

EL METODO ESPECTROGRAFICO

(Notas de Fonética experimental)

Sólo pretendemos dar, con estas notas, una explicación del método espectrográfico aplicado a la investigación fonética y lingüística.

Durante muchos años han prevalecido en materia fonética las descripciones fisiológicas o articulatorias (genéticas, en la terminología de Panconcelli-Calzia). Tuvieron el gran interés de definir, para los primeros lingüistas, la formación de los sonidos atendiendo a sus características más esenciales (punto de articulación, modo de articulación, sonoridad, etc.). La otra rama de la fonética, la acústica (genémica, según el citado fonetista), se ha desarrollado con más lentitud. La Fonética acústica no fue olvidada por los antiguos fonetistas, pero tuvo un desarrollo mucho menor que la fisiológica. Esto fue debido a que los estudios fisiológicos avanzaron mucho más rápidamente que los de la acústica.

Las investigaciones actuales de fonética deben hacerse teniendo en cuenta los dos tipos de descripción de los sonidos: el articulatorio y el acústico. Como dice Bertil Malmberg¹, era errónea la opinión de toda la fonética clásica de que cada posición diferente de la lengua daba lugar a un sonido nuevo. Hoy podemos comprobar que en la caja de resonancia bucal se compensa la variación de un órgano «A» que producía antes de su movimiento un sonido «X», con la modificación de otro órgano «B» para dar un resultado acústico análogo al primitivo «X».

A partir de las experiencias de Helmholtz, con sus resonadores, en 1856, y, más tarde, de las del abad Rousselot, de Gemelli y Pastori, con sus métodos de análisis quimográficos y electroacústicos, respectivamente, se marcaron nuevos avances en esta rama tan importante de la Lingüística.

En los años que han seguido a la segunda guerra mundial, el pro-

¹ BERTIL MALMBERG, *Le problème du classement des sons du langage* (*Studia Linguistica*, 1952, VI, pp. 1-56). Tomando por base las teorías de clasificación de Forchhammer, estudia el problema del agrupamiento de los sonidos en «tipos articulatorios» y en «tipos acústicos», siendo éstos, en su opinión, los más netamente diferenciados y los más estables.

greso de la técnica ha sido gigantesco en todos los órdenes, y, lógicamente, este progreso también afectó a los estudios experimentales de fonética. Nuevos instrumentos han ido completando los ya existentes y casi se puede decir que han iniciado una nueva era, completamente revolucionaria, en estos estudios. Desde que Potter, Kopp y Green publicaron su libro *Visible Speech*, en 1947¹, multitud de científicos se lanzaron a profundizar en las aplicaciones espectrográficas del lenguaje. Y no han sido sólo lingüistas los que se han interesado por estas experiencias, sino que también físicos, psicólogos, etc., han aportado interesantes consecuencias, producto de sus investigaciones, al interesante campo del lenguaje.

Poco después de iniciarse los estudios del método espectrográfico, una nueva técnica ha permitido componer el lenguaje sintético, ha hecho posible la creación artificial de la palabra por medio de la mano del hombre. Pero de este procedimiento hablaremos más detenidamente en otras notas, ya que está desempeñando un papel muy importante en la investigación fonética.

Este método de análisis del lenguaje descompuesto automáticamente en sus componentes se llevó a la realidad en los Bell Telephone Laboratories. El espectrógrafo se ideó con la intención de que los sordos pudiesen leer perfectamente, sobre una pantalla móvil, una conversación telefónica descompuesta, por medio de filtros, en los formantes característicos de cada uno de los fonemas. Este fue su objeto principal, tal y como viene indicado en las primeras páginas del libro *Visible Speech*, donde numerosas fotografías nos muestran los grupos de experimentación con los que se empezó a trabajar.

Al poco tiempo de ver que las pruebas realizadas con el espectrógrafo daban resultado satisfactorio, la Kay Electric Company comenzó a construir este aparato. Los espectrógrafos construidos por esta casa, antes de 1948, alcanzaban una frecuencia de sólo 3.500 c.p.s. (ciclos por segundo), lo que hacía difícil el estudio completo de muchas consonantes, cuyos rasgos espectrográficos comienzan a esa frecuencia. Más tarde, después del referido año, la escala de frecuencias espectrográficas fue aumentada hasta 8.000 c.p.s., con lo que el estudio de dichos sonidos puede hacerse ya mucho más completo.

En la figura 1 podemos ver un esquema muy simplificado del espec-

¹ POTTER, KOPP y GREEN, *Visible Speech*, Nueva York, Van Nostrand, 1947. Primer libro en el que apareció la aplicación del espectrógrafo al lenguaje. Multitud de espectrogramas muestran los rasgos acústicos de cada sonido. Libro fundamental para la iniciación en este método.

