

23. La acústica del habla

La energía del sonido es una onda de presión que consiste en la vibración de las moléculas en un medio elástico, como por ejemplo un gas, un líquido o algunas clases de sólidos. Al estudiar la producción del habla, normalmente nos ocupamos de la propagación del sonido a través del aire: las partículas de aire son perturbadas por los movimientos y las vibraciones de los órganos articulatorios, especialmente por las cuerdas vocales (§22). Pero cuando estudiamos la recepción del habla (§25), el aire no es el único medio implicado. El proceso de la audición requiere que las vibraciones en el aire se transformen en vibraciones mecánicas (a través de los mecanismos óseos del oído medio), cambios hidráulicos (a través de los líquidos del oído interno), e impulsos eléctricos nerviosos (por el nervio auditivo hasta el cerebro).

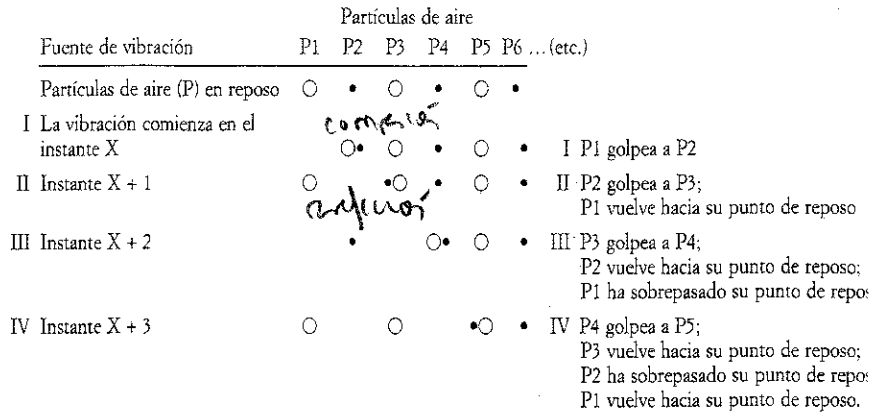
Cuando un objeto vibra, provoca un movimiento oscilante en las partículas de aire que lo rodean. Estas partículas afectan a partículas adyacentes, y el proceso continúa como una reacción en cadena mientras dura la energía. Si la vibración inicial conlleva una gran cantidad de energía, la onda sonora producida puede transmitirse a gran distancia, antes de desaparecer. Pero las propias partículas de aire no viajan a lo largo de toda esta distancia. El movimiento de cada partícula es meramente local: cada una influye en la siguiente, de la misma manera que una larga fila de fichas de dominó, colocadas unas cerca de otras, se derrumba si empujamos la primera. No obstante, a diferencia de las fichas de dominó, las partículas de aire vuelven a su posición inicial, una vez que han transmitido el movimiento a sus vecinas.

El movimiento de las ondas sonoras en el aire se ha comparado a veces con las olas que provoca una piedra cuando cae al agua; pero esta analogía no refleja la naturaleza tridimensional del movimiento. De la misma manera, el símil del dominó sólo da una idea limitada del movimiento que se produce. Una comparación mejor sería la de un globo que se infla, que se expande en todas las direcciones a la vez. También las ondas sonoras se mueven simultáneamente en todas las direcciones, desde el punto de origen.

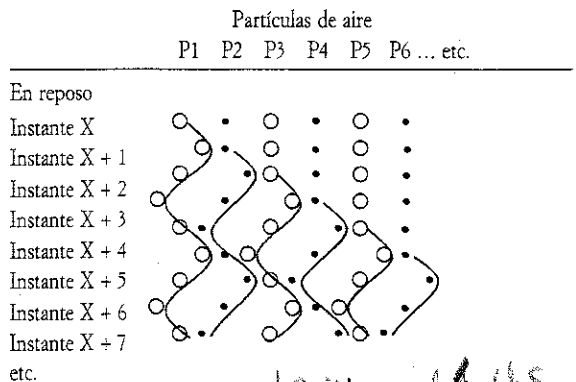
SINUSOIDES

El movimiento de las partículas puede compararse al de un péndulo o un columpio. Cuando está parado, el columpio cuelga verticalmente. Cuando se pone en movimiento, a un desplazamiento hacia atrás le sigue uno hacia adelante, a ambos lados del punto de reposo, mientras haya energía para mover el columpio. Este vaivén se conoce como *oscilación*. De igual manera, las partículas de aire oscilan alrededor de su punto de reposo. Cuando una partícula se mueve hacia adelante comprime a las partículas adyacentes, y causa un ligero aumento de la presión del aire en ese punto. Cuando se mueve hacia atrás deja de comprimir a esas partículas, y provoca una disminución de la presión. El movimiento es como el de una ola,

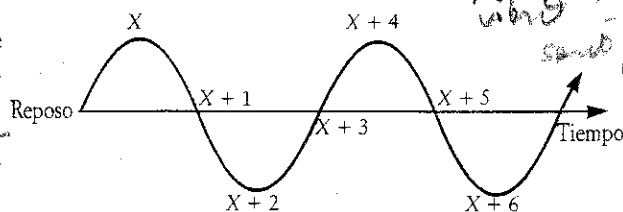
como podemos observar si seguimos el progreso de una serie de partículas, una vez que han sido puestas en movimiento por una fuente de vibraciones. En el siguiente diagrama, se representa el movimiento de cada partícula en instantes sucesivos (X + 1, X + 2, etc.), suponiendo que la fuente de vibración esté a la izquierda.



