



JORNADAS NACIONALES DE ACUSTICA

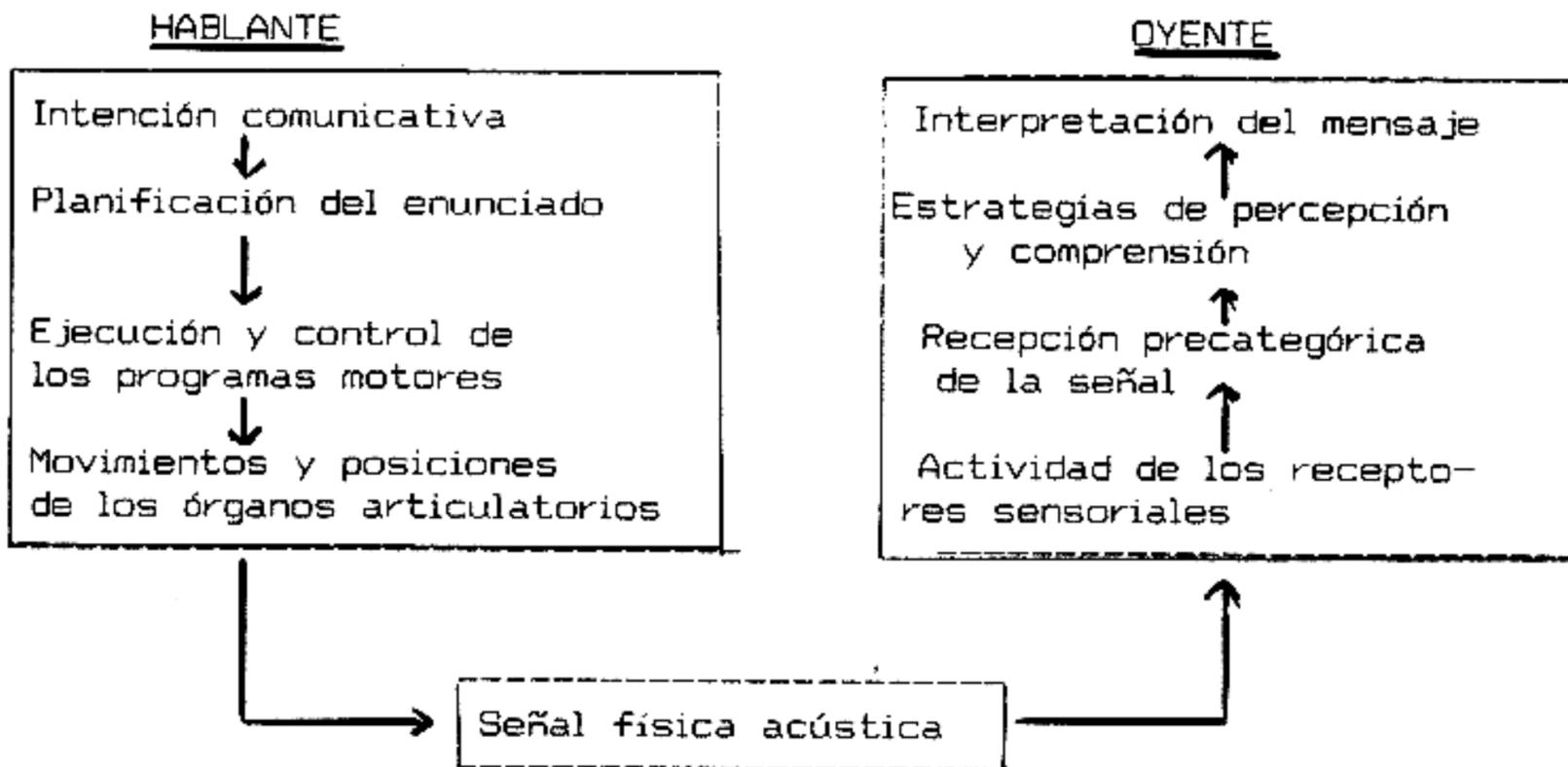
Zaragoza, Abril 1989

PERCEPCION DE LA PALABRA EN INDIVIDUOS SANOS Y CON LESION AUDITIVA: INDICACIONES Y CONTRAINDICACIONES DE LA CORRECCION CON PROTESIS AUDITIVA

A. Borragán Torre

Médico especialista en Foniatría
C/ Cádiz 9, 7º
Santander -Cantabria-

La percepción del lenguaje es un proceso complejo que forma parte del sistema de elaboración de la información. Las diferentes fases de este sistema son (Lehtonen e Hurme,1980):



Cuando un oyente recibe un mensaje oral busca extraer la información de la señal acústica, es decir, decodificar dicho mensaje. Se realiza gracias a unos procesos que decodifican la información en sucesivos niveles de abstracción: auditivo, fonético, fonológico, lexical, sintáctico y semántico. A esta decodificación se llega después de una secuencia organizada jerárquicamente con funciones de síntesis sucesiva, que se realizan a través de los mecanismos cerebrales de recepción, análisis y almacenamiento de la información.

Sólo analizaremos la decodificación fonológica que consiste en descomponer lo acústico en los elementos mínimos que poseen una función o bien en las unidades básicas. Es decir, que a través del continuum fónico se establece una cadena fonémica, una cadena de unidades básicas. Esta nos permite reconocer lo que oímos aunque nos parezca que oímos directamente las palabras o las frases.

El estudio de la percepción lingüística se ha desarrollado a partir de los años 60, gracias a dispositivos para la síntesis de la palabra (sonido sintetizado, máquinas que hablan), como el Pattern Play Back, puesto a punto en los laboratorios Haskins.