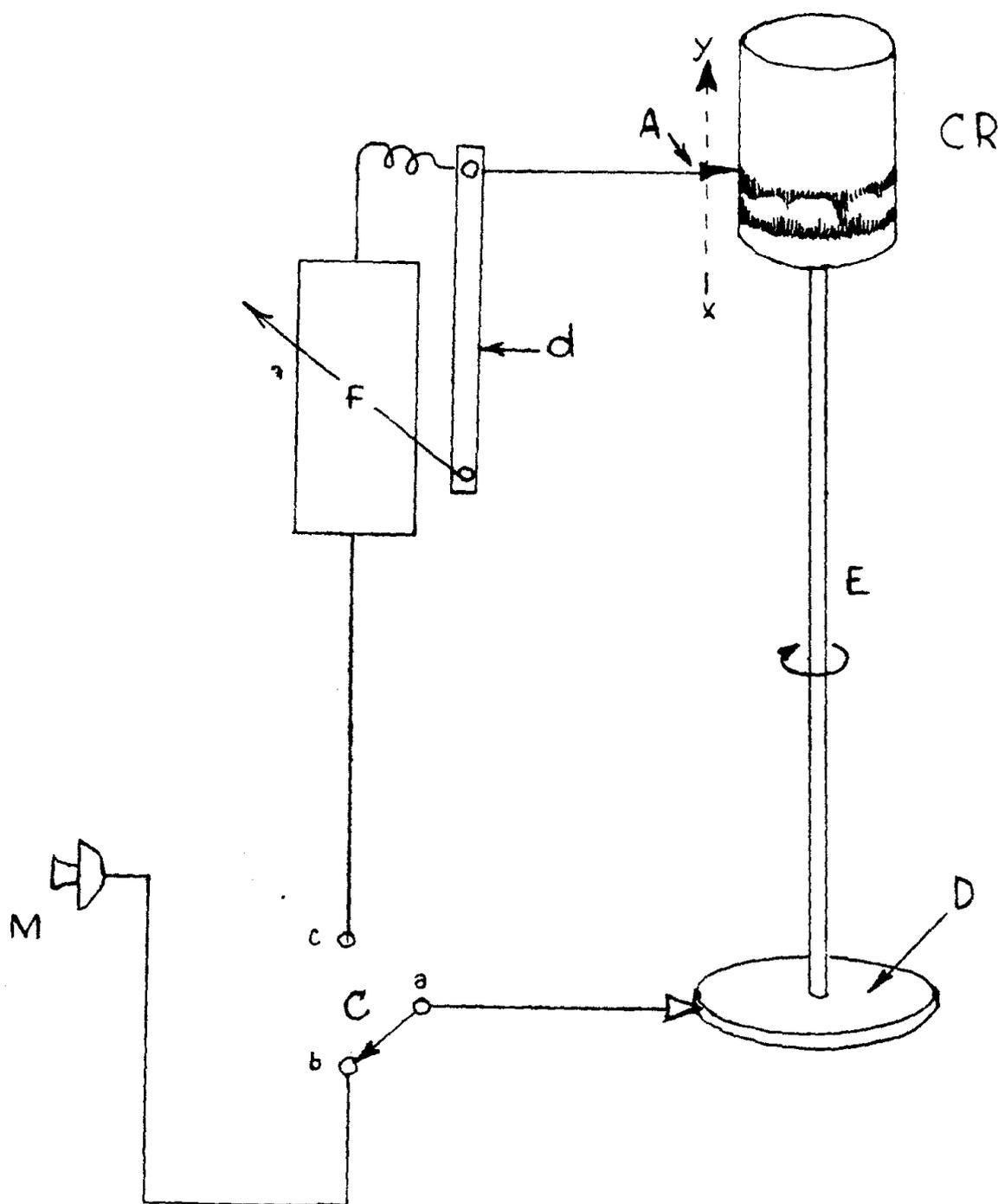


FIG. I

M Micrófono,
 C Conmutador.
 D Disco.
 F Filtro.

A Aguja inscriptora.
 CR Cilindro reproductor.
 E Eje.
 d Corredera.

trógrafo. Las partes más importantes del aparato van acompañadas de unas iniciales que permiten dirigirnos a ellas con facilidad, a lo largo de la explicación.

Esquema y funcionamiento del aparato.—Las tres partes más esenciales que constituyen el aparato son: el sistema de filtros, F; el disco magnético, D, y el cilindro reproductor, CR. Como es natural, el espectrógrafo lleva anejos otra serie de mecanismos, como el amplificador, alimentador, motor de rotación, etc., que no aparecen en la figura. Cuando el conmutador C se halla en la posición *a-b*, el aparato se encuentra a punto de poder efectuar una grabación sobre la superficie cilíndrica externa del disco magnético D.

Para efectuar esta grabación podemos emplear dos procedimientos: 1.º directamente, hablando frente al micrófono del espectrógrafo; 2.º grabando antes sobre una cinta magnetofónica y haciendo pasar esta grabación por el mismo conducto *a-b* al disco magnético D. Este segundo procedimiento es el más empleado y el más recomendable, ya que así nos queda siempre, archivado, el control auditivo de lo analizado.

Ya tenemos el problema grabado sobre el disco magnético. La duración máxima que podemos analizar en cada espectrograma es de 2,4 segundos. Una vez que hemos llegado al final de este tiempo, donde hemos calculado que terminaba la frase o la palabra para analizar, desconectamos el conmutador C de su posición *a-b*; de no hacerlo así, volveríamos a grabar encima, borrando lo anterior. Al hacer la desconexión, el conmutador C pasa automáticamente de la posición *a-b*, que tenía cuando hicimos la grabación, a la posición de reproducción *a-c*; de esta manera queda en comunicación el disco magnético con el cilindro reproductor CR. Sobre este cilindro buscamos el comienzo de la palabra o del conjunto de palabras que están grabadas sobre el disco magnético D. Podemos oír perfectamente la grabación a través de un altavoz que lleva el aparato. Calculado el punto cero, punto en el que interrumpimos la grabación sobre el disco magnético, y que, por tanto, es el principio y el final del espectrograma, colocamos alrededor del cilindro reproductor CR un papel especial, sensible, fabricado por la misma casa Kay Electric, compuesto de una substancia conductora de la electricidad. Ponemos en marcha el aparato mediante un motor colocado debajo del disco magnético D. Este y el cilindro reproductor CR, que van unidos por el eje E, giran simultáneamente según la dirección de la flecha. Al estar en movimiento el conjunto disco magnético-cilindro reproductor, la energía de la palabra grabada sobre el disco magnético pasa por los puntos de conexión *a-c*, a través del sistema de filtros, hacia la aguja inscriptora A. Esta se desliza por la corredera *d*, movida por un tornillo

sin fin, según la dirección $x-y$, paralela a la altura del cilindro. Al poner en contacto la aguja inscriptora A con el papel que está alrededor del cilindro reproductor CR, como cada uno de ellos —aguja y cilindro reproductor— actúa de polo de distinto signo, salta la descarga eléctrica desde la aguja A al cilindro reproductor a través del papel conductor. La aguja, en su camino ascendente, va trazando un helicoides sobre el papel, y lo va quemando en aquellos puntos donde hay energía. De esta manera va dejando en el papel, a sus frecuencias determinadas, el espectro o imagen del sonido.

