

Aplicaciones de la Acústica a la Logopedia

José Luis Miralles Adell
Universitat de València

1. Introducción

El uso del lenguaje en los procesos de producción y comprensión implica una serie de transformaciones de estructuras que van desde una representación cognitiva a la producción hablada en el primer caso y desde una secuencia hablada hasta la elaboración de una representación mental en el segundo. Esta secuencia de transformaciones puede representarse del siguiente modo:

Producción

Repres. cognitiva → neuroanat. → acústica →

Comprensión

neuroanat. → esquema mental

De acuerdo con este esquema las estructuras acústicas están al final de la cadena de producción e inician los procesos de comprensión. Bastaría la consideración de esta secuencia para descubrir la importancia y sentido que tiene la información acústica en los procesos de producción y comprensión lingüística. Las alteraciones de la producción pueden manifestarse al final de la secuencia bien porque desde el principio exista una patología de origen central que dificulta el correcto funcionamiento de los efectores periféricos bien porque la patología esté localizada en los mismos órganos de la fonación. En cualquier caso esta alteraciones se manifestarán en un patrón acústico desorganizado. A su vez patrones acústicos empobrecidos o alterados van a dificultar considerablemente la comprensión del lenguaje. Ade-

más de la integridad y correcto funcionamiento tanto de los mecanismos neuroanatómicos periféricos y centrales como cognitivos la adecuada comprensión dependerá de la naturaleza de la señal acústica y de sus relaciones con el sistema auditivo fundamentalmente. De este modo el sistema natural de comunicación de los seres humanos implica básicamente la actividad de un canal oral-auditivo a través del sonido producido y recibido.

2. Nociones básicas de acústica.

Antes de exponer propiamente el proceso de producción del lenguaje conviene realizar un breve recordatorio de nociones y conceptos básicos de Acústica que ayuden a comprender las aplicaciones logopédicas.

Dentro de las múltiples definiciones que podemos encontrar del sonido, a los efectos de tema que ahora nos interesa, podemos considerarlo como *“una perturbación que se produce en un cuerpo, se transmite a través de un medio y finalmente estimula un sistema auditivo”*. Esta perturbación puede ser caracterizada de acuerdo con los siguientes parámetros acústicos:

Frecuencia: determina el número de ciclos o repeticiones por unidad de tiempo, convencionalmente el segundo. Se mide en hercios (Hz) y el sistema auditivo humano responde entre 20 y 20.000Hz. La frecuencia de vibración determina el tono, percibido como grave o agudo, de un sonido. La frecuencia es inversa al periodo, teniendo en cuenta que el periodo mide la duración temporal de un ciclo:

$$F = 1/T$$

Amplitud: es una propiedad física determinada por la cantidad de energía del sonido. Convencionalmente se toma el decibelio (dB) como unidad de medida. El sistema auditivo responde a un rango entre 0 y 130 dB aproximadamente. 0dB marcan en el umbral mínimo de audición y 130 corresponden a un umbral doloroso y con riesgo de lesiones auditivas. La amplitud física determina la intensidad percibida como parámetro psicológico.

Fase: hace referencia al instante temporal en el que se encuentra un punto de la oscilación. No es evidente su correlato psicológico.

Duración: es directamente la medida temporal de un sonido

Espectro o composición espectral: representa la amplitud de los componentes de un sonido en función de la frecuencia. El espectro fundamenta el timbre de un sonido entendido como aquella propiedad que permite caracterizar o individualizar una fuente sonora diferenciandola de las demás.

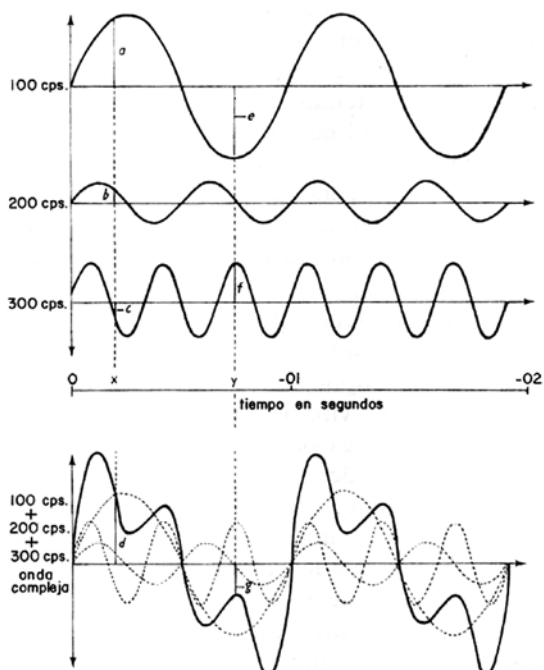
La información que transmite el habla se organiza fundamentalmente sobre la *frecuencia, la duración y la amplitud*, de manera especial sobre los