Si unimos con una línea las posiciones de cada partícula, podremos apreciar el movimiento ondulatorio, que seguidamente presentamos en vertical.



Se puede trazar un gráfico de la onda de presión que se forma cuando las partículas se mueven de esta manera; es lo que se conoce como sinusoide. Las ondas sinusoidales se suelen trazar de izquierda a derecha, a ambos lados de una línea horizontal que representa el paso del tiempo. El movimiento simple de una sola partícula tendría este aspecto:



Sonidos puros

Las ondas más sencillas, como las que aparecen en esta página, tienen forma de senoide, y consisten en una sola pulsación vibratoria, que se repite a intervalos constantes y produce un *sonido puro*. Estos sonidos difícilmente se escuchan en la vida diaria. La mayoría de los sonidos son compuestos, al estar formado por varios patrones simultáneos de vibración. Par producir un sonido puro, hace falta una máquina electrónica especial, o un instrumento como el diapasón. Cuando se golpea un diapasón (abajo), éste vibra sobre una sola nota. Las puntas de la horquilla oscilan con un ritmo constante. Si se acerca el diapasón al oído, se puede percibir un sonido puro



propiedad de la onda

A la oscilación de una partícula de aire se la llama *ciclo*, al número de ciclos que se suceden en un segundo se le llama *frecuencia* de un sonido. La frecuencia se solía medir en «ciclos por segundo» (cps), pero a esta unidad se le ha dado el nombre de *hertzio* (por el físico alemán Heinrich Rudolf Hertz (1857-94), que fue el primero en emitir y recibir ondas de radio), abreviado en Hz. La frecuencia básica de la vibración de un sonido se conoce como *frecuencia fundamental*, abreviado generalmente en F0, pronunciado «F cero».

La gama de frecuencias que puede oír un adulto joven normal es amplísima, desde unos 20 a 20.000 Hz. No se pueden oír las frecuencias que están por debajo («infrasónicas») o por encima («ultrasónicas») de estos valores. En cualquier caso, las frecuencias que se encuentran a ambos extremos de esta gama tienen poco interés para el habla: las frecuencias de habla más importantes se encuentran entre los 100 y los 4.000 Hz. La frecuencia fundamental de la voz de un hombre adulto, por ejemplo, es de unos 120 Hz, y la de una mujer de unos 220 Hz (pág. 128).

La frecuencia de un sonido puro guarda relación con la sensación de tono, que es nuestra percepción de que un sonido es «más alto» o «más bajo». En general, cuanto más alta es la frecuencia de un sonido, más alto es el tono que percibimos. Pero nuestra percepción del tono también se ve afectada por la duración y la intensidad del estímulo sonoro. Las nociones de «frecuencia» y «tono» no son idénticas: la frecuencia es un hecho físico, objetivo, mientras que el tono es una sensación psicológica, subjetiva (pág. 144).

Longitud de onda

La *velocidad* a la que el sonido viaja a través del aire tiene un *valor constante*, que habitualmente (según las condiciones de temperatura) corresponde a unos 343 m. por segundo. Todos los sonidos que tienen la misma energía tardarán el mismo tiempo en desplazarse de A a B. Durante el tiempo necesario para que se produzca un ciclo de la vibración, una onda sonora recorre una cierta distancia, que constituye la *longitud de onda* del sonido. Puesto que la velocidad a la que se mueve el sonido es constante, resulta que *cuanto más alta sea la frecuencia, más corta será la longitud de onda*. Dicha relación se expresa con una sencilla fórmula: $\lambda = (C/F)$, donde C es la velocidad del sonido, F es la frecuencia, y lambda es la longitud de onda. De este modo, un sonido de 500 Hz tendrá una longitud de onda de $(343 \text{ m.} / 500 \text{ Hz}) = 69 \text{ cm.}$; la de un sonido de 1.000 Hz sería de 34 cm.

La importancia de la longitud de onda se aprecia en nuestro modo de recibir el sonido: cuando una onda se aproxima a un objeto, si su longitud de onda supera el tamaño del objeto, tiende a «combarse» alrededor de éste, mientras que si la longitud de onda es menor, tiende a rebotar. Así, por ejemplo, cuando las ondas sonoras se acercan a la cabeza, es más probable que sean retenidas las frecuencias más bajas, que tienen una longitud de onda mayor, antes que las altas, factor que puede tener una importancia considerable cuando se trata de asistir a personas con problemas de audición (pág. 266).

La mayoría de las fuentes de sonido producen grupos complejos de vibraciones, y ese es siempre el caso del habla. El habla implica la utilización de ondas sonoras complejas porque es el resultado del empleo de muchas fuentes distintas de vibración dentro del tracto vocal (§ 22). Cuando se combinan dos o más ondas puras de frecuencia distinta, el resultado es una *onda compleja*. Existen dos tipos de onda compleja. En uno de ellos la onda sonora se repite: hay un patrón de vibración *periódico*. En el otro no hay tal repetición: las vibraciones se producen al azar, de forma *aperiódica*. En el habla se emplean ambos tipos. Los sonidos vocálicos, por ejemplo, presentan un patrón periódico, mientras los sonidos como el de la [s] son aperiódicos.