Filtros.—Dijimos más arriba que, al conectar el conmutador C con el punto c , pasaba la energía acumulada en el disco magnético a través del sistema de filtros F, hasta la aguja inscriptora A.

Un filtro acústico es un sistema de transmisión pasiva que sólo permite el paso de una determinada gama de frecuencias a través de él. O de otra manera: que sólo es capaz de dejar pasar aquellas frecuencias para las que es sensible. De este modo, frecuencias superiores o inferiores a su capacidad quedan sin pasar.

La onda acústica de la palabra es un sonido complejo que necesitamos analizar para ver sus tres componentes más importantes: número, amplitud y frecuencia de las vibraciones. Este análisis, como dice Bertil Malmberg¹, puede hacerse por tres procedimientos: 1.º Por el estudio de un análisis matemático de la curva compleja, según el teorema de Fourier, que nos enseña que cualquier onda compuesta puede ser descompuesta en un número de curvas sinusoidales, múltiplos de la fundamental. 2.º Con la ayuda de un filtro acústico, y 3.º Por el oído, lo que exige que sea sumamente sutil para poder separar cada uno de los componentes que integran el conjunto total del sonido.

El primer procedimiento es largo y complicado; el tercero, muy difícil, casi imposible de conseguir. Nos queda como más viable el segundo, el de los filtros acústicos, que, por la nitidez con que se presentan los armónicos múltiplos del fundamental, pueden ser objeto de un estudio exacto y no muy complicado. Del análisis de sus resultados trataremos más adelante.

Estos sistemas de filtros son, en realidad, el elemento más importante del método espectrográfico. Porque un espectro es, en realidad, el resul-

¹ B. MALMBERG, *La Phonétique*. Col. *Que sais-je?*, P. U. F., 1960, 2.ª edición. Pequeño manual de fonética, donde, sin abandonar el aspecto tradicional, considera la fonética y sus problemas bajo el punto de vista de los más recientes hallazgos.

tado de la dispersión de un conjunto de radiaciones, o de un conjunto de ondas, hablando en términos acústicos. Pues bien, este sistema de filtros no tiene otro objeto que el de dispersar, según unos módulos constantes, la energía acústica que llega a ellos.

En el espectrógrafo de la Kay Electric hay dos sistemas de filtros: uno, constituido por una serie que no deja pasar cada uno de ellos nada

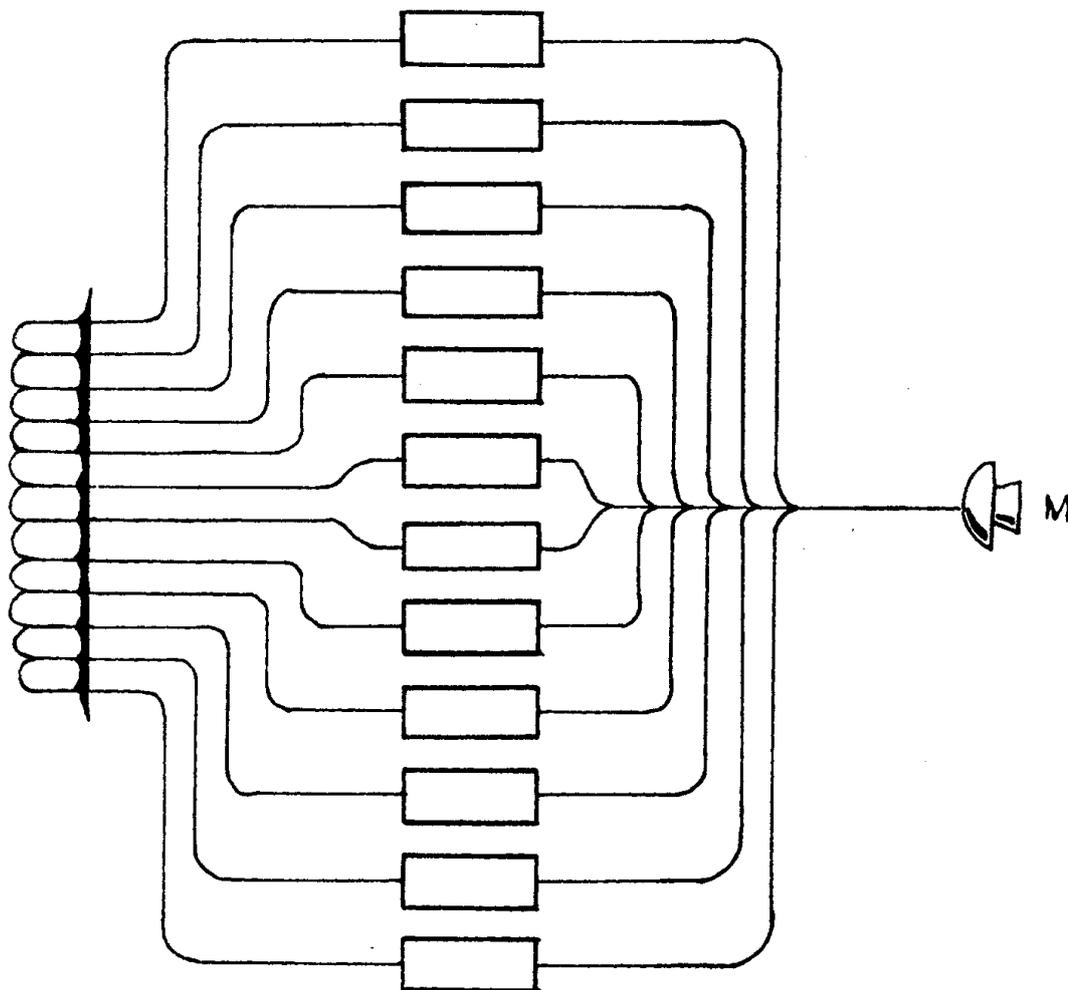


FIG. 2

más que frecuencias de 45 c.p.s., y que es el que da origen a los espectrogramas llamados de banda estrecha; y otro, en los que cada uno de los filtros tiene un paso de frecuencias mucho más amplio: 300 c.p.s. Este último da origen a los espectrogramas llamados de banda ancha.