dos primeros parámetros. Desde una perspectiva fonética las características acústicas que permiten diferenciar y reconocer unos sonidos como distintos de otros son su contenido frecuencial y la duración. La distancia a la que se transmite la información y la prosodia, que transmite información emocional del hablante, están soportadas además, por las variaciones de la amplitud.

El espectro, parámetro de carácter más global e integrador, del que depende el timbre, está en la base de las propiedades cualitativas de la voz que habitualmente es evaluada como más o menos agradable, brillante, apagada, etc.

Atendiendo a la complejidad los sonidos pueden clasificarse en *tonos puros* y *sonidos complejos*. Los primeros sólo se presentan en el laboratorio y en audiometría tonal. Todo lo que percibimos son señales complejas. Estas señales complejas pueden descomponerse en una serie de tonos simples mediante análisis de Fourier de manera que es posible conocer los componentes de cualquier sonido complejo (fig. 1). En la figura 1 la señal compleja puede descomponerse en tres componentes simples de 100Hz., 200Hz. y 300Hz. respectivamente.

Figura 1. Sonido complejo formado por tres componentes simples de 100 Hz., 200 Hz. y 300 Hz.

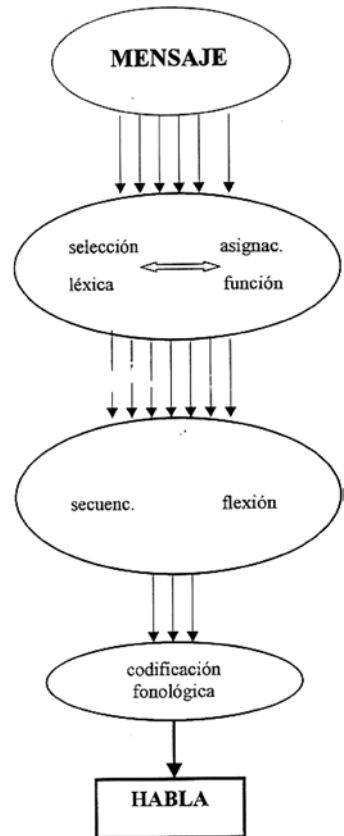


3. Producción del lenguaje.

En general los distintos modelos de la producción del lenguaje suponen que se trata de un proceso complejo con una primera fase de naturaleza

estrictamente cognitiva en la que se activan representaciones abstractas del conocimiento en forma de un tema de comunicación, se seleccionan las unidades léxicas y sus relaciones gramaticales. En un segundo momento estas representaciones se codifican neurológicamente y se transmiten, vía sistema nervioso periférico, desde las estructuras nerviosas centrales hasta los efectores periféricos que producen el habla (fig. 2).

Figura 2. Modelo de producción del lenguaje.



Algunos procesos son de naturaleza cognitiva pero otros implican la actividad de estructuras neuromusculares. La producción del habla se explica aceptablemente mediante la Teoría de la fuente y el filtro del habla (Fant 1960; Lieberman y Blumstein 1988). El habla puede ser considerada como la resultante de la interacción, técnicamente la convolución, de una fuente de vibración y un tracto de resonancia de manera que podemos decir

$$H = Fv * Tr$$

donde Fv representa las cuerdas vocales o fuente de vibración y Tr el tracto de resonancia. A efectos prácticos puede considerarse despreciable el efecto de radiación de la energía en los labios principalmente y en las paredes de la boca.

La voz se genera mediante vibración de las cuerdas vocales utilizando como fuente de energía el aire de la respiración. En la fase prefonatoria las cuerdas se encuentran adducidas. Debido a la presión suglotal deben abrirse y por la acción combinada de esta presión más el efecto Bernoulli se produce la vibración de las cuerdas. Esta señal acústicamente está compuesta de una frecuencia fundamental (F_0) y armónicos, que son frecuencias múltiplos de la fundamental. En general en el hombre adulto la frecuencia fundamental puede tener un rango entre 90 y 230 Hz., en la mujer llegaría hasta los 400 y en el niño hasta los 500 Hz. con valores medios de 128Hz. 225Hz. y 265Hz. respectivamente (Boone 1977).