Armónicos

de resonancia

El sonido producido por un objeto que vibra de forma periódica implica más que una simple senoide (pág. 132). La misma vibración genera también otras cantidades de energía, ligadas a la senoide básica según una sencilla relación matemática: *todas son múltiplos de la frecuencia fundamental*. Es decir, una F0 de 200 Hz originaría una serie de frecuencias relacionadas de 400 Hz, 600 Hz, etc. Estos múltiplos se conocen como *armónicos* o *hipertonos*, y se numeran por orden. En física (pero no en música) F0 cuenta como el primer armónico, o sea que en el presente ejemplo 400 Hz sería el «segundo armónico», 600 Hz el «tercer armónico», etc. Este sistema constituye un marco adecuado para el estudio de las vocales, de algunas consonantes, y de los patrones de entonación (págs. 135-7).

Según cuál sea la naturaleza del objeto que vibra (por ejemplo el material de que está hecho, o su espesor), se establecen grupos de armónicos diferentes, que nosotros percibimos como diferencias de *timbre*. La diferencia que apreciamos entre dos voces o dos instrumentos musicales que producen un sonido con el mismo tono y la misma intensidad, es un contraste en el timbre, causado por la diferencia de los armónicos.

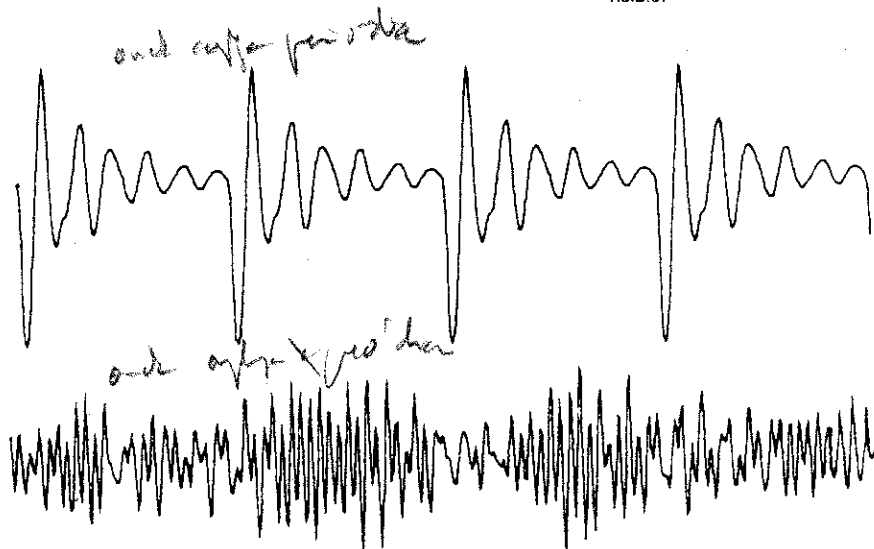
Una manera de relacionar la noción física de frecuencia con nuestra percepción del tono consiste en relacionar las notas musicales que nos resultan familiares con la frecuencia fundamental. El Do mayor tiene una frecuencia de 264 Hz, y la frecuencia de las notas que están por encima de éste en la escala diatónica de Do son las siguientes:

Do	528 Hz
Si	495 Hz
La	440 Hz
Sol	396 Hz
Fa	352 Hz
Mi	330 Hz
Re	297 Hz
Do	264 Hz

La nota La es la nota que toca el oboe cuando la orquesta está afinando. En comparación, la nota más alta de un piano de siete octavas es de 3.520 Hz, y la más baja de 27,5 Hz.

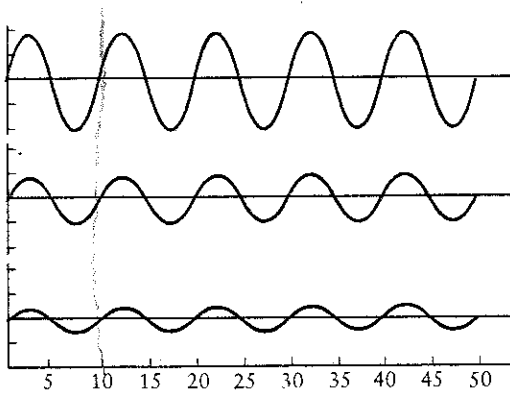
Ondas vocales

He aquí las ondas sinusoidales típicas de la vocal [a:] y de la consonante [s]. El espacio de tiempo que se representa es el mismo para ambos sonidos. En el caso del sonido vocálico el patrón periódico se puede apreciar claramente, pero en el caso de la [s] no hay ningún patrón visible.



AMPLITUD E INTENSIDAD

La distancia alcanzada por una partícula de aire a un lado y a otro del punto de reposo se conoce como *amplitud* de la vibración. Cuanto mayor es la amplitud, mayor es la intensidad del sonido y, por efecto también de otros factores (como la frecuencia y la duración), mayor es la sensación de altura que percibimos. En el siguiente diagrama podemos ver tres ondas de igual frecuencia pero distinta amplitud. En todos los casos una vibración completa dura 10 msec (por lo tanto la frecuencia es de 100 por segundo, o 100 Hz). Pero la amplitud de (a) es el doble que la de (b), y la de (b) el doble que la de (c).



Para medir la altura de un sonido tenemos que tener en cuenta tanto la amplitud como la frecuencia, dos factores que están relacionados con la energía con que el sonido se produce. El término *intensidad* se emplea para referirse a la potencia global de un sonido, y es una noción muy útil para el estudio del habla, donde las ondas sonoras son complejas, y la altura de un sonido no se relaciona claramente con ninguno de los componentes acústicos.

