

**LA INVARIACIÓN ACÚSTICA EN EL PUNTO DE ARTICULACIÓN  
DE LAS OCLUSIVAS DEL ESPAÑOL.**

**BELÉN NÚÑEZ-ROMERO LINARES**  
*Laboratori de Fonètica, Facultat de Filologia,  
Universitat de Barcelona*

### RESUMEN

El objetivo de este trabajo es comprobar si, para cada punto de articulación de las consonantes oclusivas castellanas, el espectro muestreado de corta duración a partir del comienzo de la explosión presenta propiedades acústicas invariantes. S. E. Blumstein y K.N. Stevens llevaron a cabo esta teoría para la lengua inglesa con un resultado positivo. Pero una de las pretensiones de su teoría es que esto mismo se va a encontrar en todas las lenguas. Los investigadores americanos defienden que los espectros de las oclusivas se adaptan a tres tipos diferentes de plantillas: difusa-ascendente para las alveolares -dentales en nuestra lengua-, difusa-descendente para las labiales y compacta para las velares. En este trabajo se han realizado de una forma sistemática los mismos experimentos llevados a cabo por S.E. Blumstein y K.N. Stevens y se ha comprobado que los resultados que éstos obtuvieron para la lengua inglesa no se cumplen en la lengua castellana, excepto para las consonantes velares, que se ajustan perfectamente a la plantilla establecida -compacta-, alcanzando el 100% de efectividad. Sin embargo tan sólo el 48% de las labiales se ajusta exclusivamente a la plantilla difusa-descendente (si se estudian las primeras 26m.s. en el espectro). Si se estudian las primeras 10 m.s., tan sólo se ajusta el 31'5%. Con las dentales ocurre algo similar. Si se estudian las primeras 26 m.s., se adapta exclusivamente al patrón difuso-ascendente sólo el 26'5% de las dentales. Si se analizan las primeras 10 m.s. el resultado es del 40%. Por tanto ni las consonantes labiales ni las dentales alcanzan el 85% esperado de adaptación a las plantillas ideadas para ellas.

### ABSTRACT

The aim of this study is to check whether the short-time spectra sampled at the onset of Spanish stops, present invariant acoustic properties that specify consonantal place of articulation. S.E. Blumstein and K.N. Stevens (1979) hypothesized that the spectra match three classes of spectral shapes or templates: diffuse-rising, diffuse-falling and compact, corresponding to alveolar -dental in Spanish-, labial and velar consonants, respectively. In our study we have followed systematically, the experimental procedures designed by Blumstein and Stevens to test whether this hypothesis could also be proved in Spanish. However, our results were not as encouraging as theirs: only 48% of the labial consonants matched against the diffuse-falling shape -using a 26 ms time window-. When we used a 10 ms window, only 31,5% matched the template. Dental consonants were better classified when we used the shorter-time window. Velar consonants were a case apart: 100% of the

spectra were correctly classified by the templates.. Therefore, the overall results obtained for Spanish do not support Blumstein and Stevens's hypothesis, since neither labials, nor dentals were consistently classified by their corresponding templates.

## **1. INTRODUCCIÓN**

La teoría presentada por S.E.Blumstein y K.N.Stevens (1979) hipotetiza que la señal del habla está altamente estructurada al contener patrones acústicos invariantes que configuran los rasgos fonéticos y que permanecen invariantes a través de los hablantes, los contextos fonéticos y las lenguas.

Parten de la suposición de que, para cada punto de articulación de las consonantes oclusivas, el espectro muestreado de corto tiempo a partir del comienzo de la explosión debería presentar propiedades acústicas invariantes, las cuales sólo especificarían el punto de articulación de la consonante sin importar qué vocal vaya detrás.

Para comprobar su hipótesis estudiaron un muestreo del espectro de un amplio número de sílabas compuestas de consonante oclusiva - sorda y sonora- seguida y precedida de vocal, en diferentes hablantes.

Para el análisis de los espectros han sido ideadas unas plantillas que recogen las tres formas espectrales: difuso-ascendente para las consonantes alveolares, difuso-descendente para las consonantes labiales y compacto para las velares.

El objetivo de este trabajo ha sido realizar para la lengua castellana y de forma sistemática los mismos experimentos que han llevado a cabo los investigadores S.E.Blumstein y K.N.Stevens para el inglés, con el fin de comprobar si los hallazgos de estos autores se cumplen en nuestra lengua.

## **2. METODOLOGÍA**

Para el estudio de la lengua inglesa han sido analizadas las consonantes oclusivas labiales, velares y alveolares. Puesto que la fonética inglesa difiere de la castellana, ha sido necesario estudiar, en lugar de las alveolares, nuestras dentales. Además, nuestras oclusivas sordas no son aspiradas, a diferencia de las *dl* inglés.

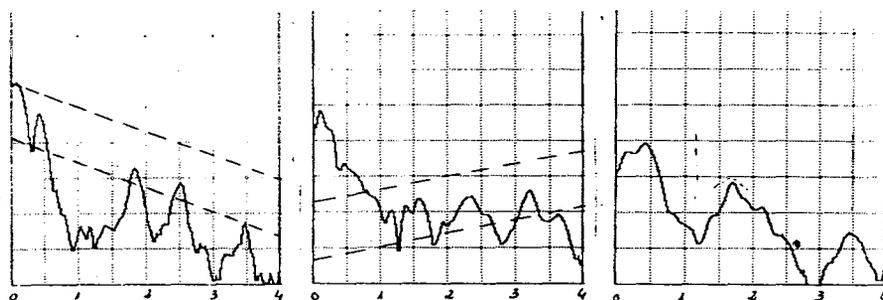
Otra diferencia respecto del inglés está en que nuestra lengua no posee palabras que finalicen con consonante oclusiva. Por ello se han analizado exclusivamente sílabas compuestas por consonante oclusiva seguida de vocal.

