

**LA PERCEPCION CATEGORIAL DE /b-p/ EN
ESPAÑOL BASADA EN LAS DIFERENCIAS DE
DURACION.**

EUGENIO MARTINEZ CELDRAN
Director del Laboratorio de Fonética de la
Universidad de Barcelona

RESUMEN

El objetivo de este estudio era comprobar si en español la duración del silencio de las oclusivas labiales era una pista suficiente para distinguir /b/ y /p/ y, además, si sucedía de forma categorial, como ocurrió en inglés. Se han realizado dos tests: uno de discriminación y otro de identificación. Se han utilizado 23 estímulos sintéticos para conseguir realizaciones de las palabras: *cava*, *capa* y *cappa*. Las duraciones del silencio iban desde 26,4 ms hasta 220 ms en intervalos de 8,8 ms. Los resultados demuestran que, a pesar de las diferencias en las realizaciones fonéticas de ambas lenguas, la duración del silencio es una pista suficiente para distinguir estas categorías y que existe la percepción categorial basada en esa misma pista. Además, como sucedió en inglés, hay una tercera categoría que sin ser fonológica también se percibe correctamente y que posee una duración muy larga: /pp/.

ABSTRACT

The aim of this study was to confirm whether the durations of silence were a sufficient cue to distinguish between /b/ and /p/ in Spanish and whether it happened categorically as in English. Two tests were done: the first one was an identification test and the second one was a discrimination test. Twenty-three synthetic stimuli with phonetic realizations of the words *cava*, *capa* and *cappa* were prepared. The durations of silence ranged from 26,4 ms to 220 ms with 8,8 ms intervals. The results show that, despite the existing differences between phonetic realizations in both languages, the categorial perception based on durations of silence in labial stops is present in Spanish, as well as in English, and also that the durations of silence are a sufficient cue to distinguish between /b/ and /p/. As is the case in English, there is a third category /pp/ that, without being phonemically distinctive, is also perfectly perceived and has a rather long duration.

1. INTRODUCCION

1.1. La percepción categorial.

Cuando en 1957 Liberman, Harris, Hoffman y Griffit empezaron a experimentar con el sintetizador Pattern Playback, se encontraron con que estímulos que formaban parte de un continuum físico eran clasificados en categorías por los oyentes. Comprobaron además mediante tests ABX de discriminación que los estímulos que pertenecían a distintas categorías eran fácilmente discriminados, mientras que los que pertenecían a la misma no eran tomados como desiguales a pesar de sus diferencias físicas. A esta forma de percibir los estímulos se le denominó *percepción categorial* (Repp, 1984).

Así, pues, el ser humano percibe el habla de forma discreta a pesar de que los estímulos físicos sean continuos. Así lo expresan Liberman, Harris, Eimas, Lisker y Bastian (1961:175) cuando dicen que *hemos encontrado que la discriminación de una diferencia dada es considerablemente más aguda a través de los límites de los fonemas que en el centro de las categorías fonemáticas*. Es decir, cuando dos estímulos físicos diferentes representan dos fonemas distintos son perfectamente diferenciados en los tests de discriminación; mientras que si los dos estímulos físicos distintos representan el mismo fonema, entonces no se suelen distinguir entre ellos.

Al principio pareció que este fenómeno era exclusivo del habla humana; pero investigaciones posteriores demostraron que si los estímulos extralingüísticos estaban categorizados, como en el caso de la música, también eran percibidos de forma categorial (Pisoni, 1978). En el lenguaje, las vocales y ciertos rasgos prosódicos no son percibidos categorialmente (Fry, Abramson, Eimas y Liberman, 1962); se comportan, pues, como los estímulos extralingüísticos no categorizados.

1.2. Discriminación vs Identificación.

Conviene aquí diferenciar claramente entre

estos dos tipos de tareas a las que son sometidos los informantes en los tests. La tarea de discriminación consiste en diferenciar entre sí varios estímulos. El típico test ABX de discriminación consiste en presentar tres estímulos al informante, de los cuales los dos primeros (AB) son diferentes y el tercero (X) es igual que A o que B. El oyente ha de decir si X es igual que el primero o el segundo.