Decibelios

Para medir la intensidad del sonido necesitamos un nivel básico de referencia de la presión del sonido en el aire, que esté reconocido internacionalmente. Este nivel de presión del sonido tomado como referencia indica el límite a partir del cual un sonido puede ser oído (que se define tradicionalmente como 0.0002 dinas por centímetro cuadrado, donde *dina* es la unidad de medida de la presión). Los valores por encima de este nivel de referencia se miden en unidades conocidas como *decibelios* (db) (llamadas así por Alexander Graham Bell (1847-1922), in-

ventor del teléfono). Por consiguiente, si decimos que un sonido es de 90 db, significa que tiene una intensidad de 90 db por encima del nivel de referencia.

Nosotros somos capaces de oír una vasta gama de intensidades de sonidos. Un grito es un millón de veces más potente que un susurro. Se ha calculado que el oído humano es sensible a unos 10 millones de millones (10 elevado a 13) de unidades de intensidad. Para que los analistas puedan manejar cifras tan grandes, las intensidades de los sonidos se relacionan entre ellas de forma proporcional, utilizando una tabla de logaritmos. Un aumento de 10 db equivale a multiplicar la intensidad por dos: 30 db es el doble de alto que 20 db, 40 db es el doble de alto que 30 db, etc. De esta manera, 10 elevado a 13 puede reducirse a una escala de 130 decibelios, escala que refleja con más exactitud el modo en que percibimos la diferencia de intensidad entre los sonidos.

Se pueden calcular los valores medios de intensidad de cada uno de los sonidos del habla. En la siguiente tabla (según D. B. Fry, 1979, pág. 127), los valores de los sonidos ingleses, expresados en decibelios, se han puesto en relación con el sonido de menor intensidad, [θ] (como en inglés *thin*), al que se da el valor 0. Las vocales abiertas son los sonidos más intensos, seguidas por las vocales cerradas; las fricativas y oclusivas débiles se encuentran en el extremo opuesto de la lista (§ 27; para las convenciones de transcripción, ver apéndice 2). De acuerdo con esto, en una palabra como *thorn* el aumento de intensidad entre el primer sonido y el segundo es de unos 30 db.

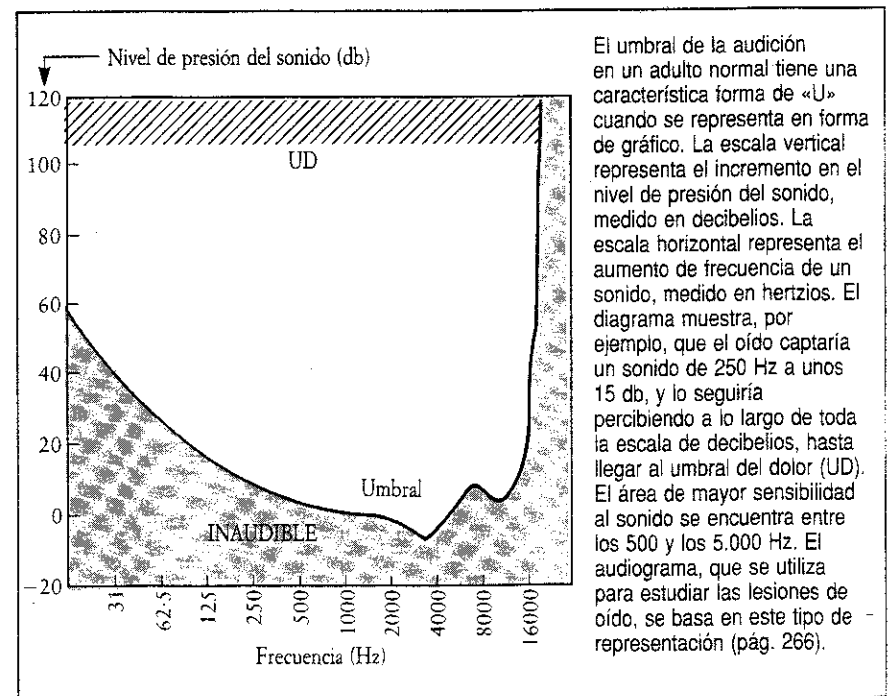
ɔ	29	e	23	l	20	ʒ	13	ð	10
ɒ	28	i:	22	ʃ	19	z	12	b	8
ɑ:	26	u:	22	ŋ	18	s	12	d	8
ʌ	26	ɪ	22	m	17	t	11	p	7
ɜ:	25	w	21	tʃ	16	g	11	f	7
a	24	r	20	n	15	k	11	θ	0
ʊ	24	j	20	dʒ	13	v	10		

Habla y sonidos cotidianos

La intensidad propia de las distintas clases de habla se puede apreciar comparándola con la intensidad media de algunos sonidos corrientes (según D. B. Fry, 1979).

0	umbral de audición
10	susurrar de las hojas
20	tictac de un reloj (pegado al oído);
30	jardín silencioso; conversación susurrada
40	zona residencial, sin tráfico
50	oficina tranquila, máquina de escribir
60	conversación a 1 m.; coche a 10 m.
70	tráfico urbano muy intenso a 30 m.
75	timbre del teléfono a 3 m.; gritos
80	tren del metro; música alta en la radio
90	taladradora a 1 m.
100	claxon de coche a 5 m.; fortissimo orquestal
110	taller de calderería
120	banda de rock con amplificación potente
130	avión de reacción de cuatro motores a 30 m.