Hay que tener en cuenta que la frecuencia fundamental depende de las características físicas de la fuente de vibración, de modo que su variabilidad puede ser muy considerable, de acuerdo con las distintas características anatómicas de los sujetos. Aparte del valor que tienen los datos acústicos de las cuerdas vocales en Logopedia en la medida en que adquieren valores anómalos, en algunas lenguas como el chino, las mismas vocales y consonantes tienen un significado diferente al variar la frecuencia fundamental.

La calidad de la voz depende pues de factores como la aducción y abducción de las cuerdas, la periodicidad de la vibración, amplitud y simetría de la onda generada. A su vez las características vibratorias están relacionadas con la viscosidad, densidad y elasticidad de las estructuras anatómicas glóticas (Hirano 1974).

La señal acústica generada en la glotis es transferida al tracto oral y nasal, en su caso, que actúan a modo de resonadores. El tracto de resonancia oral puede ser considerado como un tubo de aproximadamente 17,5 cm. de longitud en el hombre y de sección variable, determinada por las posiciones de la lengua.

El tracto de resonancia amplifica unas frecuencias de la onda generada en la glotis y atenúa otras. La bandas de frecuencias se convierten en los *formantes* y son aquellas regiones de energía acústica que transportan información a partir de la que se reconocen los fonemas. Es conveniente diferenciar los armónicos de los formantes. Los primeros corresponden a información directamente relacionada con la frecuencia fundamental de las cuerdas vocales mientras que los formantes son regiones de energía acústica amplificadas por el tracto vocal.

4. Caracterización acústica de la voz y Logopedia.

Teniendo en cuenta las características frecuenciales de la vibración es posible, en términos generales, diferenciar tres patrones vibratorios que determinan los tres principales registros vocales: voz dicrótica, modal y falsete. La voz dicrótica se caracteriza por su baja frecuencia, entre 20 y 60Hz. y un largo periodo de cierre glótico.

En el registro modal, el habitual en el habla, las cuerdas realizan una adducción total y completan un ciclo en el rango de 100-300Hz. La voz de falsete se consigue mediante el cierre de los dos tercios posteriores de las cuerdas permitiendo que vibre únicamente el tercio anterior con las cuerdas vocales delgadas, cortas y tensas produciendo la correspondiente elevación del tono.

A partir de estos patrones vibratorios y aplicando los principios de la mecánica a la vibración de cuerdas, ha sido posible obtener datos normativos de carácter acústico que, a su vez permiten evaluar el ajuste o desviación de la voz respecto de dichos patrones. A partir de los parámetros acústicos que caracterizan las ondas sonoras y teniendo en cuenta que la voz normal consiste en ondas periódicas más ruido aleatorio es posible desarrollar una serie de índices y parámetros acústicos que pueden ser aplicados a la evaluación de la voz normal o patológica. Las variaciones en *frecuencia, amplitud, la relación señal-ruido y el temblor* se han convertido en indicadores valiosos en el estudio de la voz de tal manera que existen distintos programas informáticos que realizan cálculos y comparaciones con los patrones normalizados.

Frecuencia. En Baken (1987) es posible encontrar datos normativos en función del género, edad y tarea ejecutada. Si bien es cierto que han sido obtenidos para hablantes de lengua inglesa hay que tener en cuenta la posibilidad de generalización en la medida en que se trata de valores determinados por estructuras anatómicas y funcionales no determinadas culturalmente. Aun cuando los programas informáticos arriba mencionados dan diferentes valores referidos a la frecuencia de vibración de las cuerdas es posible resumirlos en dos parámetros básicos: la frecuencia fundamental promedio y las variaciones ciclo a ciclo. Una desviación estadísticamente significativa de la media y del rango respecto de los valores que se consideran normales para una determinada población, atendiendo a la edad y el género, pueden indicar algún tipo de anomalía que podrá ser confirmada mediante la observación directa de las cuerdas con las técnicas pertinentes.