El lenguaje es un sistema simbólico de signos con el cual el hombre comunica lo que siente y piensa. Su característica fundamental es que con un número limitado de signos (29 grafemas y 24 fonemas) se puede realizar un número infinito de combinaciones, pudiendo transmitir así infinidad de pensamientos (Borragán Torre, 1988). Cada fonema se opone a otro por una serie de características. A esto se le denomina oposición fonológica y condiciona diferencias, en la producción lingüística y en la percepción auditiva. Ejemplo de lo anterior son la existencia de palabras en el español que se diferencian sólo por mínimas variaciones: parejas mínimas.

capa(Kapa)	cata(Kata)	caca(Kaka)		
cava(caba)	cada(kada)	caga(Kaga)		
cafa(Kafa)	caza(caDa)	casa(casa)	calla(Kaja)	caja(kaxa)
catcha(katsa)				
cala(Kala)	cara(Kara)	carra(Kara)		
cama(Kama)	cana(Kana)	caña(Kana)		

Se podría pensar que esta percepción por oposición de los sonidos lingüísticos venga determinada por las características intrínsecas del oído humano, es decir, que la percepción sea la misma para todos. Pero no es así, ya que la percepción de una determinada lengua viene condicionada por el aprendizaje en el periodo de su adquisición (en la primera infancia), resultando el oyente o el hablante sensible a las diferencias sonoras que son transformadas por el sistema fonológico, resultando insensible a las diferencias que no lo son (Fry D.B., 1986). Esto condiciona un espacio perceptivo donde aparecen todos los fonemas y donde se mueven todas las diferencias fonémicas de esa lengua.

La percepción no puede basarse sobre mínimas diferencias percibidas. Cuando se percibe un rostro, se reconocerá a pesar de que exista una luz tenue o una posición del objeto no habitual. También se reconocerá como idéntica una palabra (por ejemplo, buhardilla) producida por 50 hablantes diferentes, o bien esa

misma palabra producida por una misma persona en 50 ocasiones distintas. Las diferencias acústicas y visuales en todos los casos son grandes, y el oyente debe ignorar mucha información y quedarse con lo esencial. Estas características esenciales se denominan "rasgos sonoros". La capacidad de abstraer las unidades mínimas es esencial para la percepción lingüística, ya que los sonidos que un oyente escucha están impregnados de una enorme variabilidad: voces de hombres, de mujeres, de niños, diversos ruidos ambientales, música, el contexto fonético, la coarticulación de los fonemas, la posición de la palabra, diferencias sociolingüísticas, diferencias de sexo y edad, la idiosincrasia del hablante o el estilo del enunciado, la prosodia, el acento y la entonación, las características paralingüísticas. La percepción de los rasgos acústicos depende de las relaciones que se establecen entre ellos (por ejemplo, la cantidad física o intensidad) y no de sus valores absolutos. Es decir, que estas diferencias no dependen tanto de parámetros físicos intrínsecos, sino de la forma en que el cerebro capta y organiza las relaciones entre estas diferencias físicas. Algunos rasgos acústicos se presentan en sistemas fonológicos de diversas lenguas (presencia/ausencia de un breve silencio, contraste entre sonido periódico y no periódico...).

Los rasgos sonoros se pueden clasificar según la duración, la frecuencia y la intensidad. A pesar de ser múltiples los índices o los rasgos sonoros en la cadena hablada, sólo una de estas dimensiones será la preponderante e influirá sobre el efecto perceptivo. Se han hecho múltiples experimentos mediante la síntesis de palabras por ordenador, donde variaban diversos parámetros de los índices acústicos y así se establecían unos test de percepción para medir dichos índices.

Podemos enumerar los índices presentes en cada rasgo sonoro (Fry D.B., 1986):

RASGO DE DURACION de la cadena fónica

1.-Consonantes oclusivas. Interrupción en el flujo sonoro: silencio (oclusivas sordas) o casi silencio (oclusivas sonoras, africadas). La duración absoluta del silencio no es constante, incluso varía en un mismo hablante. Es dependiente del contexto fonético, de la coarticulación y del grado de énfasis de la secuencia sonora.