Alrededor de los 120 db, la sensación de audición se convierte en sensación de dolor.



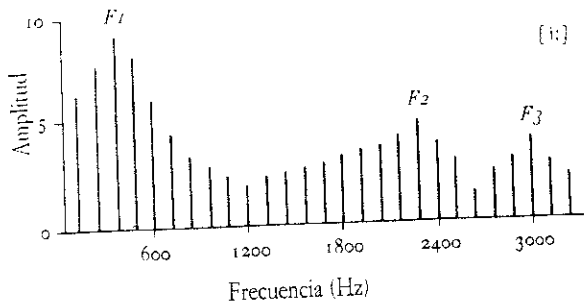
El umbral de la audición en un adulto normal tiene una característica forma de «U» cuando se representa en forma de gráfico. La escala vertical representa el incremento en el nivel de presión del sonido, medido en decibelios. La escala horizontal representa el aumento de frecuencia de un sonido, medido en hertzios. El diagrama muestra, por ejemplo, que el oído captaría un sonido de 250 Hz a unos 15 db, y lo seguiría percibiendo a lo largo de toda la escala de decibelios, hasta llegar al umbral del dolor (UD). El área de mayor sensibilidad al sonido se encuentra entre los 500 y los 5.000 Hz. El audiograma, que se utiliza para estudiar las lesiones de oído, se basa en este tipo de representación (pág. 266).

El sonido ambiental

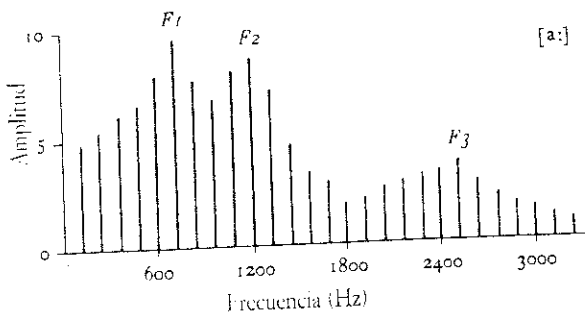
El tiempo que tardan las ondas sonoras en desvanecerse en el nivel de lo inaudible se conoce como «tiempo de reverberación». En una habitación, las paredes y los muebles absorben la energía. Los materiales fibrosos, como las cortinas y las alfombras, absorben bien el sonido, mientras que las superficies duras y densas lo reflejan. En las aulas modernas se suelen usar estos últimos materiales, que producen altos niveles de ruido («ruido ambiental»), y a menudo son la causa de que sea difícil para los niños oír lo que se dice.

ESPECTROS

Se puede hacer un análisis acústico de una onda compuesta y presentar sus distintos componentes en forma de un *espectro* de sonido. Un análisis de espectro es un gráfico en el que el eje horizontal representa la frecuencia y el eje vertical representa la amplitud. Por ejemplo, los numerosos componentes acústicos de la vocal [i:], que representan la resonancia producida en el tracto vocal durante la articulación, se pueden mostrar en forma de *espectrograma* de la siguiente manera:



La vocal [a:] proporciona un espectro distinto, lo que refleja una configuración completamente diferente del tracto vocal.



Nótese que, en estos espectros, la amplitud de algunas frecuencias es mucho mayor que la de otras. De hecho, en cada caso se pueden ver varios «picos» de energía acústica, que reproducen los principales puntos de resonancia del tracto vocal. Dichos picos se denominan *formantes*, y se numeran del más bajo al más alto: «primer formante» (F₁), «segundo formante» (F₂), etc. En el espectro de [i:], arriba (proferido por un hombre con una frecuencia fundamental de 120 Hz), F₁ alcanza 360 Hz, F₂ 2.280 Hz, y F₃ 3.000 Hz.

La estructura de los formantes es un rasgo fundamental de los sonidos del habla. Todas las vocales y algunas de las consonantes tienen formantes. El patrón de los formantes (especialmente la disposición de los dos primeros) es lo que nos permite distinguir las vocales, o reconocer que se está repitiendo la «misma» vocal, incluso cuando la pronuncian hablantes distintos. Además, los formantes de las vocales también pueden ayudar a identificar la naturaleza de los sonidos consonánticos contiguos.

