráfico que se construye a partir de la relación del primer y segundo formante. Para estos efectos se requiere de la emisión de todas las vocales por separado.

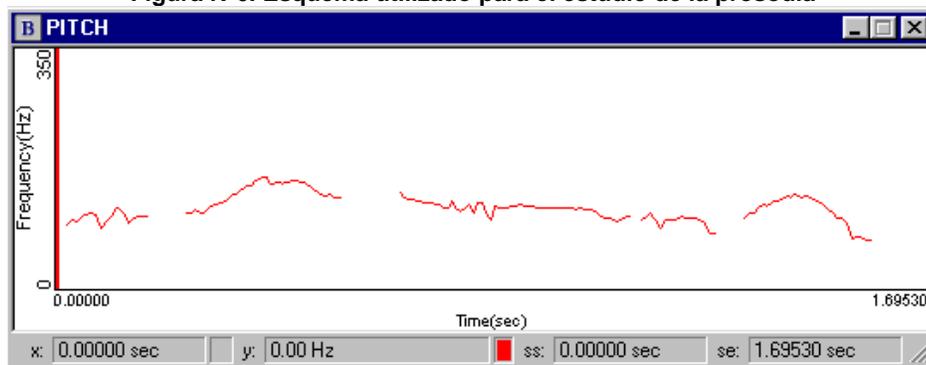
Figura N°5. Carta de formantes para los fonemas vocálicos de la lengua Española



2.3. Estudio de la prosodia

La prosodia se asocia a las variaciones en el tiempo de la frecuencia fundamental (Fo) de la vibración de las cuerdas vocales, las cual determina la melodía de la voz, tanto a nivel de entonación (frases) como acentuación (palabras). Es uno de los principales rasgos fonéticos que varía localmente dentro de una misma lengua. Se realiza mediante la evaluación del habla espontánea.

Figura N°6. Esquema utilizado para el estudio de la prosodia



3. Aplicaciones en estudios de Patologías Vocales

De igual forma podemos clasificar las aplicaciones en el estudio de emisión vocal mediante los tipos de análisis realizados:

3.1. Estudio de patrones objetivos de la emisión de voces

Gran parte de los criterios que se utilizan actualmente para describir las características vocales se basan en criterios auditivos que muchas veces suelen ser subjetivos. La aplicación de las herramientas de análisis espectral describe de modo objetivo la emisión de una voz lo que permite unificar conceptos dentro de los especialistas.

3.2. Descripción detallada de patologías vocales

Las herramientas presentadas permiten implementar un completo examen acústico de la voz, identificando tanto el comportamiento general de la voz (incluyendo trastornos resonanciales) como el detalle de la emisión fonada a nivel de las cuerdas. Para estos efectos se requieren múltiples herramientas de análisis espectral y estadístico. Estas metodologías siempre consideran la comparación de los patrones vibratorios en estudio con bases de datos de emisión normal de voz que permiten identificar el tipo de patología.

Figura N°7. Estudio de Patologías de la voz mediante el contraste de espectro completo