Quizá menos conocido es el concepto de *jitter*. Teniendo en cuenta que la producción de la voz es una actividad dinámica al considerar las características de la vibración ciclo a ciclo se observa cierta variabilidad entre ellos. Estas variaciones, al analizar una vibración ciclo a ciclo son normales cuando ocurren dentro de un rango determinado de valores pero son un indicador de la presencia de una patología cuando se encuentran fuera del rango de normalidad. Las variaciones mínimas que ocurren en cada ciclo son necesarias para que la voz pierda el carácter metálico y monótono propio en ocasiones de determinadas voces sintetizadas. No obstante cuando las variaciones del jitter se encuentran fuera de los valores normativos indican una falta de control en el movimiento de las cuerdas, que puede estar

asociado a un trastorno neurológico o a la presencia de una modificación anatómica.

Amplitud. Ya se ha mencionado que la amplitud depende de la cantidad de energía que contiene la señal acústica y que su correlato psicológico es la intensidad. En condiciones normales, a un metro de distancia, a un oyente le llegan 60dB. Si bien se ha dicho con anterioridad que las dimensiones básicas que transmiten información diferencial sobre el habla son la frecuencia y la duración, la evaluación de la amplitud debe ser tomada en cuenta tanto en la disfonías hiperquinéticas como en las hipoquinéticas, bien para su reducción o para su amplificación, pero de manera especial en los problemas que presenta la voz profesional, por ejemplo en el caso de los cantantes. Al igual que ocurre con la frecuencia la amplitud presenta también variaciones ciclo a ciclo denominadas shimmer.

Alteraciones tanto de la frecuencia fundamental como de la amplitud o intensidad pueden ser observadas en tipos distintos de patologías vocales teniendo en cuenta que en la producción del habla intervienen factores neurológicos, biomecánicos, aerodinámicos y acústicos. Así ciertas alteraciones neurológicas presentan inestabilidad del tono, la intensidad y la cualidad vocal y son percibidas como *temblores* en la voz, es decir, como fluctuaciones hasta 15 Hz. De este modo el análisis acústico en la enfermedad de Parkinson muestra altos valores de temblor e inestabilidad del patrón acústico aun cuando no sea posible determinar si estas fluctuaciones dependen de la alteración funcional del sistema nervioso central o de mecanismos periféricos.

Relación armónicos/ruido (NHR). Si tenemos en cuenta que la voz consiste en vibración periódicos de las cuerdas más ruido aleatorio otro indicador de interés que puede observarse en la información que contiene la señal acústica viene dado por la relación señal / ruido. El ruido puede proceder de la presencia de aire de la espiración que no ha sido transformado en vibración pero también es ruido la presencia de señal acústica que no corresponde con la estructura de armónicos de la frecuencia fundamental producida por alteraciones de las cuerdas. En general cuando el ruido acústico, es decir, aquella información acústica que se desvía de la composición armónica, sustituye a los valores armónicos la voz queda caracterizada por la presencia de ronquera tanto mayor cuanto peor es la relación señal / ruido.

La voz susurrada o cuchicheada se caracteriza precisamente por la presencia de turbulencia en el espectro acústico generada porque la glotis no está completamente cerrada ni abierta atenuándose en consecuencia la amplitud de la vibración de la F_0 . En este caso la intensidad de estas turbulencias puede enmascarar la de los componentes armónicos con la consiguiente pérdida de cualidad vocal.

Los parámetros acústicos relacionados con F_0 , amplitud, NHR y temblor se utilizan en la evaluación y tratamiento de voces patológicas pero tienen también su extensión al estudio de las variaciones prosódicas, que transmiten, entre otras, información del estado emocional del locutor. La laringe es especialmente susceptible a pequeñas fluctuaciones de sistemas como el nervioso, respiratorio, vascular o linfático en la medida en que el cuello es un lugar del cuerpo humano por el que pasan en estrecha confluencia estos sistemas.

Desde esta perspectiva emocional la información acústica puede ser utilizada eficazmente, por ejemplo, en el entrenamiento de actores o cantantes.

La voz modal, en términos promedio, exhibe pulsos glóticos en forma de pico de sierra, la mayor parte de la energía se concentra por debajo de 1kHz con componentes armónicos, el jitter es inferior al 1% y la energía de las vocales suele ser 1,2dB mayor que la de fricción. En el caso de los estados de ira por ejemplo, la energía de la vocal es superior en 1,5 o 2.0 dB a la de fricción, los pulsos glóticos son muy pronunciados con pendientes poco suavizadas y elevación de la intensidad y el jitter ligeramente mayor que en la voz modal. La voz triste o melancólica se caracteriza por una pequeña diferencia en intensidad, menor que 1 dB, entre los segmentos vocálicos o con vibración de la cuerdas y la fricción, alta presencia de ruido debida a la salida excesiva de aire durante la fonación que enmascara principalmente los componentes en altas frecuencias del espectro. En ocasiones, especialmente al principio y al final de las frases aparece voz dicrótica.