La relación entre acústica y articulación

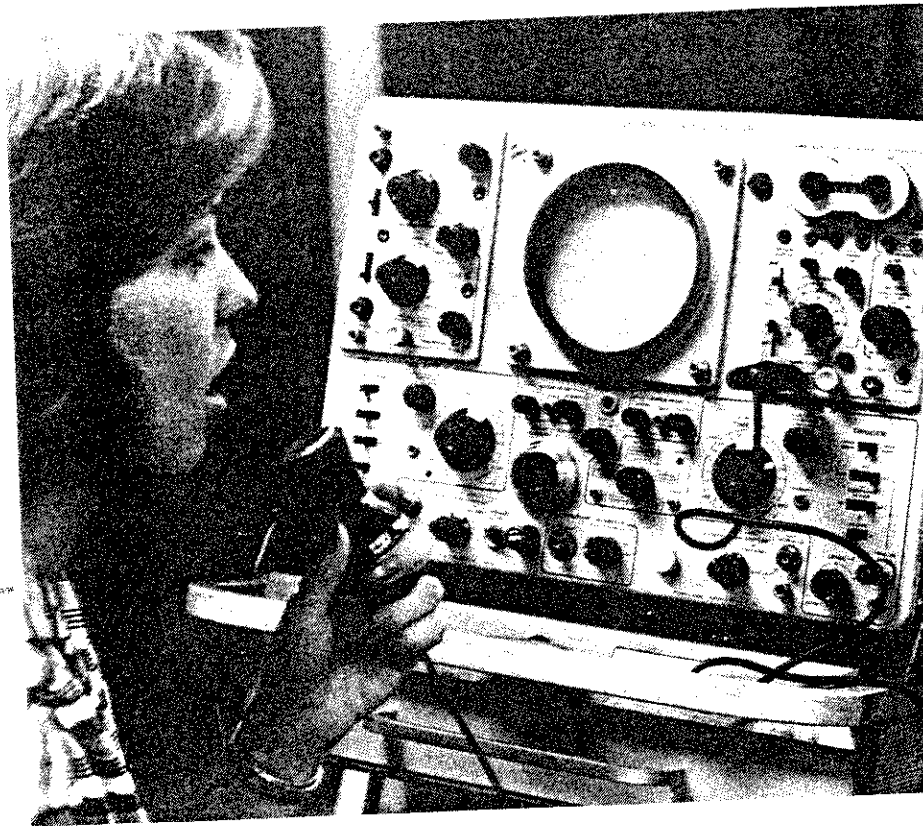
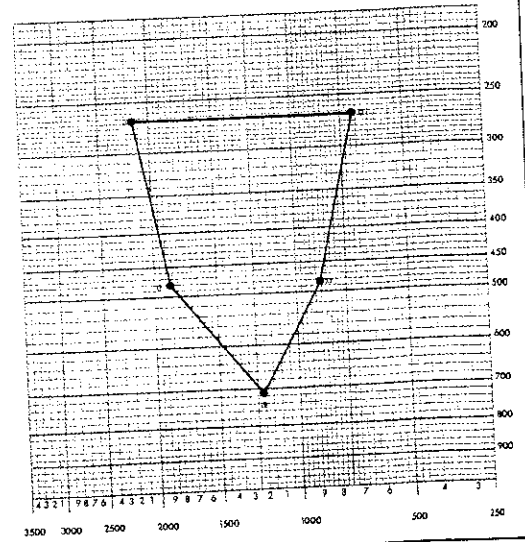
Estudiando a distintos hablantes, es posible calcular las principales frecuencias del primer y del segundo formante de las vocales de una lengua. Así se ha hecho también para el español, con los siguientes resultados (según E. Martínez Celdrán, 1984):

	F ₁	F ₂
[i]	267	2112
[e]	489	1889
[a]	711	1222
[o]	489	889
[u]	267	711

Podemos situar estas cifras en un gráfico en el que en el eje vertical aparece la frecuencia de F₁, y en el eje horizontal la de F₂. El patrón resultante es muy similar al que aparece cuando se describen las vocales de acuerdo con el punto de

articulación en el tracto vocal (véase el «trapecio» vocálico, pág. 154). No obstante, no coinciden de forma exacta, porque el esquema articulatorio se basa sólo en

el punto donde la oclusión de la lengua es mayor, mientras que el esquema acústico deriva de la resonancia del tracto vocal entero.



Analizador espectral de tiempo real. Este instrumento utiliza un tubo de rayos catódicos para mostrar el espectro en constante cambio de las ondas compuestas. Debido a la velocidad con que se suceden los cambios, para un análisis más profundo es necesario que el investigador «pare la acción», por ejemplo conservando la onda en un osciloscopio de almacenamiento, fotografiando la pantalla, o utilizando la pantalla de un ordenador.

EL ESPECTRÓGRAFO

Durante la década de 1940, se diseñó el espectrógrafo para analizar y presentar los espectros acústicos. Esta máquina graba la voz, analiza las diferentes frecuencias de las ondas sonoras mediante una serie de filtros electrónicos, mide la intensidad de cada frecuencia, y luego presenta el resultado de forma visual, utilizando una aguja para trazar las marcas sobre una tira de papel satinado especial. En esta página y en la siguiente presentamos una serie de espectrogramas que representan varios sonidos.

En esta clase de espectrogramas se muestran tres dimensiones del sonido.

1. El tiempo se representa horizontalmente; en el espectrógrafo más usado (el Sona-Graph Kay), se pueden grabar 2,4 segundos de habla sobre el papel, que se lee de izquierda a derecha. Cada media pulgada de papel (1,27 cm.) representa una décima de segundo de habla.
2. La dimensión vertical ofrece información sobre la frecuencia, desde 0 Hz (la línea inferior) hasta 8.000 Hz. Es una escala lineal: cada pulgada en vertical representa 2.000 Hz.
3. La tercera dimensión es la de la intensidad, indicada por el grado de oscuridad de las marcas sobre el papel. Cuanto más intensa es la señal, más oscuras son las marcas que deja la aguja. De esta manera, las frecuencias de intensidad muy baja o nula aparecen como zonas de papel blanco.