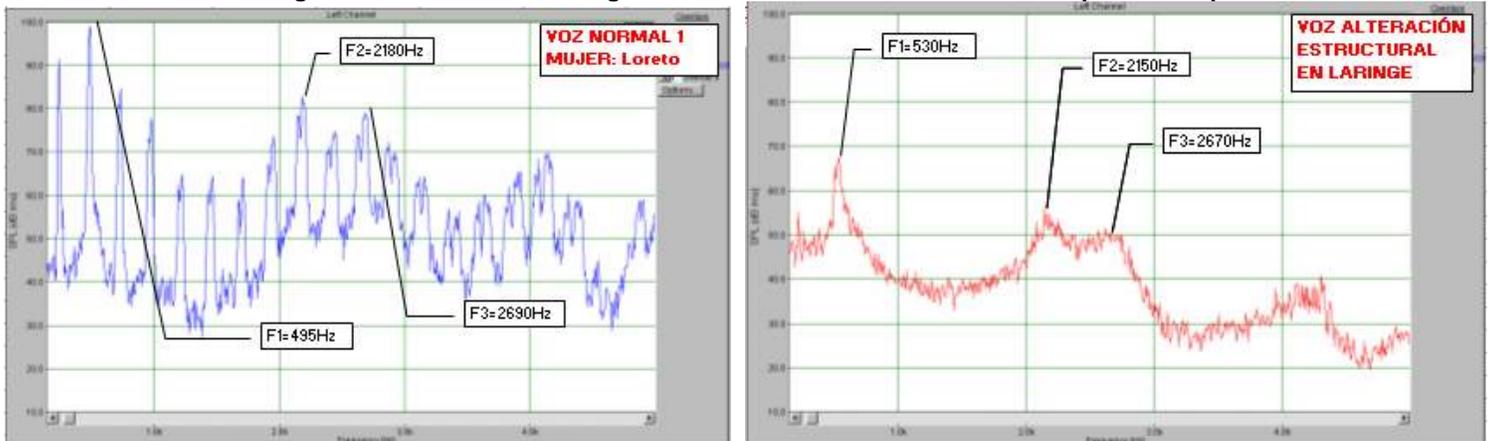
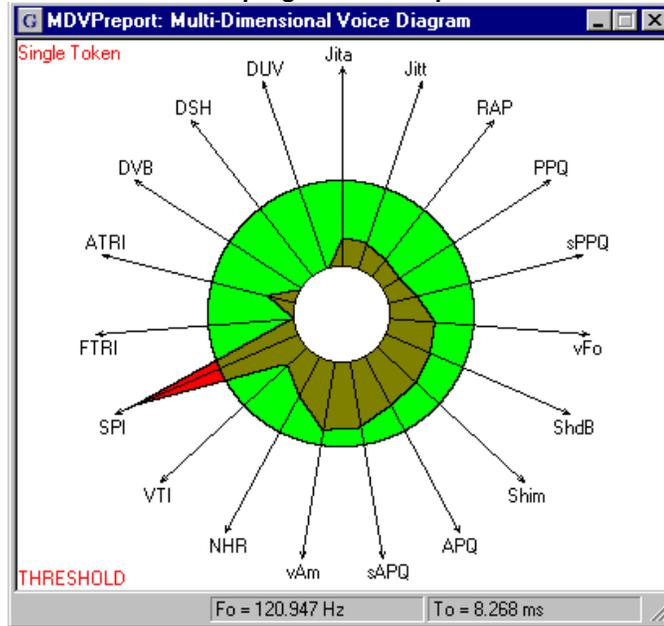


Figura N°8. Resultado del análisis del programa MDVP para una voz con fonación deficiente



3.3. Seguimiento de tratamientos fonoaudiológicos

Gran parte de los tratamientos fonoaudiológicos requiere de un número de sesiones donde el paciente va presentando una evolución en su trastorno. Las herramientas de análisis espectral entregan una descripción objetiva de la evolución del paciente, la cual se está desarrollando como parte de la ficha médica del paciente. Por otro lado, la visualización gráfica de los parámetros sonoros de la voz facilita el tratamiento con cada paciente.

4. Otras aplicaciones

Existen una gran variedad de otras posibles aplicaciones de las herramientas de análisis espectral en el estudio de la voz humana. Una de ellas consiste en el Reconocimiento e identificación de voces mediante el uso de las mismas herramientas de análisis espectral presentadas. Con estas herramientas es posible identificar patrones de voz característicos de una persona, aunque no mediante reconocimiento automático de estos patrones. Es posible lograr identificar patrones tales como su relación entre formantes, duración de fonemas característicos, puntos de articulación, entre otros, los cuales son requeridos en el modelo estadístico que permitirá diferenciarlos de otros patrones de voz. De esta forma este procedimiento sólo se realiza mediante post proceso de las señales en estudio.

5. Conclusiones

Se ha presentado en este trabajo cómo herramientas de análisis ampliamente usadas por especialistas en Acústica pueden tener interesantes aplicaciones en distintas áreas del conocimiento, como es el caso de las ramas asociadas al estudio de la voz humana. La aplicación de estas herramientas entrega una nueva perspectiva que presenta grandes beneficios tanto en el diagnóstico y seguimiento de patologías vocales, así como en aplicaciones al estudio de los lenguajes.

Hoy en día los estudios de la voz y la audición humana requieren cada vez más de equipos multidisciplinarios de especialistas que aporten desde sus distintas perspectivas los conocimientos necesarios para lograr un fin común, y la participación de Ingenieros especialistas en Acústica en estos equipos se hace cada vez más necesaria y conciente.

6. Bibliografía

Zañartu M., Barrera A., Espinoza V. "Guías de Laboratorio de Física Acústica", Escuela de Fonoaudiología, Universidad Mayor, 2002.

Dimitar D. Deliyski, "Acoustic Model and Evaluation of Pathological Voice Production", Proceedings: 3-rd Conference on Speech Communication and Technology EUROSPEECH'93, Berlin, Germany.

Crocker M., "Handbook of Acoustics", Wiley-Interscience, 1998.

Quilis A.M., "Fonética Acústica de la Lengua Española", Editorial Gredos, 2000.

Jackson Menaldi, "La Voz Patológica", Editorial Medica Panamericana, 2002.