La identificación consiste en dar un nombre al estímulo que se está oyendo. En el caso del habla, se trata de decir el fonema que representa el estímulo percibido.

1.3. Hipótesis de trabajo

En estudios anteriores (E. Martínez Celdrán, 1991a, 1991c) he estudiado la importancia de la duración en la identificación de los fonemas /p, t, k/ frente a sus respectivos /b, d, g/; de modo que ese hecho me llevó a defender que el rasgo pertinente que actuaba en estos casos era el rasgo tenso frente al laxo. En el estudio, realizado en 1992, veía que la situación actual es consecuencia del proceso histórico de lenición que sufrieron los romances occidentales en su evolución. En la lenición toda la serie de fonemas oclusivos latinos sufrió una relajación que se transmitió en cadena. Suponiendo que A. Martinet (1955) tenga razón, la relajación empezó por la simplificación de las geminadas latinas, continuó en la sonorización de las sordas simples y acabó con la espirantización de las sonoras. Pero en ese último estudio traté de demostrar que las tres series de oclusivas mantenían sus diferencias fundamentalmente a través de la duración del silencio que constituye la tensión de cada uno de esos sonidos y que simplificación, sonorización y espirantización eran sólo tres aspectos superficiales y redundantes del mismo proceso subyacente: la pérdida de duración en el silencio o la oclusión. Para demostrarlo llevé a cabo un experimento de identificación de estímulos, que ahora retomo y completo con otro de discriminación basado en los mismos materiales que el anterior.

Mi hipótesis, pues, sostiene que bastan las diferencias en la duración del sonido para que interpretemos perceptivamente [pp], [p] y [β] intervocálicas y que el modo de percibir las funciona categorialmente. Vamos a comprobar esta hipótesis presentando dos tests auditivos a unos informantes, en los que se ofrecen estímulos que sólo se diferencian por su duración. Naturalmente, a la hora de preparar los tests, partimos de algunas premisas que ya conocíamos por otros experimentos perceptivos llevados a cabo con anterioridad (E. Martínez Celdrán, 1991b). Como el contexto es intervocálico no podemos categorizar una [b] y una [β], sólo podemos pedir la identificación de la segunda frente a [p], por ejemplo. Además, como [β] no es explosiva, tenemos que eliminar la explosión; pero, para no tener que modificar en absoluto ninguno de los estímulos, la explosión va a ser suprimida en todos, esperando que en las bilabiales no sea necesaria como ya habíamos comprobado en otros experimentos. Por último, confiábamos en que los informantes percibirían la geminada a pesar de no existir en castellano, porque es un sonido muy diferente del simple; lo que buscábamos es que lo reconocieran como diferente de [p] y [β]. Además, como los informantes eran, en general, bilingües catalán-castellano, esperábamos que también tendrían una buena categorización de la geminada ya que teóricamente existen algunas de ellas en catalán, aunque no tienen un "status" fonológico. Por otra parte, tomo sólo las labiales porque son las únicas que no necesitan la explosión para ser identificadas como tales, como ya se ha dicho.

2. METODO

2.1 Los estímulos

Los estímulos fueron preparados con un sintetizador de voz: PCF8200 de Philips y con el programa VOXSY de la casa TEVOX de Barcelona (E. Martínez Celdrán 1991a:90 ss). El contexto era ka_a. Las dos aes tenían transiciones descendentes hacia la consonante interior en los segundos y terceros formantes, como corresponde a