Si bien las limitaciones de esta máquina no permiten efectuar mediciones perfectas sobre el papel, el despliegue visual hace posible un reconocimiento inmediato para todo el que esté acostumbrado al análisis acústico, lo que ha convertido al espectrógrafo en un instrumento de investigación habitual. Se puede mostrar claramente el contraste entre sonidos distintos, al igual que la influencia de unos sonidos sobre otros cuando se mezclan en la cadena hablada. Además, en la actualidad se están haciendo grandes progresos en el diseño espectrográfico que permiten obtener análisis más exactos, gracias a los métodos informáticos. En particular, ahora se dispone de espectrógrafos digitales que pueden almacenar dos espectrogramas en pantalla, lo que permite una comparación más sistemática y más precisa de los detalles.

Tipos de espectrograma

Cuando se hace un espectrograma se puede elegir entre dos posibilidades. Si se elige la opción «estrecho», la máquina analizará la gama de frecuencias del habla en bandas pequeñas (normalmente de 45 Hz), lo que hace que los armónicos destaquen con gran claridad. Si la máquina se coloca en «ancho», el análisis se hará utilizando bandas de

frecuencia mucho más amplias (normalmente de 300 Hz), lo que hace que destaquen claramente los formantes. El análisis de banda ancha es más útil para los objetivos de la fonética, en la mayoría de los casos. Con el espectrógrafo también se puede obtener información sobre la amplitud de un sonido. La intensidad de cada componente de la frecuencia es analizada e

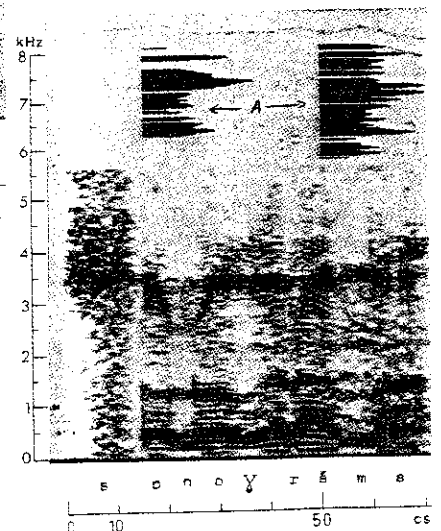
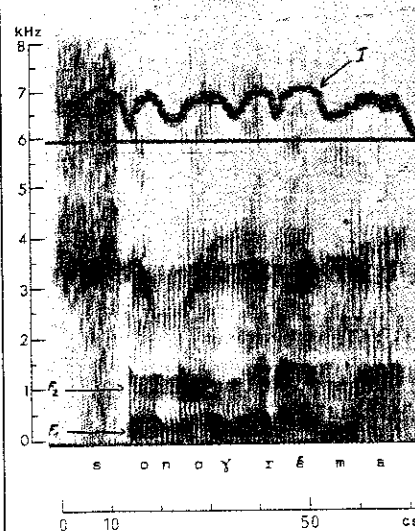
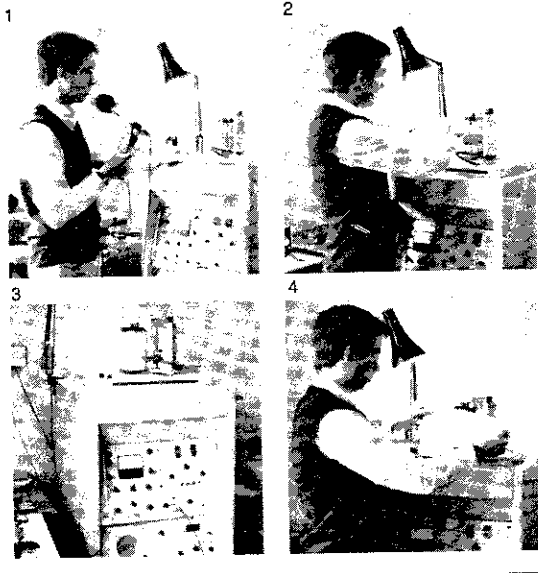
impresa en la parte superior del espectrograma, obteniéndose así una imagen de la amplitud. Abajo aparecen las versiones de banda estrecha y de banda ancha de la palabra «sonograma», la primera acompañada de una imagen de la intensidad, y la segunda de una imagen de la amplitud (según A. Quilis, 1981).

Cómo se hace un espectrograma

1. Se graban unas palabras en el espectrógrafo, bien mediante un micrófono, bien desde un magnetofón.
2. Se coloca un papel satinado especial alrededor del tambor que se encuentra sobre la máquina, y se pone una aguja en contacto con el papel.
3. Cuando se pone en

marcha la máquina, el tambor gira. Entonces la aguja sube por el tambor, analizando las distintas frecuencias del habla. Si en una frecuencia determinada hay energía, la aguja deja una marca en el papel.

4. Cuando la aguja llega al final del tambor, se desconecta la máquina y se quita el papel. El espectrograma está listo para ser interpretado.



Los rasgos acústicos de las vocales y las consonantes

Vocales Todos los sonidos vocálicos en el habla normal presentan dos o, normalmente, tres formantes, que en un espectrograma de banda ancha aparecen como espesas bandas negras. Se pueden apreciar claramente en los espectrogramas de las vocales largas [i:], [a:] y [u:], pronunciadas aisladamente. Las estrias verticales representan las vibraciones de las cuerdas vocales.