En expresiones emocionales dominadas por el miedo la diferencia de intensidad entre segmentos vocálicos y fricativos es mínima. La frecuencia fundamental exhibe marcadas alteraciones que se manifiestan en valores altos de jitter. La frecuencia fundamental aparece elevada y con pocas variaciones de intensidad prosódica. El análisis de voz aburrida tiene características de la voz triste aunque con ausencia de irregularidades en la frecuencia fundamental. De alguna forma podría caracterizarse como voz modal pero con muy pocas variaciones dinámicas en el contorno de la F_0 y por tanto con muy poca variabilidad prosódica y transmisión emocional.

5. El tracto de resonancia.

El siguiente elemento en la producción del habla, más allá de las cuerdas vocales, es el tracto de resonancia o filtro, mecanismo también de naturaleza acústica.

Ya se ha señalado que la acción del tracto de resonancia consiste en amplificar unas bandas de frecuencia de la señal acústica generada en la glotis y atenuar otras. En general un resonador puede ser considerado como un cuerpo que "suena" cuando es estimulado por una fuente de sonido previa-

mente activada. Las propiedades acústicas de un cuerpo dependen de sus características físicas, en este caso las de un tubo, dado que así puede ser considerado el tracto oral, de su longitud y sección. Sin embargo el tracto oral no debe ser considerado un tubo de longitud y sección uniforme sino más bien como una serie de tubos de longitud y sección variables que depende básicamente de la posición de la lengua dentro del tubo, de la forma de los labios y de la elasticidad de las paredes de la boca. No obstante el papel acústico fundamental en las variaciones que presenta este tracto es el de la lengua.

La energía acústica emitida en el habla alcanza hasta los 8.000Hz. y no aparece en forma continua sino como bandas de frecuencia amplificadas o atenuadas. Estas regiones amplificadas se denominan *formantes* (Tabla 1) y contienen la energía acústica necesaria para una adecuada percepción del habla. En la medida en que la fonética acústica depende de la articulatoria, los formantes exhibirán propiedades relacionadas con las diferentes formas de emisión de los fonemas. Así, en el caso de las vocales los formantes se caracterizan por su estabilidad y duración frente a la variabilidad que manifiestan las consonantes. Teóricamente debe ser posible distinguir hasta 6-8 formantes en las emisiones vocálicas, sin embargo la energía acústica se concentra en los primeros o más bajos siendo prácticamente despreciables los formantes de altas frecuencias. De hecho para la percepción adecuada de las vocales son necesarios sólo los dos primeros formantes, F1 y F2, y asimismo a partir de los dos primeros formantes es posible sintetizar vocales bien diferenciadas. Téngase en cuenta que los sistemas de telefonía suelen filtrar la información acústica por encima de los 3.500-4.000 Hz y el lenguaje sigue transmitiendo adecuadamente la información.

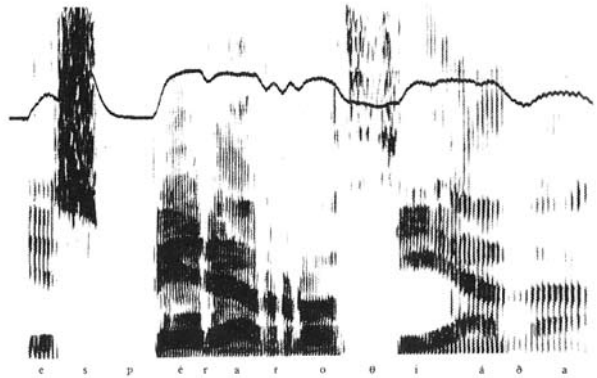
Tabla 1.
Formantes de las vocales en castellano

	f1	f2	f3	f4	f5
i	290	2250	2876	3390	3660
e	444	1986	2629	3300	4350
a	700	1300	2500	3280	4180
o	523	820	2584	3085	3600
u	295	650	2446	3210	3930