una bilabial. En el hueco sólo va a haber silencio: desde 3 tramas hasta 25. Una trama ('frame', en inglés) es una línea de datos frecuenciales que posee 8,8 milésimas de segundo. En este hueco, todos los parámetros tendrán valor cero, puesto que la oclusiva carece por completo de valores formánticos, sólo es silencio colocado entre vocales. Por tanto, preparamos una serie de estímulos que iban desde las 3 tramas hasta las 25; es decir, se comenzó en 26,4 ms y se fue aumentando, de 8,8 ms en 8,8 ms que suponía cada trama, hasta completar un total de 23 estímulos. El último de 25 tramas poseía 220 ms. Comenzamos con 3 tramas porque, según nuestra propia percepción, por debajo de las 26,4 ms ya no se oía [β] y llegamos hasta las 25, porque 220 ms ya era más que suficiente para percibir la geminada, siempre nos fiábamos de nuestra propia percepción. (Denominamos los estímulos con el número de tramas; así, el estímulo 5 es el estímulo que posee 5 tramas ($5 \cdot 8,8 = 44$ ms); por tanto, el número no indica el orden de los estímulos.)

Véanse tres sonogramas: el primero posee 35,2 ms de silencio en la oclusiva intervocálica; el segundo tiene 96,8 ms y, el tercero, 211,2 ms.

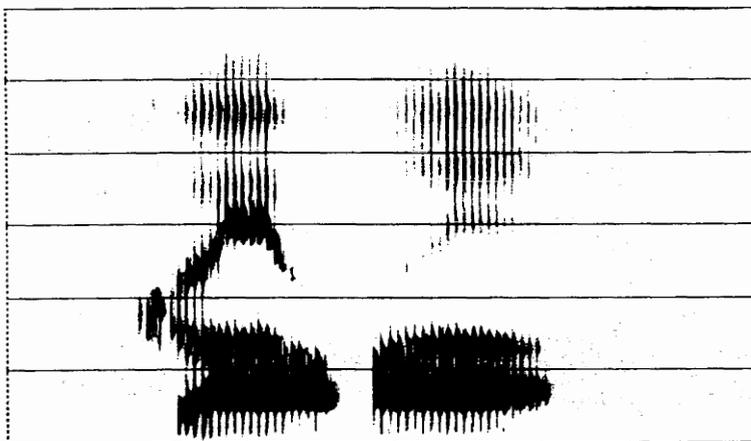


FIG. 1 Sonograma obtenido con el DSP 5500: estímulo sintético que posee 4 tramas de silencio en la posición intervocálica y que fue interpretado como [β].



FIG. 2 Sonograma obtenido con el DSP 5500: estímulo sintético que posee 11 tramas de silencio en la posición intervocálica y que fue interpretado como [p].

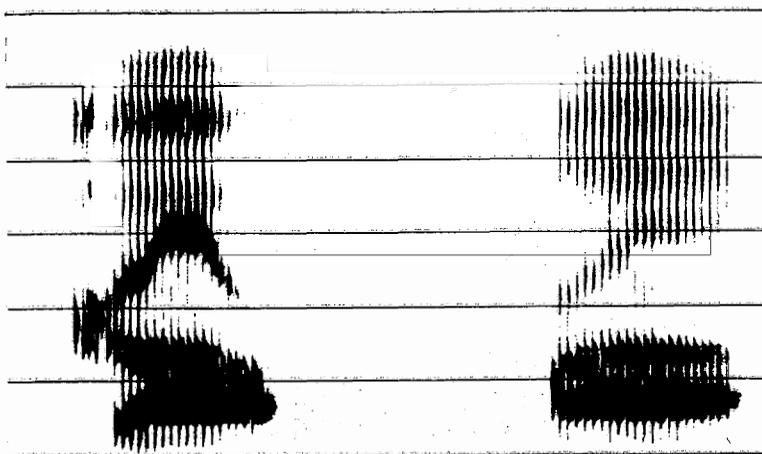


FIG. 3 Sonograma obtenido con el DSP 5500: estímulo sintético que posee 24 tramas de silencio en la posición intervocálica y que fue interpretado como [pp]

Todos los estímulos son iguales a excepción del número de tramas intervocálicas con ceros, que es lo único que hemos variado.