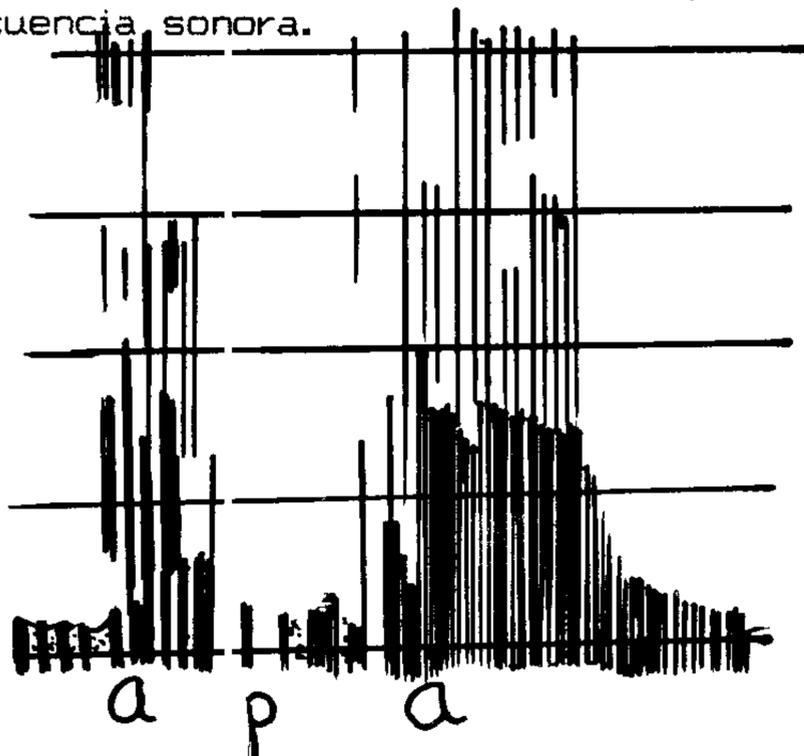


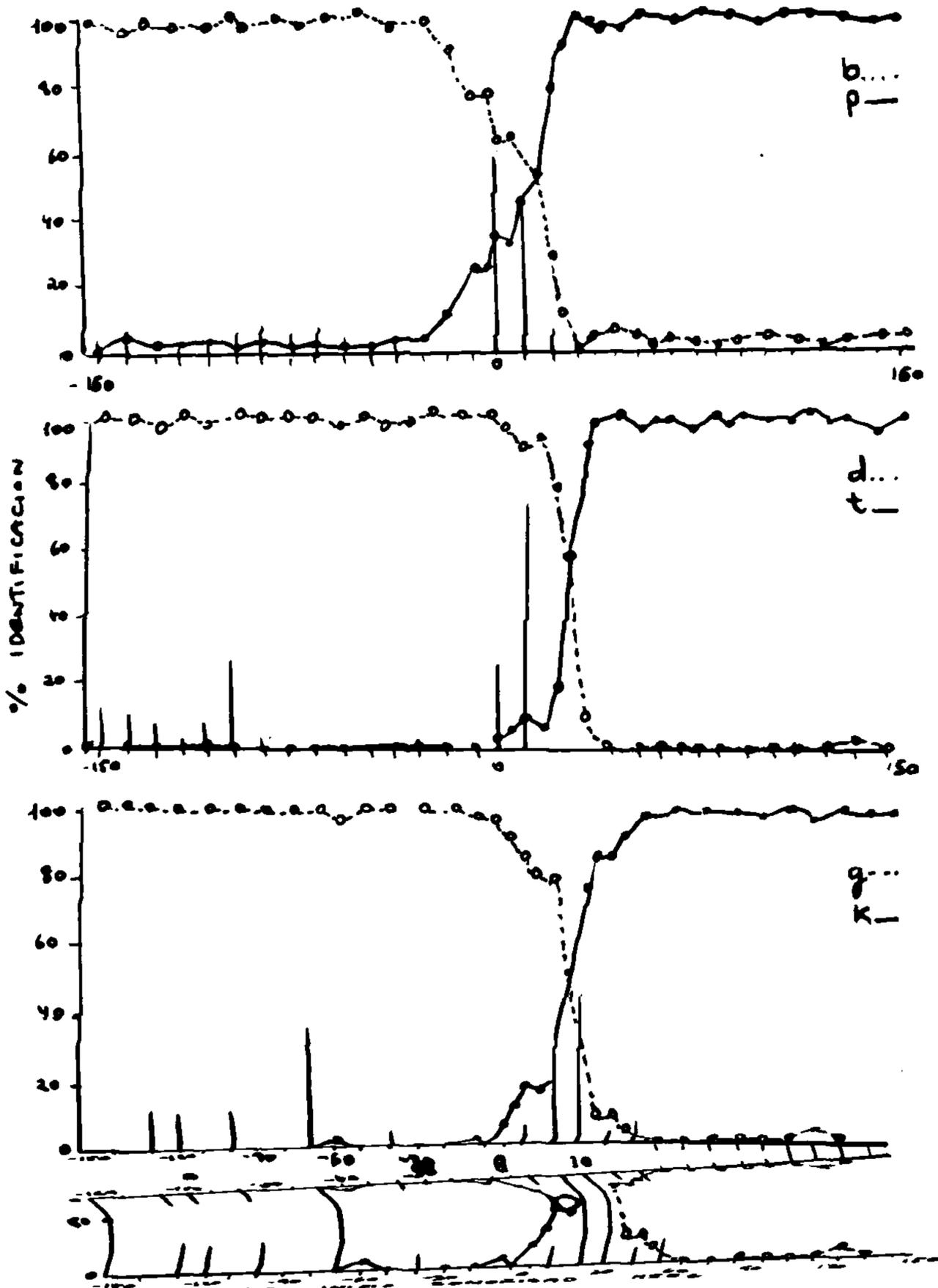
Fig. Espectrograma del fonema /p/ coarticulado. Se muestra la fase de silencio

2.-Distinción sordo-sonoro:

a. Duración del tiempo de silencio: cuando la duración es corta, corresponde a un fonema sonoro; si es larga, le corresponde un fonema sordo. Entre ambas zonas bien delimitadas siempre hay una zona de incertidumbre o de frontera que separa dichas unidades fonémicas.

b. VOT: tiempo de inicio de la sonoridad: se estudia el inicio de la sonoridad de una consonante oclusiva seguida de una vocal(sonora). En las consonantes sordas el VOT es positivo, es decir, hay un periodo de intervalo (de silencio) antes de que se inicie la sonoridad; en las sonoras es negativo, ya que la sonoridad se inicia incluso antes de la emisión de la consonante sonora (b, d, g). La longitud del periodo de inicio de la sonoridad es directamente proporcional a la aspiración de las consonantes ;por esto, en español donde la aspiración es escasa, el tiempo que media entre la explosión y el inicio de la sonoridad es corto.

Fig. VOT: porcentajes de identificación en los test con estímulos sintetizados. Los histogramas muestran identificación enunciado



natural (Fry, 1986).

RASGO DE FRECUENCIA

Es la distribución de la energía acústica sobre el espectro sonoro. En los sonidos vocálicos las resonancias del aire al pasar por el tracto vocal, originan unas concentraciones de energía a unas determinadas frecuencias a partir de los 250 Hz, pero sólo los componentes hasta 2500 Hz son determinates para reconocer los sonidos vocálicos. En este rango de frecuencias se encuentran las dos primeros formantes. En los recuadros siguientes se observa la frecuencia (en Hz) de los dos primeros formantes de las vocales del español en locutores masculinos y femeninos según Quilis y Esgueva (1983).