De esta manera, la onda acústica, transformada ya en onda eléctrica por la acción del micrófono, llega frente a la muralla de filtros. Cada filtro está colocado a una frecuencia determinada; dejará pasar sólo

aquel número de frecuencias que están en su región y comprendidas dentro de sus límites: 45 c.p.s. o 300 c.p.s., según utilicemos el selector de banda estrecha o el de banda ancha. De esta manera, si suponemos que el límite superior de un filtro de banda estrecha está colocado a 2.000 ciclos por segundo, sólo pasarán a través de él las frecuencias comprendidas entre 1.955 c.p.s. y 2.000 c.p.s. Si en vez de utilizar el filtro de banda estrecha empleamos el de banda ancha, y lo suponemos también situado en la misma región, y con el mismo límite superior, entonces pasará una gama de frecuencias comprendidas entre 1.700 c.p.s. y 2.000 c.p.s.

Así, la energía que llega al sistema de filtros es descompuesta en una serie de zonas de frecuencias que se transmiten a la aguja inscriptora A. Esta, cuando llega a la región de cada una de esas frecuencias, efectúa una descarga eléctrica que deja una marca sobre el papel sensible enrollado alrededor del cilindro reproductor CR.

En la figura 2 tenemos el esquema de un sistema de filtros del *Visible Speech*. La energía que es descompuesta por estos filtros pasa a reunirse de nuevo sobre el espectrograma o sobre una pantalla y así se pueden ir leyendo los componentes acústicos del sonido. Recientemente H. Sund, en Estocolmo, ha ideado un nuevo procedimiento, basado también en la misma técnica del espectrógrafo. En vez de tener una limitación en el tiempo, como requiere la obtención de un espectrograma, cuando hay que analizar cantidades mayores de sonidos es más conveniente emplear otro método más rápido y de más bajo costo. De esta manera, el sonido descompuesto por medio de los filtros se refleja en una pantalla. Una cámara fotográfica agregada al espectrógrafo recoge sucesivamente todas las imágenes del sonido. Además, en este procedimiento ha sido aumentada la frecuencia de los sonidos hasta 10.000 c.p.s.

Espectrogramas.—El producto de la descomposición de un sonido o grupos de sonidos en sus componentes acústicos es lo que recibe el nombre de espectrograma. En las figuras 3 y 4 podemos ver dos espectrogramas obtenidos en la Sección de Fonética del C. S. I. C. de Madrid. Es la misma frase, *Dulcinea del Toboso*, pronunciada por el autor, en sus dos realizaciones: con el filtro de banda ancha y con el de banda estrecha, respectivamente. Estos espectrogramas alcanzan una frecuencia de 6.000 c.p.s., el de banda ancha, y de un poco más de 4.500 c.p.s. el de banda estrecha. No se ha dejado que alcanzasen los 8.000 c.p.s. con el objeto de mostrar en el mismo espectrograma otras posibilidades del método.

Es preferible emplear el filtro de banda estrecha para obtener las «secciones» de los sonidos sonoros, para la descomposición en armónicos

y para el estudio de los cambios de melodía. El filtro de banda ancha, por el contrario, es muy bueno para el estudio de los sonidos sordos, y es también mejor para ver las intensidades, así como el juego de los formantes.

Como ya dijimos, el espectrograma puede analizar una duración máxima de 2,4 segundos, en una extensión lineal de unos 31 cm. El límite máximo de frecuencia que alcanza es de 8.000 c.p.s., también llamados Hertz (Hz), o su múltiplo Kilohertz (KHz), equivalente al Kilociclo. Esta escala de frecuencias tiene una altura sobre el espectrograma de unos 9,7 cm. El tiempo empleado en obtener un espectrograma es de unos cinco minutos; a este tiempo hay que añadirle, claro está, el empleado en hacer la grabación.

Un espectrograma nos ofrece a primera vista, inmediatamente, tres representaciones: 1.^a Sobre el eje de abscisas, y de izquierda a derecha, podemos medir la duración total del espectrograma y la de cada uno de los sonidos. 2.^a En el eje de ordenadas, la escala de frecuencias. 3.^a En el mayor o menor negror de los formantes vemos la intensidad con que se pronuncia un sonido con relación a otro. Así, en la figura 3, se nota el menor grado de negror de la /o/ final de *Toboso*, con relación a la tónica o a la protónica de la misma palabra. En esta frase, que ha sido pronunciada procurando matizar bien todos los sonidos, sólo aparecen dos vocales relajadas: la /i/ de *Dulcinea* y la ya referida /o/ final de *Toboso*, en las que se ve menor ennegrecimiento, particularmente del segundo formante. También se nota esta relajación en una separación mayor de las líneas verticales que indican las vibraciones de las cuerdas vocales: nótese cómo en las tónicas la separación es menor.

En la figura 3 aparece la descomposición del sonido en sus **f o r m a n t e s** o conjunto de frecuencias características del timbre de un sonido.