2.2. Test de identificación

El experimento de identificación ya se había llevado a cabo para el inglés (Alvarez González (1978), Liberman et alii (1961), etc.). De hecho, hemos seguido muy de cerca los pasos de Alvarez González. Para el test de identificación, se hizo una mezcla aleatoria. Mezcla proporcionada por un programa de ordenador. Así, pues, el orden de presentación de los estímulos fue el siguiente: 12, 25, 7, 22, 15, 16, 4, 14, 21, 5, 18, 23, 10, 3, 13, 20, 8, 9, 6, 17, 11, 24, 19 (el número hace referencia a la cantidad de tramas y, por tanto, a la duración del estímulo).

En el test de identificación se presentó una hoja con tres posibilidades: kappa - kapa - kaba, y los informantes (34 estudiantes de primero de Filología de la Universidad de Barcelona), después de oír tres veces cada estímulo, tenían que poner una cruz en la casilla de la secuencia que creían escuchar.

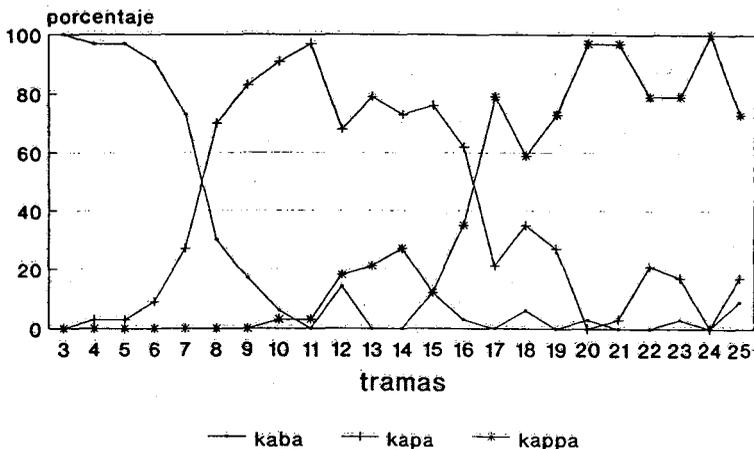
2.3. Test de discriminación

El test de discriminación fue de tipo AX; es decir, se presentó también en orden aleatorio cada uno de los estímulos con todos los demás: desde el estímulo con 3 tramas hasta el de 11 tramas y, luego, se añadieron los estímulos de las tramas 21, 22, 23 y 24, que no se compararon con todos los demás, sólo se hizo con algunos para una simple comprobación. Este test se pasó a 12 informantes diferentes de los que habían realizado el test de identificación. Como se pretendía sólo comprobar si la discriminación funcionaba de igual manera que la identificación se omitieron los estímulos de las tramas 12 a 20 y el de 25. De esta forma reducíamos el número total de estímulos que se debían escuchar y que fueron 40 parejas finalmente.

3. RESULTADOS

3.1. Test de identificación:

Duración de las oclusivas Test auditivo



El gráfico muestra con exactitud las vicisitudes del test. Los datos se dan en tantos por ciento. Se observa que kaba es identificada entre 3 y 7 tramas, es decir, entre 26,4 y 61,6 ms; en este intervalo tenemos, pues, la interpretación laxa (o sonora, si se prefiere, aunque no existe ninguna pista de sonoridad). Además, se puede ver que cuanto menor es la duración, mayor es la identificación de [β]. Entre 8 y 16, podemos situar la identificación de [p], es decir, entre 70,4 y 140,8 ms. Y a partir de 17 (149,6 ms), la identificación es de [pp]. Obsérvese además que se dan tres puntos óptimos. 3 para [β] y 24 (211,2 ms) para [pp], ambas con un 100% de respuestas. Y 11 (96,8 ms) para [p], con un 97%. También [β] consigue el 97% con 4 y 5 tramas. [pp] repite el 97% con 20 y 21 tramas.

Todas ellas podríamos considerarlas puntos óptimos de la duración de las respectivas consonantes. Por otra parte, en las zonas fronterizas (7 y 8 / 16 y 17) bajan los porcentajes, como era de esperar.