♂	F1	F2
/i/	264,5	2317,5
/e/	453,8	1995,01
/a/	657,28	1215
/o/	474,5	888,4
/u/	293,5	669,08

♀	F1	F2
/i/	240,75	2841,7
/e/	491,6	2252,08
/a/	663,75	1167,8
/o/	510,75	981
/u/	243	628,8

Sin embargo, existe una gran variabilidad entre los sonidos vocálicos (mucho más que entre las consonantes), por eso hay que hablar de un diagrama de dispersión.

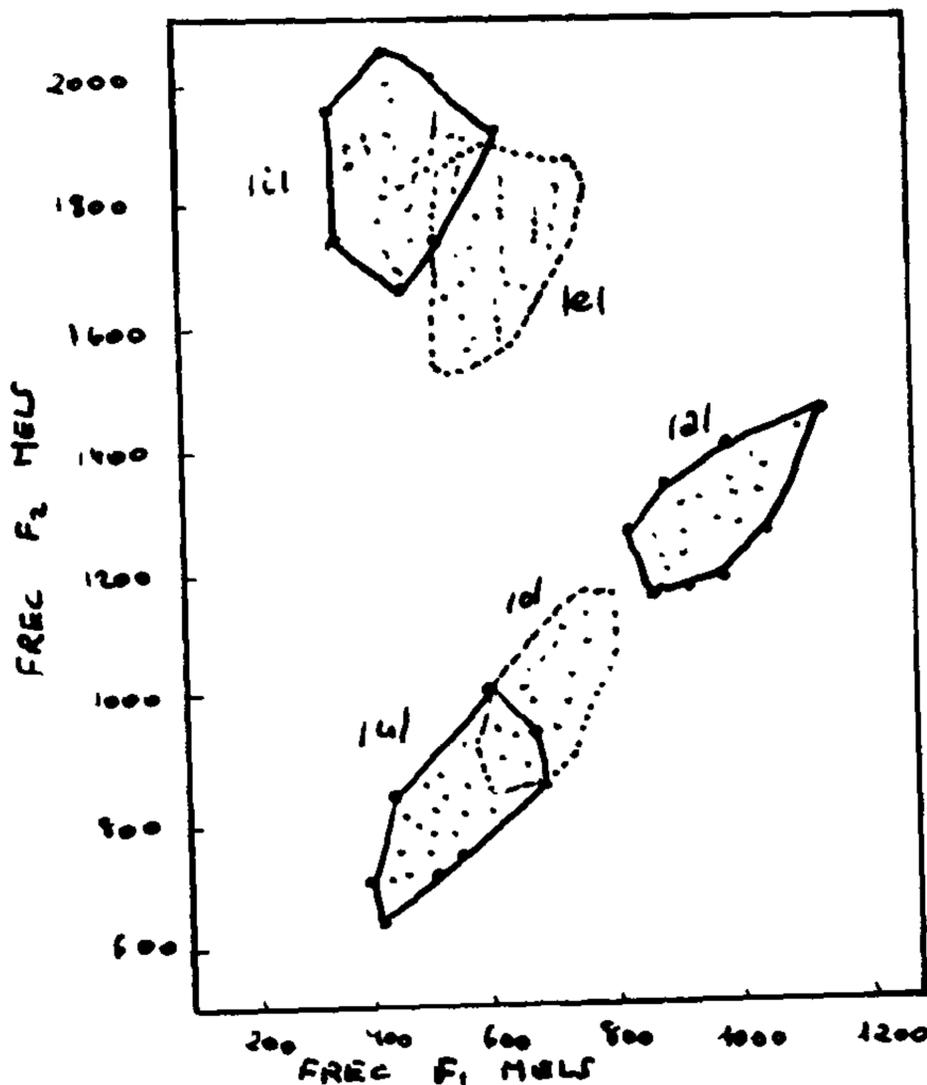


Fig. Diagramas de dispersión de las vocales españolas

A pesar de estas diferencias, el oyente reconoce las vocales quizá por la redundancia del sistema lingüístico: un oyente por el contexto lingüístico puede suponer cual va a ser la vocal que se presentará en una sucesión fonémica sin realizar una percepción de todos los elementos acústicos.

1. Índice de transición de las formantes: cambios rápidos en la frecuencia de las formantes por la articulación de las consonantes. Estos son rápidos y netos porque las explosivas son muy veloces.

a. Transición de la segunda formante. Liberman (1957) obtuvo, a través de un sintetizador de la palabra la producción de oclusivas sonoras diversas, variando la transición del 2º formante. Este rasgo se utiliza para la discriminación del punto de articulación: la transición negativa del 2º formante indica el rasgo bilabial; la 0 o ligeramente positiva para la alveolar y una transición muy positiva para la articulación velar. Se puede usar, asimismo, para la diferenciación de las nasales.