Cada uno de los formantes no es un tono único, sino una zona en la que se pone de relieve un conjunto de armónicos, cada uno situado a una frecuencia determinada. Por eso se suele llamar también «zona de formantes». De ahí que en los espectrogramas de banda estrecha, en los que el sonido está descompuesto en sus armónicos, haya que hacer, cuando no se está habituado, una pequeña extrapolación, con el fin de localizar el conjunto de armónicos que darán lugar al formante. La medición de las frecuencias en las vocales se hará tomando el punto medio de la zona de formantes, en relación a la línea cero, línea de referencia más baja.

Uno de los avances más importantes que supuso para la Fonética

el método espectrográfico fue, como dice Pierre Delattre ¹, el papel de las transiciones de los formantes en la percepción de las consonantes: el quimógrafo e incluso el oscilógrafo no dejaban ver, en las consonantes oclusivas, por ejemplo, otra cosa que una interrupción seguida de un ruido de explosión. El espectrógrafo ha puesto de relieve los cambios rápidos de frecuencias que reflejan los movimientos articulatorios que unen el centro de la consonante con el centro de la vocal. La comprobación de los datos que nos proporciona el espectrógrafo ha sido realizada con pleno éxito por medio del lenguaje sintético. Además de estas transiciones en la percepción consonántica, puede verse perfectamente el juego de los formantes en el paso de un sonido vocálico a otro.

En los espectrogramas que acompañamos se ve claramente la sonoridad o sordera de los sonidos consonánticos, en nuestro caso, por ejemplo, entre la /d/ y la /t/, por la presencia o ausencia de un formante bajo. Cuando hay una completa cerrazón de la región bucal, no aparecen formantes altos que son los característicos de las vocales. En estas condiciones sólo pasa a través de las paredes de la boca un sonido pequeño que procede de las vibraciones de las cuerdas vocales. Este sonido aparece reflejado en el espectrograma en una región de frecuencia muy baja que no llega al punto cero. Esta «barra de sonoridad», que aparece, como ya se ha indicado, sólo para las consonantes sonoras, es su primer formante. Aunque en el espectrograma no aparezcan más formantes, sin embargo, podríamos encontrar otros formantes más altos para las oclusivas /b/, /d/, /g/, aplicando las «secciones» en un punto determinado de su dimensión temporal.

La representación del sonido /l/ de las palabras *Dulcinea* y *del* está constituida también por una serie de formantes de características muy similares a los vocálicos. Por eso Jakobson, al hacer la clasificación de los fonemas, forma un núcleo aparte con las líquidas por participar de los rasgos consonánticos y vocálicos. Los formantes altos de la /l/ varían según la posición adoptada para la articulación del sonido siguiente. El primer formante de este sonido aparece a una altura mayor que los de las consonantes nasales /m/ y /n/, que son iguales. Las frecuencias de estos formantes también son más elevadas que las de las oclusivas /b/ y /d/, también iguales. El orden, de menor a mayor frecuencia, para estas consonantes es el siguiente: /b/ y /d/; /m/ y /n/; /l/.

En las vocales, según las investigaciones hasta ahora realizadas, los

¹ PIERRE DELATTRE, *Les indices acoustiques de la parole: Premier rapport (Phonética, 1958, pp. 108-118 y 226-251)*. Relación detallada de todas las investigaciones realizadas, hasta esa fecha, en el estudio de las características acústicas de los sonidos. Metodológicamente, está referido más bien a la síntesis del lenguaje.

dos primeros formantes son suficientes para caracterizar su timbre; en algunos casos también se tiene en cuenta el tercer formante. En las vocales anteriores, el segundo formante es alto y está colocado, por tanto, próximo al tercer formante. En la percepción de estas vocales, como dice Delattre, hay una equivalencia relativa entre estos dos formantes más próximos y un solo formante que tuviese una frecuencia media entre los dos. Por el contrario, en las vocales posteriores, en las que el segundo formante está bastante alejado del tercero y muy próximo al primero, para la percepción sería equivalente a un formante que tuviese una frecuencia media entre los dos.

Basado en este grado de separación o acercamiento de los formantes vocálicos es como se ha llegado a un rasgo distintivo de las vocales al considerarlas en oposición según su grado de densidad o difusión; llamando, a las que tienen un mayor acercamiento entre sus dos primeros formantes, «densas» o «compactas», por oposición a las «difusas» en las que la separación de las zonas de formantes es mayor.

Cada uno de los formantes vocálicos, o de las modificaciones que aparecen en ellos, no son debidos a un solo órgano de la cavidad bucal, sino al conjunto de las variaciones que pueden sufrir los diversos órganos de la palabra. Fue primero Pierre Delattre^{1,2} y, más recientemente Gunnar Fant³, los que más han trabajado para encontrar la relación exis-

¹ P. DELATTRE, *Un triangle acoustique des voyelles orales du français* (*The French Review*, 1948, XXI). Partiendo de los resultados acústicos proporcionados por el espectrógrafo, construye un triángulo, mejor un pentágono, de las once vocales orales francesas. Inicia el estudio de las relaciones acústicas y fisiológicas.

² P. DELATTRE, *The Physiological Interpretation of Sound Spectrograms* (*PMLA*, 1951, LXVI). Demuestra las relaciones que existen entre los tres primeros formantes y la posición de los órganos bucales.