3.2. Test de discriminación

Resultados: datos en porcentaje

		Respuestas de igualdad								
		3	4	5	6	7	8	9	10	11
3		92	92	75	50	42	8	0	0	
4	8		100	83	83	8	0	0	0	
5	8	0		92	75	92	17	25	0	
6	25	17	8		100	83	17	50	0	
7	50	17	25	0		58	25	25	8	
8	58	92	8	17	42		67	75	33	
9	92	100	83	83	75	33		92	67	
10	100	100	75	50	75	25	8		92	
11	100	100	100	100	92	67	33	8		

Respuestas de diferencia

Comprobación de los estímulos más largos:

(iguales/diferentes)

21/3: respuestas 17/83

22/4: respuestas 8/92

23/11: respuestas 8/92

22/24: respuestas 83/17

El test de discriminación nos da el mismo resultado que el de identificación. Los estímulos 3, 4, 5 y 6 entran dentro de la misma categoría. 7 y 8 constituyen la frontera. 9, 10 y 11 entran dentro de una nueva categoría. Lo mismo sucederá con los estímulos 22 y 24, que pertenecen a la misma categoría, mientras que 22 y 4 o 23 y 11 corresponden a categorías diferentes.

Como muestra de estos resultados ofrecemos tres gráficos muy significativos, ya que representan los dos extremos y la frontera:

Discriminación [b/p]

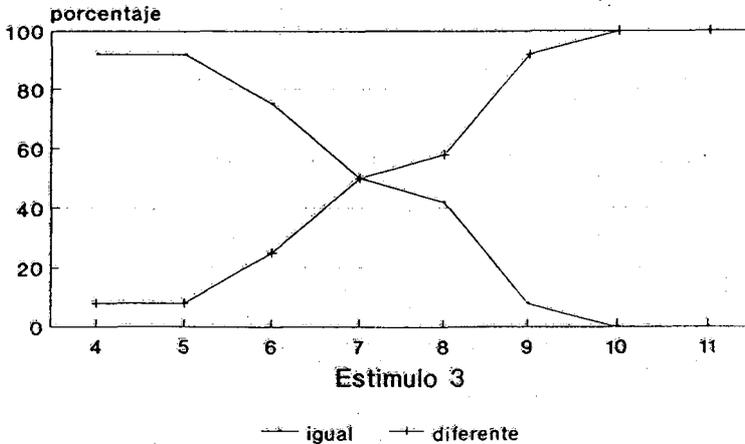


Gráfico 2: El estímulo con 3 tramas (26,4 ms) se muestra aquí igual que los estímulos con 4, 5, y 6 tramas; mientras que es totalmente diferente de los estímulos con 9, 10 y 11 tramas.

Discriminación [b/p]

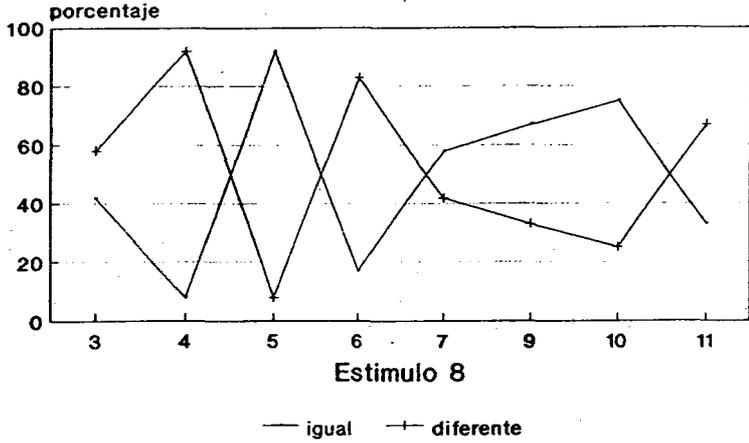


Gráfico 3: El estímulo con 8 tramas constituye la frontera entre ambas categorías, por eso muestra cruces continuos, que son signo de la indecisión de los oyentes cuando lo perciben frente a los demás.