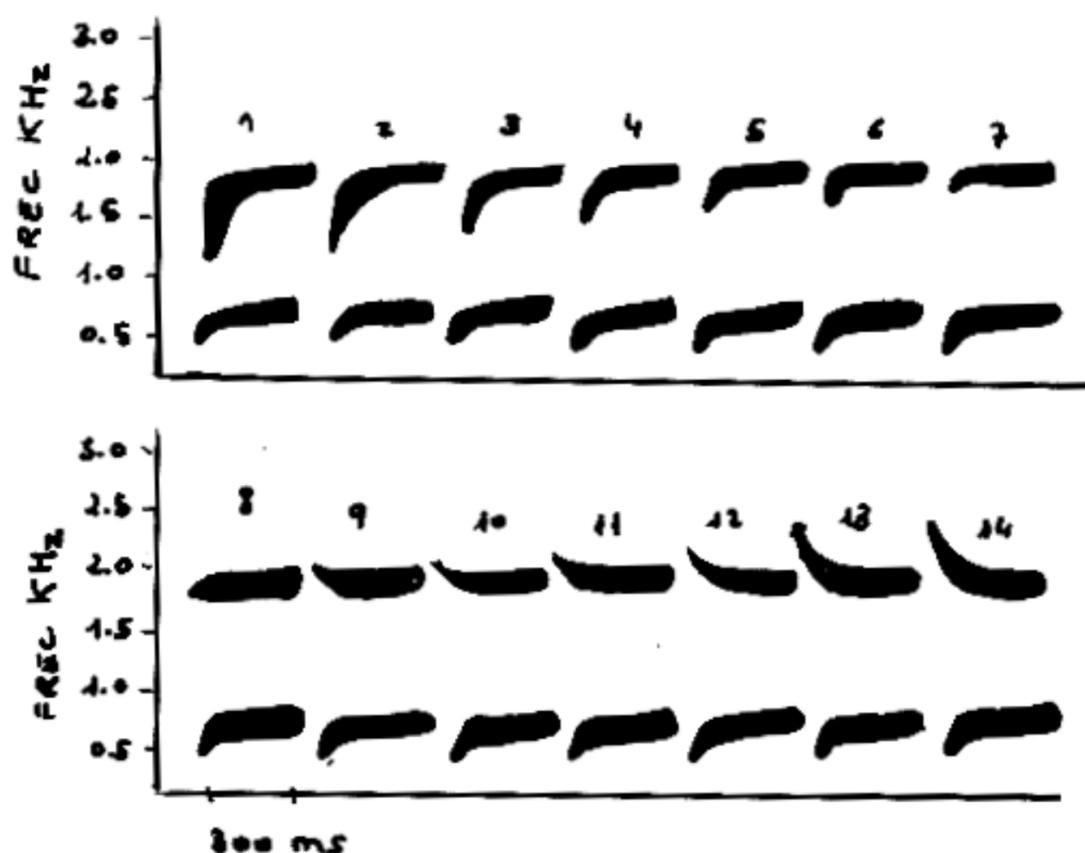


Fig. Transición del 2º Formante (Liberman et al., 1957)

b. Transición del 2º formante más largo, al moverse los articuladores más lentamente. Esta diferencia permite distinguir otros fonemas como /r/, /l/ (para ambas es necesario un estudio del 3º formante), /j/, /w/.