La representación acústica del habla se realiza en los *espectrogramas*. Un espectrograma es la representación de la frecuencia, la duración y la intensidad de la energía acústica que contiene un fonema, los tres parámetros, como ya se ha señalado que soportan la información relevante para una adecuada percepción del habla. Convencionalmente en el eje de abscisas se representa la duración, en el de ordenadas la frecuencia y la amplitud

puede determinarse, de manera aproximada, en base a la cantidad de “negro” de la representación gráfica. No obstante, en la actualidad los paquetes informáticos de análisis de la voz y el habla son capaces de cuantificar, aunque sea de manera relativa, los valores correspondientes a estas dimensiones acústicas.

Figura 3. Espectrograma de la expresión “espera rociada”



Conviene señalar en este lugar que estos programas dan valores absolutos acerca de la duración y la frecuencia, sin embargo no ocurre lo mismo con los datos que ofrecen sobre la intensidad de la emisión. La intensidad registrada depende, entre otras variables, de la distancia del locutor respecto del mecanismo de grabación, y además de las características físicas de este mecanismo transductor. Eso quiere decir que una emisión grabada a distancias diferentes daría valores diferentes de intensidad. Para obtener una medida absoluta de la intensidad emitida es necesario un calibrado previo del equipo de registro. Calibrado que se realiza normalmente con un calibrador con características acústicas homologadas. En general los programas de análisis acústico del lenguaje no permiten tal calibración y por ello los valores que dan acerca de la intensidad son sólo relativos y referenciados a una base arbitraria que incorpora el propio sistema.

El espectrógrafo puede realizar dos tipos de análisis en función de los filtros seleccionados, de banda ancha y de banda estrecha. El filtro de banda ancha característicamente tiene un ancho de banda de 300 Hz mientras que el de banda estrecha es de 45 Hz. Evidentemente estos dos tipos de filtros dan información diferente que puede ser utilizada para fines diferentes. En el análisis de banda ancha entre el límite inferior del filtro y el superior pasan 300 Hz mientras que en el de banda estrecha sólo pasan 45 Hz. Ello quiere decir que el análisis de banda ancha tiene peor resolución frecuencial que el de banda estrecha, sin embargo da mejor información acerca del dominio temporal mientras que lo contrario ocurre con el de banda estrecha, que favorece la separación entre componentes o armónicos pero con peor información temporal. Habitualmente se utiliza el análisis de banda ancha para una primera caracterización. No obstante puede procederse a un análisis de

banda estrecha si se observan alteraciones en el espectrograma inicial o bien con propósitos específicos como puede ser en el análisis de los armónicos en el caso de cantantes que presentan alguna anomalía.

Precisamente por la información que da, donde además de la acción del tracto de resonancia puede visualizarse la f_0 , el espectrógrafo puede ser utilizado no sólo para la investigación sino también como un instrumento de retroalimentación muy válido en el entrenamiento y rehabilitación vocal.

Los valores formánticos aparecen alterados por patologías diferentes, edad o cualquier estado variable que afecte bien a las características anatómicas o funcionales de etiología mecánica o neurológica, del tracto de resonancia. En este contexto es posible establecer algunos principios de carácter general para modificar las estructuras articulatorias. Así, puede afirmarse que *“todas las frecuencias de los formantes disminuyen uniformemente en la medida en que el tracto se alarga”*. Es bien conocido que los tubos más largos resuenan con frecuencias más bajas que los cortos. Ciertamente no es fácil modificar la longitud del tracto oral, sin embargo es posible una modificación de hasta un 10% tanto por elevación y descenso de la laringe como por protrusión y retracción de los labios. Como resultado de estas modificaciones cambia el “brillo” vocálico más brillante o apagado.

Un segundo principio establecería que *“todas las frecuencias de los formantes disminuyen uniformemente con el redondeamiento de los labios y se incrementan con su expansión”*. El redondeamiento de labios implica una oclusión parcial de la boca y fácilmente puede comprobarse como en esa situación el habla adquiere una cualidad oscura. No se trata aquí de proponer ejercicios de rehabilitación logopédica; simplemente se citan los principios anteriores con carácter ejemplar para señalar las posibilidades de la aplicación de los conocimientos de la acústica de la voz y habla a la investigación, la evaluación y el tratamiento logopédicos.