Discriminación [b/p]

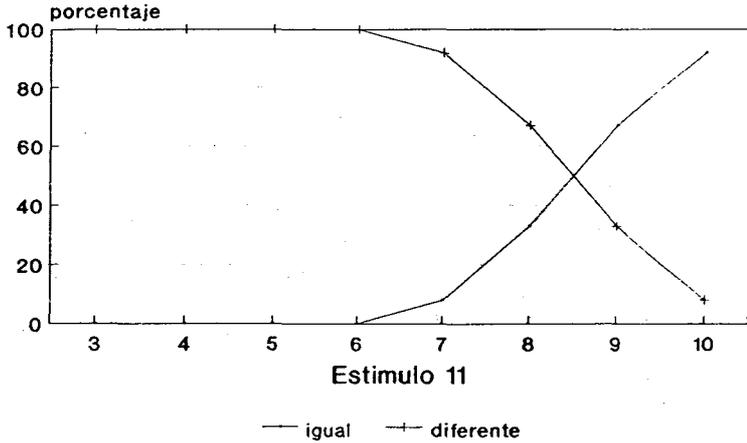


Gráfico 4: El estímulo con 11 tramas (96,8 ms) se muestra aquí igual que los estímulos con 9 y 10 tramas; mientras que es totalmente diferente de los demás.

4. DISCUSION

Creemos que estos experimentos validan suficientemente nuestra hipótesis. Tanto desde el punto de vista de la discriminación como desde el de la identificación, la duración es la pista más importante que permite distinguir entre /b/, /p/ y /pp/ y la duración ofrece además una categorización clara de las tres unidades.

La llamada geminación es simplemente una duración extralarga (a partir de las 150 ms), que no dudamos que pueda dar la sensación de que se reparte entre dos sílabas, sobre todo en posición intervocálica y si son sonidos con cierre continuado. Las "sordas" simples tienen una duración intermedia entre 70 y 140 ms. Y las "sonoras" tienen una duración siempre por debajo de los 70 ms. No ponemos en duda tampoco que en estos últimos sonidos existan de modo concomitante vibración de las cuerdas vocales, pero creemos que para nuestra percepción dichas vibraciones no son el elemento esencial de estas consonantes, sino su duración. Y ya que hemos hablado de duración, este elemento lo ponemos en relación con la tensión y distinguimos tres grados:

- pp - hipertensa
- p - tensa
- b - laxa

El experimento de identificación ha obtenido los mismos resultados que el realizado por Alvarez González (1978) para el inglés y el de Liberman et alii (1961). Obsérvese el gráfico que él proporciona en p. 122 y compárese con el nuestro y se verá un parecido extraordinario. Quizás, nuestros resultados han sido mucho más nítidos, más resolutivos. En los suyos nunca se alcanzan porcentajes superiores al 90%. Entre [p] y [pp] hay una zona de confusión entre los 150 y los 190 ms. En nuestro caso hay también, en esa zona, confusiones, pero no en tal alto grado. Las diferencias existentes son lógicas teniendo en cuenta la diferente realización de las oclusivas inglesas y las españolas. En ambos casos, los resultados ponen en evidencia la importancia de la duración en la identificación de estos sonidos.

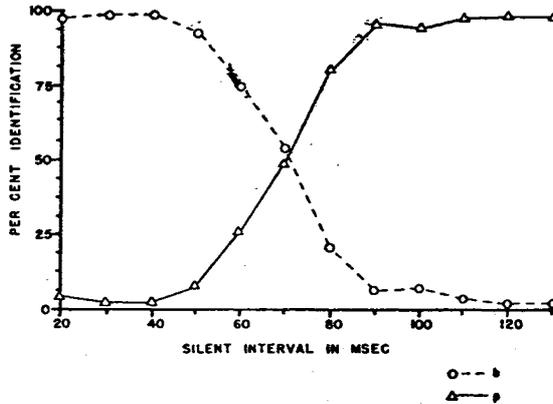


Gráfico 5 (Apud Liberman et alii, 1961:183) Compárese este gráfico con el número 1 en las 11 primeras tramas y se comprobará que son prácticamente iguales.