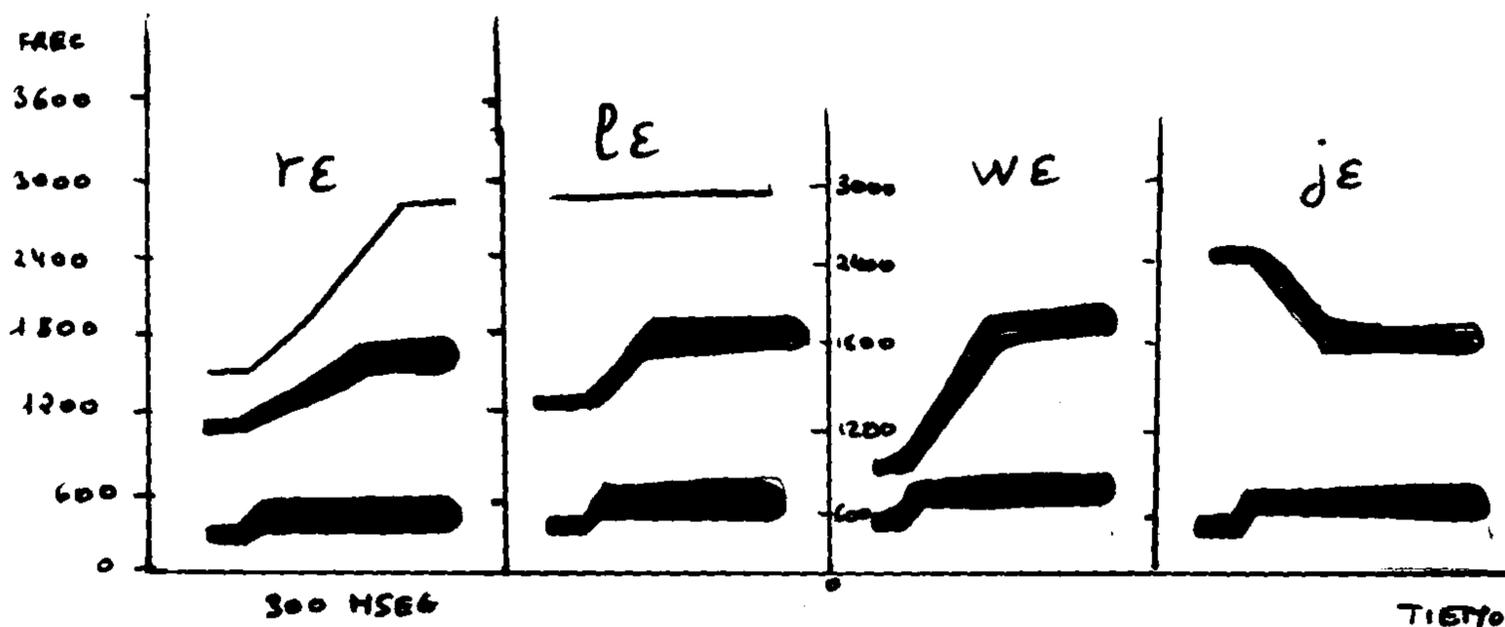


Fig. Transición del 2º formante

c. Frecuencia del ruido: la explosión de ruido que acompaña a la apertura en una consonante oclusiva presenta frecuencias diversas según el punto en que la consonante se genera. Cooper (1957) relacionó ruidos explosivos de diferente frecuencia con unas vocales sintéticas. Una explosión a una frecuencia media seguida de /i/ o /u/ se percibe como /p/, pero si se sigue de /a/ se percibe como /k/. Será necesario el uso del rasgo de transición de la 2ª formante para discriminarlos. El índice de frecuencia de los ruidos desarrolla un papel importante en la percepción de los sonidos fricativos. El efecto de filtro del canal vocal modifica el ruido difuso al pasar por una estrecha costricción, y produce ruidos con una banda de frecuencia limitada: si la fricción oscila entre 1800 y 4500 Hz. con una media de 2000 a 3000 Hz. se percibe un ruido bajo /s/; si la fricción va más allá de 4000 Hz. (con puntas hasta 7000 Hz) se percibe ruido alto /s/.

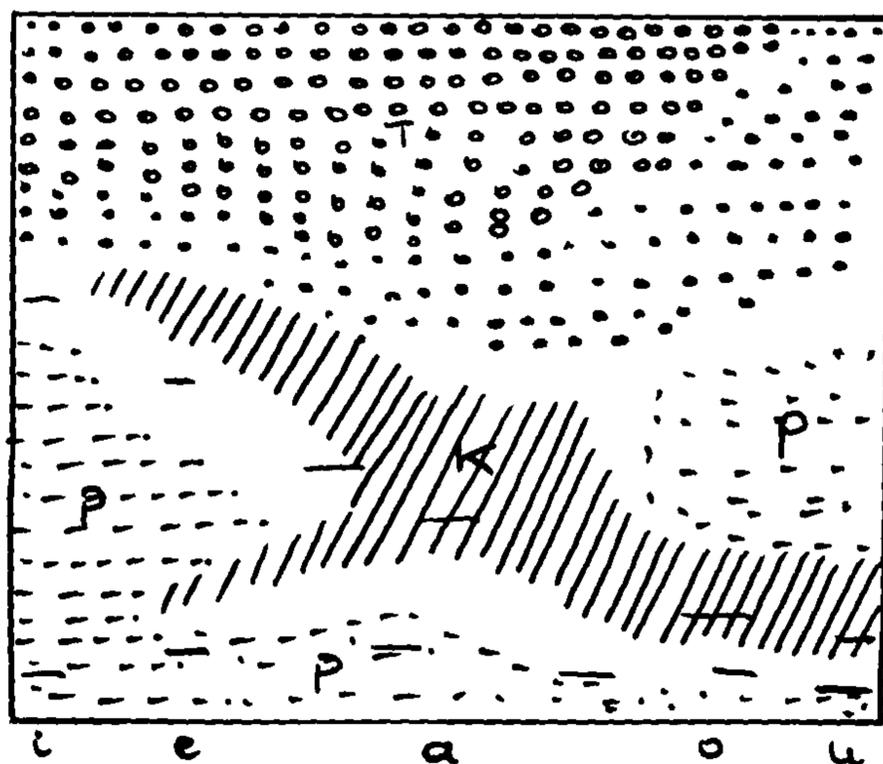


Fig. Ruidos de explosión de frecuencias diversas (Cooper et al., 1952).

RASGO DE INTENSIDAD

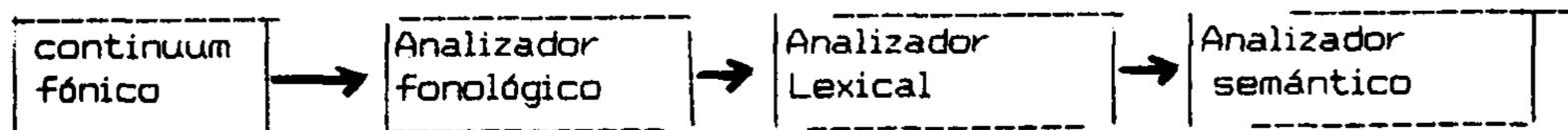
Sirve para reconocer la estructura silábica de los mensajes, es decir, para diferenciar sílabas y palabras. También para el reconocimiento de los modelos rítmicos y de las variaciones del acento y del énfasis en diversos modelos de entonación. Asimismo, influye en el reconocimiento de los fonemas, ya que la costricción ocasiona una disminución de la intensidad total de 0 o cerca a 0. Las vocales abiertas tienen un sonido más intenso en la cadena fónica (las vocales cerradas tienen una intensidad de 6-7 dB menos).