En definitiva, la evaluación subjetiva o perceptiva, técnica irremplazable en Logopedia, presenta limitaciones como la variabilidad de criterios de evaluación y clasificación, falta de acuerdo en la terminología descriptiva, dependencia del entrenamiento, etc. Sin embargo instrumentos habitualmente utilizados en la evaluación subjetiva de la voz como la escala GBRAS (ver Hirano 1981) que permite valorar la cualidad vocal en base al grado de anormalidad, aspereza, aspiración, astenia y tensión o la investigación llevada a cabo por el equipo de Hammaberg (1986, 1995) ponen de manifiesto un alta correlación entre las características perceptivas de la voz informadas por expertos y las propiedades acústicas de la misma.

Hasta ahora se han comentado las relaciones de la Acústica con la Logopedia a través de los procesos de producción, en los que estas relaciones parecen más evidentes. No obstante también la comprensión, aunque en menor medida, puede beneficiarse de dichos conocimientos.

6. Comprensión del lenguaje.

La comprensión de un mensaje implica la elaboración de un esquema de representación mental. Esta elaboración, de acuerdo con una teoría modular de la mente (cita) cuya validez provisional está puesta de manifiesto tanto por investigaciones experimentales como por observaciones clínicas, los procesos mentales tienen una naturaleza compleja y se caracterizan por la posibilidad de poderse descomponer en una secuencia de microprocesos y operaciones. En el caso de la comprensión del lenguaje esta secuencia de operaciones podemos reducirla a tres microprocesos básicos:

1. recepción sensorial de la señal, que corresponde a la actividad de los receptores sensoriales periféricos
2. acceso al léxico: activación de toda la información que contienen las unidades léxicas
3. procesos de alto nivel: análisis sintáctico, asignación de roles y funciones, integración de información, inferencias.

Considerando que los dos últimos microprocesos son básicamente de naturaleza cognitiva que actúan sobre el nivel representacional de la información, la actividad de los receptores sensoriales opera directamente sobre la base física del lenguaje y sus alteraciones pueden ser de consecuencias muy graves tanto en la comunicación humana como en el desarrollo de las funciones cognitivas.

Aunque sólo sea por motivos didáctico hay que recordar que la percepción del lenguaje, fase previa a su comprensión, responde a las siguientes características:

- a. es un proceso interactivo bottom-up y top-down que depende tanto de la información que tiene el sujeto almacenada en memoria como de la que recibe. En este sentido es posible decir que opera sobre:
 - b. información cognitiva
información acústica
 - c. un procesamiento adecuado exige ciertas condiciones de los receptores periféricos, con independencia de las habilidades cognitivas.
 - c.1. umbral absoluto: determinado audiométricamente. Habitualmente ésta es la única exploración que se hace de los sujetos cuando se sospecha una deficiencia auditiva. En general y al margen de la técnica de exploración utilizada, los distintos tipos de audiometría, impedanciometría, potenciales evocados, la audiometría convencional lo que busca es la cuantificación de la pérdida auditiva en términos de dB por frecuencias de manera que pueda determinarse la cantidad mínima de señal acústica capaz de provocar la respuesta del sistema auditivo.

c.2. umbral diferencial: el umbral diferencial está determinado por la variación mínima de energía acústica capaz de producir en el sistema sensorial una respuesta de discriminación. Este umbral, en la percepción auditiva depende de dos funciones que debe realizar el sistema

c.2.1. selectividad frecuencial y temporal

c.2.2. integración frecuencial y temporal

Dos de las características específicas del lenguaje oral son la alta variabilidad y la ausencia de segmentación. Cada locutor tiene unas características anatómicas que producen diferencias en los patrones acústicos generados, incluso los mismos fonemas varían al modificarse el contexto en el que se producen. De otra parte y a diferencia del lenguaje escrito en el que las palabras aparecen separadas por blancos, en el lenguaje oral unas palabras están unidas a otras dado que en la secuencia acústica no hay silencios de separación. Al mismo tiempo, y en situaciones de comunicación natural, los sonidos del habla aparecen enmascarados por ruidos ambientales que ocurren simultáneamente. En estas circunstancias la calidad de la información acústica adquiere especial importancia para el adecuado reconocimiento de la información verbal.