El test de discriminación ha obtenido resultados muy parecidos a los encontrados por Liberman et alii (1961) para el inglés.

La duración es una línea que nosotros hemos podido categorizar en tres partes; pero estamos seguros de que si una lengua puede categorizar cuatro, se hubieran podido establecer cuatro cortes en la línea del tiempo. De hecho, los ingleses "oyeron" [b] y, los españoles [β] - suponemos-, puesto que en posición intervocálica los españoles no poseen [b].

Por otra parte, en ambas lenguas se percibió esa tercera categoría que no es fonológica en ninguna de las dos lenguas, lo cual nos plantea una duda razonable sobre el planteamiento que estos autores realizan al indicar que la categorización es aprendida por la larga experiencia que tenemos con la propia lengua. Si esto fuera así, no se percibiría esa tercera categoría que no posee "status" fonológico. Además, ¿hasta qué punto las categorías son fonológicas?, como pretende la mayor parte de autores que han hablado de esta cuestión. Creemos que existen categorías fonéticas que pueden ser percibidas en una lengua dada aunque no tengan una

función fonológica, como sucede con la geminación. En todo caso, habría que estudiar por qué unas categorías fonéticas no se perciben si no tienen función fonológica y otras sí.

5. CONCLUSION

Mi hipótesis ha sido confirmada: la duración del silencio es un pista suficiente para la identificación y la discriminación de las oclusivas labiales y, además, podemos afirmar que existe percepción categorial de /b/ y /p/ basada en las diferencias de duración exclusivamente.

6. REFERENCIAS

- ALVAREZ GONZALEZ, J.A. 1978. "El análisis por síntesis", *Revista Española de Lingüística*, 8/1, pp.117-124, Gredos, Madrid.
- FRY, D.B., ABRAMSON, A.S., EIMAS, P.D. y LIBERMAN, A.M. 1962. "The identification and discrimination of synthetic vowels", *Language and Speech*, 5, pp.171-179.
- LIBERMAN, A.M., HARRIS, K.S., EIMAS, P.D., LISKER, L. y BASTIAN, J. 1961. "An effect of learning on speech perception: The discrimination of durations of silence with and without phonemic significance", *Language and Speech*, 4, pp.175-195.
- LIBERMAN, A.M., HARRIS, K.S., HOFFMAN, H.S. y GRIFFITH, B.C. 1957. "The discrimination of speech sounds within and across phoneme boundaries", *Journal of Experimental Psychology*, 53, 1957, pp.358-368.
- MARTINET, A. 1955. *Economía de los cambios fonéticos*, Gredos, Madrid, 1974.

- MARTINEZ CELDRAN, E. 1991a. "Tensión frente a sonoridad en las consonantes mates del castellano", en *Fonética experimental: Teoría y práctica*, pp.131-141, Síntesis, Madrid.
- MARTINEZ CELDRAN, E. 1991b. "Relevancia de los elementos de las oclusivas sordas del castellano según la discriminación auditiva", en *Fonética experimental: Teoría y práctica*, pp.115-130, Síntesis, Madrid.
- MARTINEZ CELDRAN, E. 1991c. "Duración y tensión en las oclusivas no iniciales del español: un estudio perceptivo", *Revista Argentina de Lingüística*, 7, pp.51-71.
- MARTINEZ CELDRAN, E. 1992. "Un parámetro fonético en el fondo de la lenición protorromance: La duración", en *Homenaje a E. de Bustos Tovar*, vol I, pp.117-136, Salamanca.
- PISONI, D.B. 1978. "Speech perception", en ESTES, W.K. (Ed.), *Handbook of Learning and Cognitive Processes*, 6, pp. 167-233, Hillsdale, N.J., Erlbaum.
- REPP, B.H. 1984. "Categorical perception: Issues, methods, findings", en LASS, N.J. (Ed.), *Speech and language: Advances in Theory and practice*, vol 10, pp.243-335, Academic Press, New York.