COMBINACION DE LOS INDICES ACUSTICOS

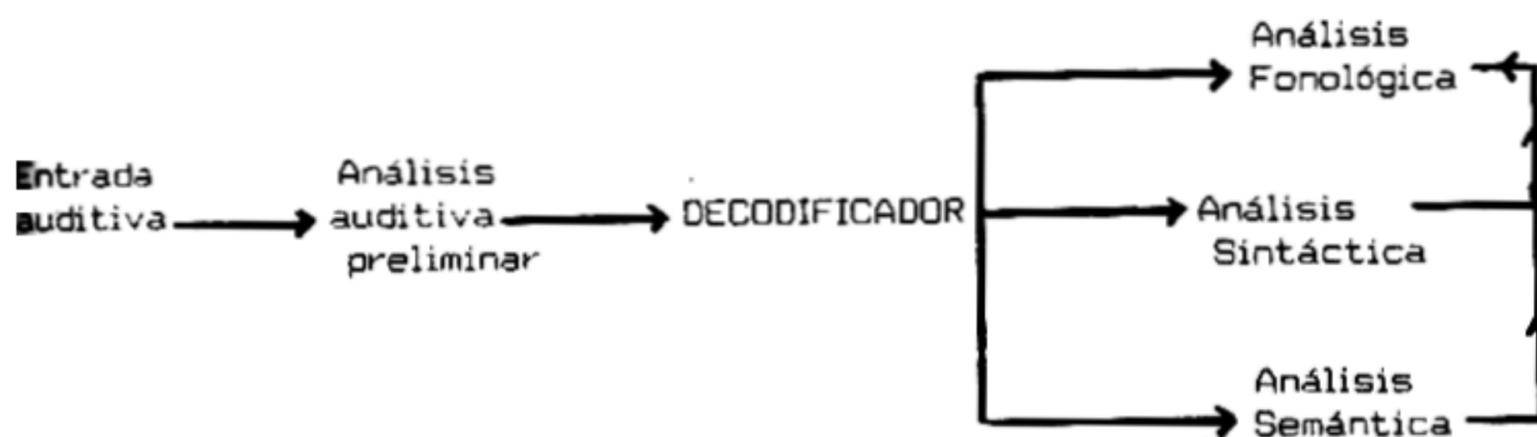
El oyente que recibe un mensaje, recibe un flujo continuo de sonidos. El cerebro lo primero que afronta es la transformación de este flujo en una cadena fonémica. Esto viene realizado por el uso de índices contextuales, por las restricciones semánticas (no siendo necesario que el cerebro analice todo el inventario fonémico de la lengua) y por los rasgos acústicos. La elección viene limitada por la presencia de un rasgo general. Por ejemplo, si existe una interrupción la elección se restringe a: p, t, k, b, d, g, ts; si existe un ruido de fricción se restringe a: f, ʃ, s, j, x. Una diferenciación más precisa se realiza con el uso del índice sordo-sonoro. En el caso de la diferenciación entre f y ʃ es imprescindible un rasgo suplementario porque ambos fonemas son ruidos a baja intensidad y alta frecuencia, y se realiza a través de la transición del 2º formante; en el /f/ hay una transición negativa, mientras que para la /ʃ/ la transición es mínima o nula. Otros índices ayudan a la percepción de diversos fonemas: alta intensidad - /s/ -, baja intensidad - /f/, /ʃ/ -, alta frecuencia - /f/, /ʃ/, /s/ -.

No hay que pensar que el funcionamiento de los rasgos acústicos en la percepción de la cadena fónica sea ordenada y bien definida y que la ausencia o presencia de un solo rasgo indique una unidad fonémica. Para una diferenciación o identificación hay disponible un número amplio de índices, pero el carácter redundante de la lengua reduce el uso que el oyente hace de ello.

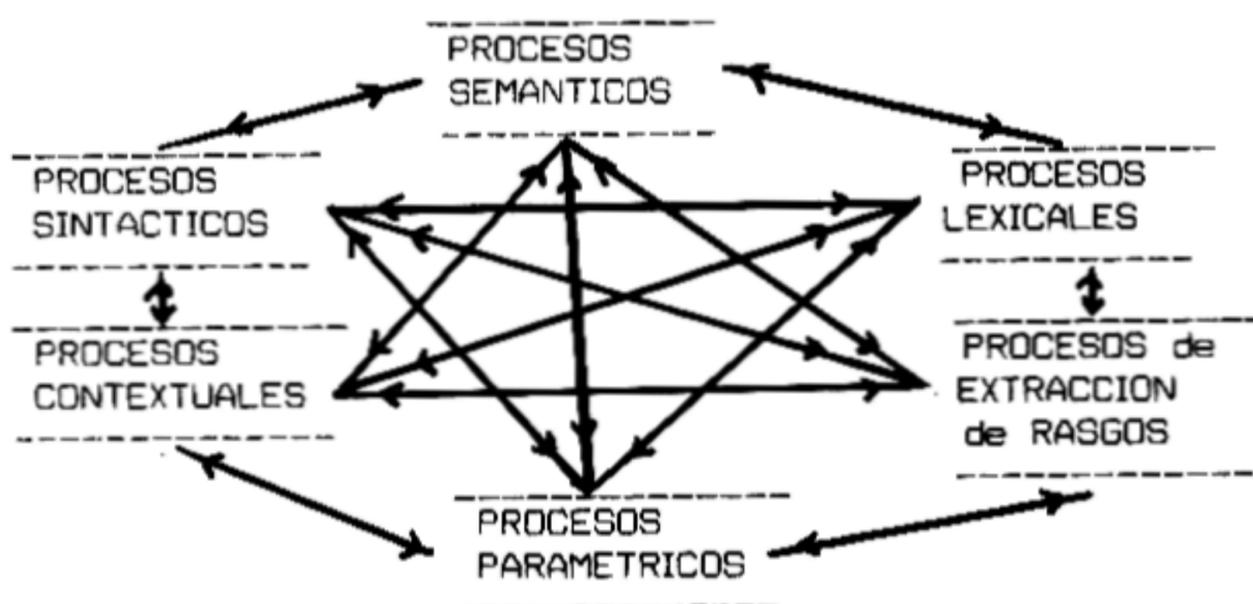
No se debe creer que para la elaboración de los sonidos de la lengua sea necesario la aplicación de unos criterios acústicos en serie.



La principal característica del cerebro humano es su capacidad de funcionar en paralelo, es decir, todos los criterios necesarios son analizados contemporáneamente y no uno detrás de otro, y el cerebro alcanza así el modelo de probabilidad.



Actualmente se hacen más hipótesis sobre un modelo de tipo interactivo.