³ GUNNAR FANT, *Modern Instruments and Methods for Acoustic Studies of Speech* (*Acta Polytechnica Scandinavica, Physics Including Nucleonics Series*, 1, Oslo, 1958). Resumen breve y claro de todos los más recientes métodos acústicos para la investigación fonética: espectrografía, oscilografía y los adelantos hechos en la síntesis del lenguaje. También expone las últimas investigaciones realizadas para hallar una relación entre la fisiología y la acústica. Otros trabajos también muy interesantes son: ROMAN JAKOBSON, GUNNAR FANT y MORRIS HALLE, *Preliminaries to Speech Analysis. The Distinctive Features and their Correlates* (*Technical Report*, núm. 13, mayo de 1952, Massachusetts Institute of Technology). Primer ensayo sobre un estudio de las distinciones acústicas de los sonidos partiendo de los nuevos avances electroacústicos. B. MALMBERG, *Questions de méthode en phonétique synchronique* (*Studia Linguistica*, 1956, X, pp. 1-44). Presentación de una serie de problemas lingüísticos y fonéticos y métodos actuales para su investigación. Una bibliografía pormenorizada de cada uno de los problemas fonéticos puede encontrarse en: DELATTRE, *Les indices...*, y en FANT, *Modern Instruments...*

tente entre los resultados acústicos obtenidos con el método espectrográfico y las posiciones articulatorias. De sus trabajos podemos deducir varios puntos interesantes:

a) El primer formante guarda una estrecha relación con la abertura del canal bucal: cuando la abertura es máxima, esto es, cuando la lengua está más separada del velo del paladar, la frecuencia de dicho formante es la más elevada; por el contrario, si la lengua se va acercando más al paladar, la abertura vocálica decrece, y la frecuencia del formante también disminuye. Así, en los espectrogramas que acompañamos puede verse cómo el primer formante de la vocal /a/ es el que tiene mayor altura, mayor frecuencia. Los de las vocales /e/ y /o/ tienen aproximadamente la misma frecuencia. Igual ocurre con los de la /i/ y la /u/. Por tanto, podemos decir que el orden que siguen las vocales de mayor a menor frecuencia del primer formante es el siguiente: /a/; /e/ y /o/; /i/ y /u/.

b) El segundo formante puede sufrir modificaciones por dos factores principales:

1.º Por la posición de la lengua: cuanto más elevada se halla situada y más anterior es su posición, más alta es su frecuencia; por el contrario, cuanto más posterior sea su situación y también más elevada, su frecuencia es menor. Mirando nuestros espectrogramas (es más fácil consultar el de banda ancha, fig. 3) nos damos cuenta del descenso del segundo formante desde la /i/ en que la posición de la lengua es la más elevada y la más anterior, hasta la /u/, en la que la lengua ocupa la posición más posterior.

Conjugando estos dos puntos que llevamos descritos, es lógico pensar que, cuanto más elevada sea la posición de la lengua, o sea, cuanto más estrecha sea la abertura vocálica y la posición de la lengua sea también más anterior, el primer formante descenderá y el segundo se elevará; así, podemos encontrar acústicamente algunos rasgos distintivos entre, por ejemplo, /i/, /j/ e /y/: para la articulación de la semiconsonante la lengua ha elevado su posición y la abertura bucal es menor que para /i/; este movimiento se traduce acústicamente en una elevación del segundo formante y en un descenso del primero. Si la lengua se eleva todavía más, la cerrazón del canal bucal es mayor, y sobreviene la fricación; se pierden ya todos los rasgos vocálicos y aparecen los consonánticos; esta pérdida se muestra acústicamente bien por una elevación y debilitamiento, o por casi una desaparición del segundo formante; también se reduce el primero a la «barra de sonoridad» de todas las consonantes sonoras.

2.º Cuanto más redondeados y más abocinados se encuentren los

labios, más baja es la frecuencia del segundo formante. Se comprueba, por ejemplo, pronunciando la vocal /i/ y luego labializándola hasta obtener el sonido de la /ii/ francesa. (Hemos de advertir antes de seguir más adelante que todos estos puntos los hemos comprobado con el espectrógrafo.)

De estos dos apartados se puede deducir que cuanto más alargada sea la cavidad bucal, más baja es la frecuencia del segundo formante.

c) El tercer formante varía con los movimientos del velo del paladar. Cuando éste desciende, tomando una posición análoga a la adoptada para la producción de las vocales nasales, la frecuencia del tercer formante se eleva, y viceversa.

En las consonantes fricativas sonoras /b/, /d/ que aparecen en los espectrogramas, al igual que la /g/, observamos que no sólo aparece el primer formante o «barra de sonoridad», sino una serie de ellos, más altos, que son, además, como una continuación de los formantes vocálicos. El que aparezcan más o menos formantes altos en estos sonidos se debe a la mayor o menor abertura del canal bucal; es evidente que, cuanto más tienda a cerrarse este canal, más se acerca su resultado acústico a la oclusión, y, por tanto, sus formantes van desapareciendo hasta quedar reducidos a la barra de sonoridad, como cuando sobreviene la oclusión.

En las consonantes fricativas sordas se nota una diferencia entre el sonido /s/ y el /θ/. El sonido /θ/ empieza a tener mayor densidad hacia los 4.500 c.p.s., mientras que el sonido /s/ comienza a unos 3.000 c.p.s. aproximadamente. Por otra parte, la fricativa interdental tiene mayor duración que la alveolar. La representación espectrográfica de la fricación del sonido /s/ empieza en la línea cero, porque no ha sido pronunciado de manera tan apical como el del castellano.

Aunque todavía no conocemos perfectamente la estructura acústica de algunas consonantes, espectrográficamente, las podemos clasificar, de igual modo que las vocales, en «densas» o «compactas» y «difusas», según la mayor o menor densidad de sus formantes. Este grado de densidad o difusión, como ya sabemos, depende de la cavidad resonadora. De esta manera las consonantes velares y palatales, densas, se oponen a las dentales y labiales, difusas. Así se ve en los sonidos /s/ y /θ/ de nuestros espectrogramas. También, por ejemplo, los sonidos /t/, /d/, /p/, /b/, por ser dentales y labiales, son difusos, mientras que /k/ y /g/ son compactos.