Semivocales y semiconsonantes. Los sonidos [j] y [w] (como en *bien* y *agua*) funcionan como consonantes en muchas lenguas, pero poseen los rasgos acústicos de las vocales [i] y [u] respectivamente; se les suele llamar «semiconsonantes» (pág. 152). Su naturaleza vocálica se puede observar claramente en un espectrograma en el que aparezcan articuladas entre dos vocales [a]. En ambos casos, los formantes se inclinan cuando la vocal cambia de timbre. Por ejemplo, en [aja] el primer formante se inclina hacia abajo y el segundo se inclina hacia arriba, cuando la lengua pasa de [a] a [i]; se vuelven a inclinar cuando la lengua vuelve a la posición original. La inclinación también afecta al tercer formante, pero afecta mucho menos al cuarto formante, que se encuentra más arriba.

Consonantes oclusivas Lo que habitualmente identifica a

una consonante oclusiva es un breve silencio, mientras en la boca se crea una oclusión, seguido de una explosión de sonido, cuando la oclusión se relaja (pág. 157). Los dos rasgos se pueden observar en un espectrograma, especialmente si dichas consonantes se articulan entre vocales. El silencio corresponde a la tira vertical de papel blanco y la relajación a una delgada «punta» de marcas, ampliamente esparcidas por el espectro. La aparición de las bandas negras de los formantes marca el comienzo de la siguiente vocal. Se puede observar claramente la diferencia entre las consonantes sordas [p, t, k] y sonoras [b, d, g] (pág. 128).

- En los sonidos sonoros la intensidad de la explosión de sonido es mucho menor que en los sonidos sordos.
- La duración del silencio es menor en los sonidos sonoros.
- Una variable importante es el tiempo que transcurre entre la relajación de la oclusiva y el comienzo de la vibración de las cuerdas vocales (representado por los formantes). Esta variable se conoce como «tiempo del comienzo de la voz» o «voice onset time» (VOT). Existe un hueco de VOT considerable en el caso de las oclusivas sordas (indicado en los espectrogramas con una X).

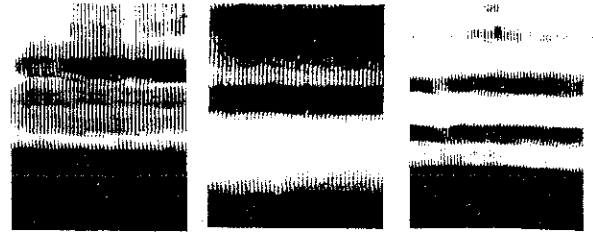
en las oclusivas sonoras el hueco

es muy pequeño o inexistente. En tales casos, la sonoridad puede comenzar antes que la explosión de sonido. En los espectrogramas también se pueden ver los distintos puntos de articulación de las consonantes oclusivas. Los sonidos bilabiales [p, b] presentan una explosión sonora en las frecuencias bajas; los sonidos alveolares [t, d] muestran dicha explosión en las frecuencias altas, y los sonidos velares [k, g] la tienen en la zona intermedia. También hay una clara diferencia en el punto de transición entre consonante y vocal: en dicho punto se produce un rápido movimiento articulatorio, que se refleja en la brusca inclinación de los formantes donde comienza la vocal, inclinación que cambia de dirección entre cada par de consonantes.

Consonantes fricativas

Estas consonantes utilizan una energía acústica casual, o ruido, representada en el espectrograma por una ancha zona de perturbación en ciertas frecuencias. Donde se aprecia con más claridad es en el caso de las sibilantes fricativas, como [s] y [ʃ], que son sonidos con una energía elevada. La energía de la [s] está muy por encima de los 4.000 Hz, mientras que la de [ʃ] empieza más abajo, alrededor de los 2.500 Hz. Aquí presentamos el contraste de estos sonidos entre vocales [a].

Vocales

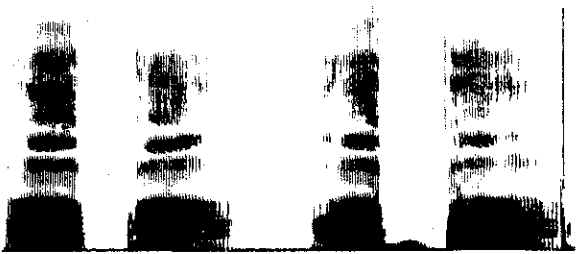


a: i: u:

Semivocales y semiconsonantes



a j a a w a



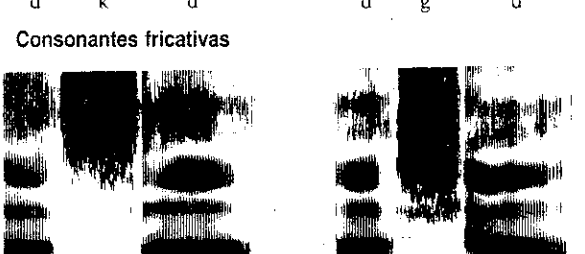
a p X a a b X a



a t X a a d X a



a k X a a g X a



a s a a ʃ a

La cadena hablada

Los espectrogramas de la derecha corresponden a unidades de sonido articuladas con cuidado, de forma aislada. Pero en la cadena hablada, los sonidos se producen con mayor rapidez, influyen unos en otros, y resulta más difícil distinguir los límites que los separan. Los espectrogramas de la cadena hablada ponen el acento en la continuidad fundamental de la lengua hablada, en el hecho de que la articulación es un proceso de cambio continuo. Es algo que siempre habría que tener en cuenta cuando se leen los símbolos, claros y aislados, de la transcripción fonética. (El ejemplo es de Martínez Celdrán, 1984, pág. 327).

«Te deseo felicidad»