Cuando un oyente está expuesto a un continuum fónico hay ciertos segmentos, en relación a un rasgo acústico, en los que no existe duda en dar una determinada etiqueta fonémica. Pero también existen zonas de incertidumbre o de frontera fonémica. Se puede pensar que el espacio perceptivo total esté dividido en subespacios donde la diferenciación fonémica no presente duda, es decir, sea insensible a la diferenciación. Pero los límites entre los subespacios presentan un gran aumento en la sensibilidad de diferenciación. Junto a la percepción de la secuencia fonémica se añaden otras informaciones como los modelos funcionales de acento y de tono del mensaje.

La unidad esencial a nivel lingüístico se piensa que sean las sílabas y no los fonemas. Se puede distinguir entre señal sonora y mensaje lingüístico, entre segmentos sonoros y segmentos fonémicos.

Las funciones lingüísticas (control del aparato articulatorio y análisis perceptivo) están localizadas sobre todo en el hemisferio izquierdo, en la región posterior izquierda, en la circunvolución temporal superior. Esto se puede estudiar a través de enfermos con afasia, en secciones del cuerpo calloso para tratamiento de la epilepsia, en la administración de amital sódico: test de Wada, con técnicas taquistoscópicas, técnicas dicóticas, afasias en sordos. El hemisferio derecho capta el significado a través de la configuración global de la palabra y no a través del análisis fonológico, es decir, percibe globalmente y no analíticamente. Este control lateralizado parece unido a la predominancia manual, al control de movimientos especializados, al control de movimientos finos y no a procesos simbólicos más elevados. Es raro encontrar otros animales con lateralización hemisférica. Una de estas excepciones es el canto de los pájaros: si se secciona el hemisferio cerebral izquierdo produce una pérdida casi total del canto: la misma lesión en la región derecha no tiene apenas consecuencias. Esta dominancia hemisférica en los pájaros parece relacionada con efectos exclusivamente motores y no perceptivos.

Hay estudios que intentan demostrar que este analizador es innato, pero no pueden explicar por qué las categorías fonémicas están determinadas lingüísticamente (son diversos los índices de VOT en fonemas semejantes en las diversas lenguas, y de ello depende el que se valore a un hablante como nativo), en vez de fisiológicamente.

La percepción no es un proceso tan sencillo y esquemático como viene aquí expresado. Ha habido, y hay, muchas controversias entre diferentes teorías o formas de concebir el proceso. Existen verdaderas dicotomías en la concepción de estos procesos:

1. modelos de tipo taxonómico donde las diversas fases tienen un valor exclusivamente conceptual versus diagramas de flujo o modelos dinámicos donde la información se recodifica en formas cada vez más abstractas.

2. Análisis en serie (que tienen un funcionamiento muy lento y hay posibilidad de arrastrar errores) versus análisis en paralelo.

3. Análisis en bottom-up : análisis jerárquico donde se realiza un análisis fonético autónomo con hipótesis de segmentación en competición (pero se obtiene un número de hipótesis múltiple que puede ocasionar a su vez un número grande de hipótesis a nivel lexical; muchas de estas secuencias probables no son válidas y serán rechazadas por el sistema sintáctico y semántico) versus Análisis top-down : sistema jerárquico que implica intercambios continuos de información entre lo analizado, existiendo una gran influencia a nivel superior (esto lo demuestra la integración mental, la corrección de errores, la interpretación de frases ambiguas, la percepción de clicks... Estas restricciones vienen

dados por el conocimiento de la lengua, de sus reglas, de sus restricciones fonotácticas, de su léxico...

4. Mezclar activos versus acti-los pasivos

Un último aspecto a analizar es la transformación de la información de acústica, o auditiva, a fonética, a fonológica. Hay que distinguir entre dos aspectos cualitativamente diferentes de un mismo acto perceptivo: el aspecto acústico y el aspecto fonético son dos modos diferentes de análisis del estímulo. Esta diferencia se aprecia al estudiar el fenómeno de percepción doble (literamen et al., 1961) cuando se presenta dicépticamente una sílaba, por un lado la "brasa" y por otro la transición fonotáctica, se percibe mediante la integración una sílaba, mientras que si se presenta la transición aislada se oye un ruido - chirp. Por esto un mismo parámetro acústico (la transición del 2º formante) se oye en dos formas diferentes, una auditiva (chirp) y otra fonética (verbal). Otra demostración de la diferencia de los aspectos acústico y fonético es el hecho de que la percepción de un estímulo verbal puede ser cambiada al superponer al sonido de un fonema la imagen visual de otro fonema diverso, por ejemplo, dar el fonema /ga/ auditivamente y /ba/ visualmente, se percibe entonces /ga/ o /ga/ b/. Si se presenta /ba/ auditivamente y /ga/ visualmente, se percibe /da/; esto es el resultado de una fusión de la información auditivo-visual (Moguck et al., 1976). También en el fenómeno de adaptación verbal (Elmas et al., 1970), que consiste en que un sujeto identifica sílabas antes y después de una exposición repetida a una de las sílabas; el efecto es que el oyente identifica menos que antes de la adaptación. Mientras el sistema auditivo se adapta al estímulo acústico, sus percepciones fonéticas eran el resultado de una mezcla de información visual y acústica, es decir, que la percepción fonética es intermodal. La información acústica la da la estructura de las formantes, el ruido de fricción, el ruido de explosión, las variaciones de las amplitudes. La información visual viene determinada por el cambio en la configuración de los labios, de la lengua, de los dientes. Lo que unifica estas cosas (la estética) es no solo la cinemática (los cambios y los ritmos de movimientos, la velocidad, la aceleración) sino también las fuerzas dinámicas que están debajo de estos cambios o las propiedades biofísicas que definen las fuerzas esas, fuerza, elasticidad (Palenatti, 1966).