El espectrograma de la figura 4 es el que se obtiene descomponiendo el sonido en sus armónicos. Este es el procedimiento más adecuado para calcular la frecuencia del tono fundamental. Como ya sabemos, en virtud del teorema de Fourier, que cada uno de estos armónicos es un múltiplo

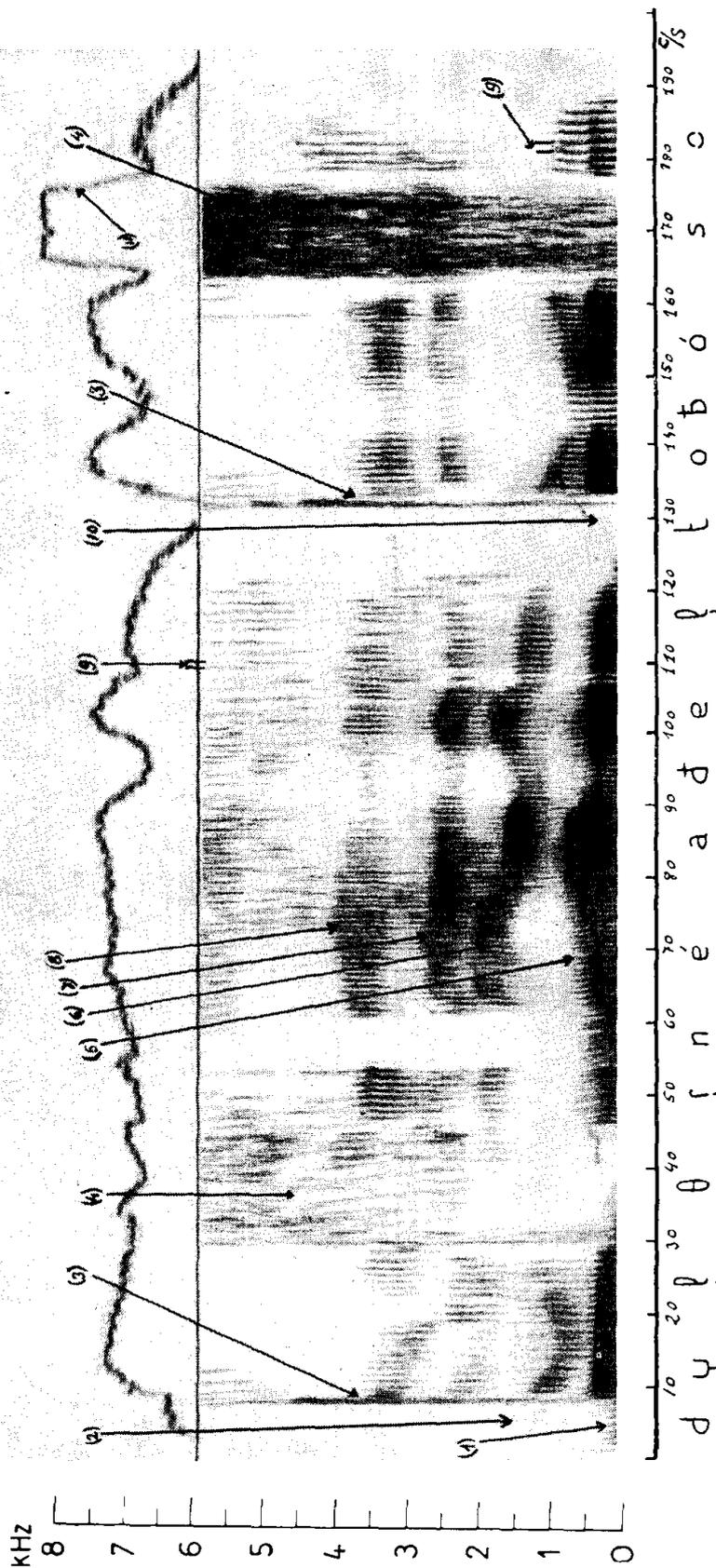


Fig. 3

- | | |
|-------------------------|---------------------------------------|
| (1) Barra de sonoridad. | (6) Segundo formante. |
| (2) Oclusiva sonora. | (7) Tercer formante. |
| (3) Barra de explosión. | (8) Cuarto formante. |
| (4) Fricación sorda. | (9) Vibración de las cuerdas vocales. |
| (5) Primer formante. | (10) Oclusiva sorda. |
| | (11) Línea de intensidad. |