Las claves de esta información se encuentran en el patrón acústico generado, que es propio para cada fonema de la lengua e incluso variable en función del locutor, el contexto, la posición y la intención de comunicación. En el caso de las vocales la información requerida para su adecuado reconocimiento, como ya se ha dicho, se encuentra en los dos primeros formantes. El reconocimiento de las consonantes, emitidas de manera más rápida y ordinariamente con menor cantidad de energía, depende de claves acústicas mucho más débiles e inestables. En estas condiciones la habilidad del sistema auditivo para la resolución de información acústica que ocurre de manera simultánea, tiene muy poca duración y frecuentemente aparece enmascarada, es decisiva para el adecuado reconocimiento de dicha información. Es decir, el umbral absoluto medido audiométricamente es una condición necesaria para una buena percepción del habla pero no suficiente. La adecuada discriminación de los sonidos del habla exige una fina sintonización con las variaciones de frecuencia y duración que caracterizan a los fonemas.

En el caso de la deficiencia auditiva se observa que tanto los umbrales absolutos como los diferenciales se encuentran alterados. Esta alteración de los umbrales absolutos explica, por ejemplo, las dificultades que tienen los sujetos con deficiencia auditiva para la comprensión del lenguaje aun cuando el audífono haya sido capaz de restaurar su umbral de audibilidad. En definitiva, muchos de los problemas observados en el rendimiento que ofrecen prótesis y ayudas auditivas mecánicas solo pueden ser compren-

didados desde la Psicoacústica, es decir, desde el estudio de las relaciones cuantitativas entre la sensación y la señal acústica.

Finalmente un somero comentario acerca de los requisitos y condiciones para una grabación fiel de segmentos de habla que luego puedan ser analizados. Si es posible la grabación deberá hacerse en un lugar silencioso de manera que la información acústica de interés no se encuentre contaminada con otros ruidos ambientales. Además es necesario disponer de un equipo de grabación fiable que no altere la señal introducida, cuyos elementos básicos son micrófono, base del equipo de registro y soporte para la señal grabada. En la actualidad y con el desarrollo de la informática quizá lo más conveniente sea utilizar el propio ordenador como instrumento de grabación. El almacenamiento de la voz en disco duro o CD permite que no se deforme la señal y sea duradera con el paso del tiempo. Es conveniente disponer del software que permita no sólo la grabación de la señal sino también su edición. En la actualidad las tarjetas de sonidos más convencionales permiten ya grabar a 16 bits. La frecuencia de muestreo en el caso del habla es suficiente que esté situada a 11 kHz de manera que con esta frecuencia de muestreo será posible analizar hasta 5.000 Hz de la señal grabada, sin embargo la amplitud de memoria de los ordenadores y la grabación digital permiten en la actualidad grabar con una frecuencia de muestreo de 44.100 Hz. con las ventajas que esto puede representar. El micrófono debe caracterizarse por una buena sensibilidad, la direccionalidad buscada y respuesta plana, de modo que no distorsione la señal. Es conveniente disponer previamente de un protocolo verbal ya elaborado que contemple la grabación de vocales aisladas, palabras y pseudopalabras, frases y texto.

Bibliografía

- Baken, R.J. (1987). *Clinical measurement of speech and voice*. Boston, Mass. Allyn and Bacon
- Boone, D.R. (1977). *The voice and voice therapy*. London, Prentice Hall
- Fant, G. (1960). *Acoustic theory of speech production*. The Hague, Netherlands: Mouton, 2nd edition. 1970
- Hammaberg, B., Fritzell, B., Gauffin, J. y Sundberg, J. (1986). Acoustic and perceptual analysis of vocal dysfunction. *Journal of Phonetics*, 14, 533-547.
- Hammaberg, R. y Gauffin, J. (1995). Perceptual and acoustic characteristics of quality differences in pathological voices as related to physiological aspects. In O.Fujimura and M.Hirano (Eds.) *Vocal fold physiology: Voice quality control* (283-304). San Diego, Singular Publishing Group.
- Hirano, M. (1974). Morphological structure of vocal cord as a vibrator and its variation. *Folia Phoniatica* 26, 89-94.

- Hirano, M. (1981). *Clinical examination of voice*. Vienna, Springer-Verlag.
- Kent, R. D. and Read, Ch. (1992) *The Acoustic Analysis of Speech*. London - San Diego: Whurr Publishers, Singular Publishing Group
- Lieberman, P. and Blumstein, S.E. (1988) *Speech Physiology, Speech Perception and Acoustic Phonetics*. Cambridge: Cambridge University Press