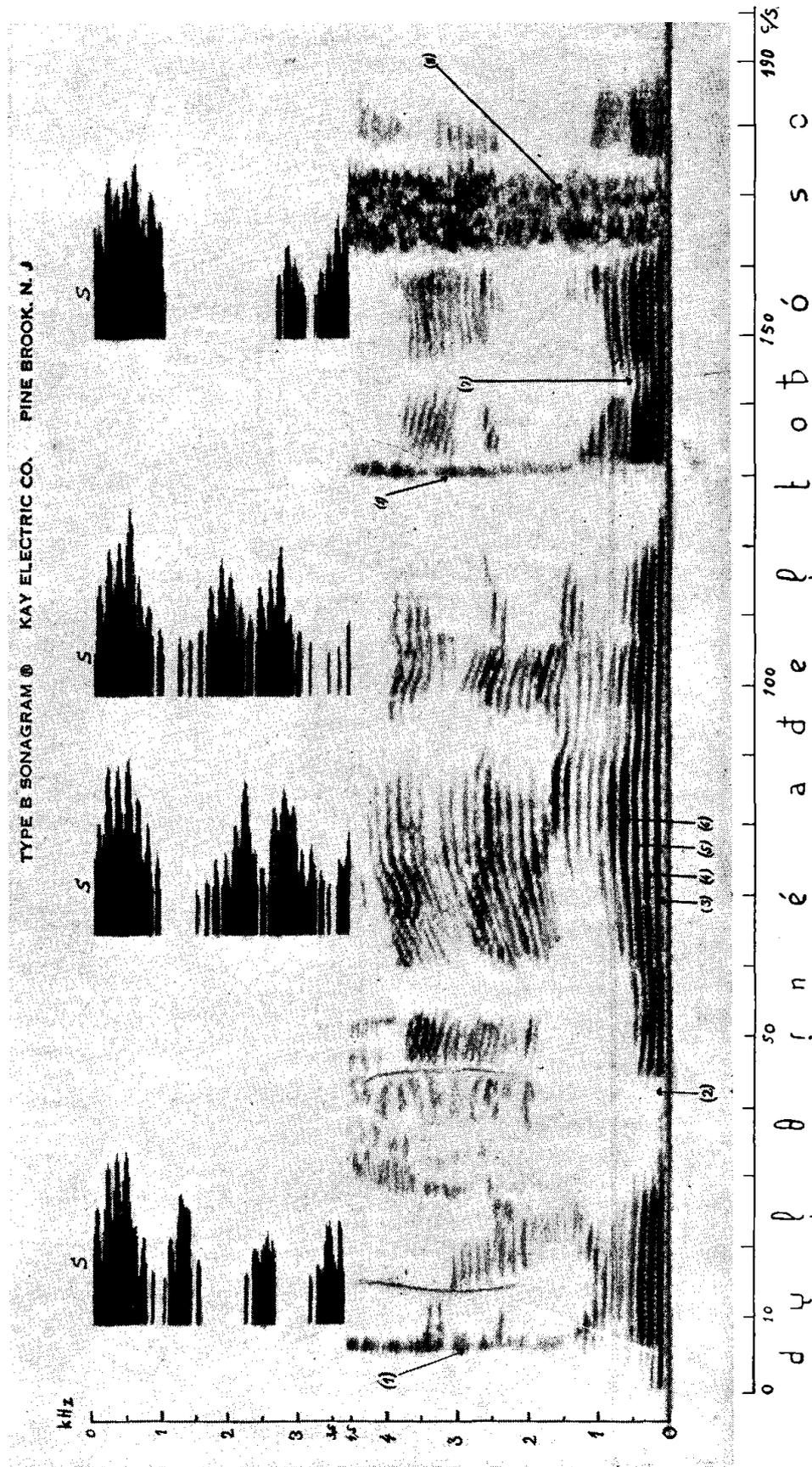


Fig. 4

- | | | | |
|-----|---------------------|-----|---------------------|
| (S) | Secciones. | (5) | Tercer armónico. |
| (1) | Barra de explosión. | (6) | Cuarto armónico. |
| (2) | Sordidez. | (7) | Armónicos de la -b- |
| (3) | Primer armónico. | (8) | Fricación sorda. |
| (4) | Segundo armónico. | | |

del fundamental, el poder hallar el tono de éste no es muy complicado. Esta es, como ya dijimos anteriormente, una de las principales ventajas del método espectrográfico. El armónico correspondiente a la línea melódica del fundamental aparece a simple vista; no hay nada más que aplicar la escala para averiguar su frecuencia. Para mayor exactitud en la medida, se pueden hacer distintas calas en los armónicos recordando siempre su propiedad de múltiplos del fundamental.

Asimismo, podemos observar también la gravedad de la voz con que ha sido pronunciada esta frase en la pequeña separación que existe entre los armónicos, ya que la frecuencia del fundamental es muy baja. En las voces femeninas, de timbre más agudo, la separación entre los armónicos es mucho mayor.

En la parte superior del espectrograma de la figura 4 tenemos representadas las *a m p l i t u d e s* de los sonidos. La «sección de amplitud» es el mejor medio de representar la intensidad de un sonido en un punto dado de su extensión en el tiempo, puesto que su frecuencia no debe variar. La escala de frecuencias de las secciones es inversa: comienza en la parte superior; en ella los formantes más altos son los de más baja frecuencia; a medida que descienden, la frecuencia aumenta. Por este sistema de las secciones de amplitud, la intensidad del sonido se obtiene con más exactitud que comparando el grado de mayor o menor oscurecimiento de las formantes. Su unidad de medida es el *b e l*, o su submúltiplo el *d e c i b e l*, que es más utilizado en las mediciones. Estas se efectúan en sentido paralelo al eje de abscisas, y de izquierda a derecha.

La *línea de intensidad total* aparece sobre el espectrograma de la figura 3. Esta representación no corresponde, como las secciones de amplitud, a la intensidad en un punto determinado del sonido, sino que se extiende a todo lo largo de la línea del tiempo, pudiendo observarse las diferencias de intensidad de unos sonidos con relación a otros. Su unidad también es el decibel, y la medición se hace verticalmente al eje de abscisas, tomando como puntos de referencia los de la línea que se extiende por debajo de la de intensidad.

Para obtener la *línea melódica del fundamental*, además del procedimiento ideado por Stanley M. Sapon (v. *RFE*, 1958-59, XLII, pp. 167-175), puede seguirse el utilizado habitualmente con el espectrógrafo. Sabemos que la frecuencia *F* del fundamental está en proporción inversa a su período *T*, o duración de una oscilación completa. O sea que:

$$F = \frac{1}{T}$$

Partiendo de este dato, si duplicamos la velocidad de la grabación que tenemos en la cinta magnetofónica, el tiempo disminuye, y, por tanto, la frecuencia del fundamental del sonido o de la frase aumenta. Al aumentar la frecuencia y estar el tiempo, como hemos dicho, más comprimido (o lo que es lo mismo: para una misma extensión del espectrograma tenemos una doble cantidad de tiempo), las oscilaciones de la melodía aparecen mucho más acusadas. Tomando entonces sobre un eje de coordenadas los puntos correspondientes a frecuencia y tiempo del fundamental, podemos trazar la línea melódica de la palabra o frase en cuestión. Es evidente que, por la reducción que sufren todos los componentes del sonido es, a veces, un poco complicado su análisis. En tales casos podemos hacer las comparaciones o comprobaciones necesarias con otro espectrograma de velocidad normal, donde los sonidos aparecen con sus características propias.

ANTONIO QUILIS MORALES.