Todo este estudio está basado en los datos obtenidos de las voces de cuatro hablantes adultos, dos voces masculinas y dos femeninas. En total son 600 las sílabas analizadas. Estas sílabas están compuestas por las consonantes oclusivas sordas [p], [t], [k] y sus correspondientes sonoras [b], [d], [g], seguidas de las cinco vocales. Cada hablante ha pronunciado estas 30 sílabas en cinco ocasiones y en órdenes diferentes para lograr un análisis más fiable. Por tanto, cada hablante ha grabado un total de 150 sílabas.

Para su estudio se ha utilizado el Sona-Graph modelo 5500. Una vez entrada cada sílaba en él, se ha trabajado con el SET UP nº14, que, además del sonograma, muestra un espectro promediado de los 4.000 primeros Hzs.

En el análisis del inglés, las pistas invariantes en cada señal se han obtenido estudiando las primeras 10-20 milésimas de segundo a partir de la relajación de la oclusiva -la explosión-. En la lengua castellana se ha analizado, por una parte, el espectro que muestra las primeras 26ms. y, por otra, el de las primeras 10ms.

Los investigadores S.E. Blumstein y K.N. Stevens defienden que los espectros de las oclusivas se adaptan a tres tipos diferentes de plantillas: difusa-ascendente para las alveolares -dentales en nuestra lengua-, difusa-descendente para las labiales y compacta para las velares. Las plantillas difusas poseen dos líneas de referencia separadas por unos 10 Decibelios. Para el castellano ha sido necesario separarlas 15 dB. Para las velares no hay una plantilla cuyas líneas de referencia indique que son compactas; basta con observar en los espectros un pico prominente entre 1.200Hz y 3.500 Hz. Si dos picos están separados por menos de 500 Hz se consideran un solo pico. (Figura 1)



DIFUSO-DESCENDENTE.  
[pe] (26 ms.)

DIFUSO-ASCENDENTE  
[du] (10 ms.)

COMPACTO  
[ka] (26 ms)

Figura 1

Los investigadores S.E.Blumstein y K.N.Stevens requieren a los espectros una serie de condiciones para poder encajar en las plantillas difusas. En la plantilla difusa-descendente al menos dos picos deben caer en las líneas de referencia, uno por debajo de 2.400Hz y el otro entre 2.400Hz y 3.600Hz. No hay ninguna condición en las amplitudes de los picos por debajo de los 1.200Hz. En la plantilla difusa-ascendente al menos dos picos deben caer entre las dos líneas de referencia, y un pico por encima de los 2.200Hz. debe ser más alto en amplitud que el otro pico por debajo de éste en frecuencia. Todos los espectros que no aceptan estas condiciones son rechazados.

Si se tienen en cuenta todas estas premisas en la lengua castellana, pocos espectros encajan en las plantillas. Por ello, sólo ha sido tenido en cuenta que al menos dos picos entrasen en las dos líneas de referencia, sin prestar atención a la frecuencia en qué aparecen. Incluso se puede decir que los primeros 1.000Hz. han sido casi siempre necesarios para clasificar los espectros.

### 3. RESULTADOS Y DISCUSIÓN

#### 3.1.- Labiales

El 81,5% de las labiales -sordas y sonoras- se adapta correctamente a la plantilla difusa- descendente. Este resultado sería satisfactorio,

puesto que los investigadores ingleses consideran efectivo el resultado igual o superior al 80%. Sin embargo, un 33% de ese total también se adapta correctamente al patrón difuso-ascendente, por lo que sólo el 48% de las oclusivas labiales pueden ser clasificadas exclusivamente como labiales sin ser confundidas. (Figura 2).

Estos resultados se han obtenido del análisis de las primeras 26ms. a partir de la relajación de la consonante. Al estudiar las primeras 10ms., los resultados son inferiores: el 31,5% se adapta sin confusión al patrón difuso-descendente, el 18% se adapta a ambos patrones y el 46% no puede ser clasificado. (Figura 2).

### LABIALES

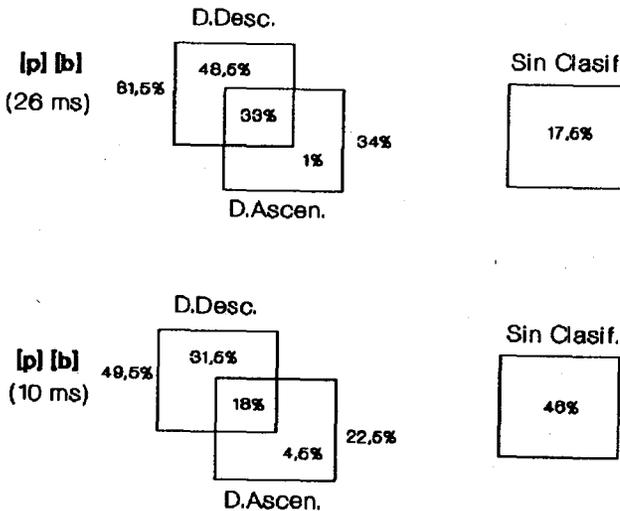
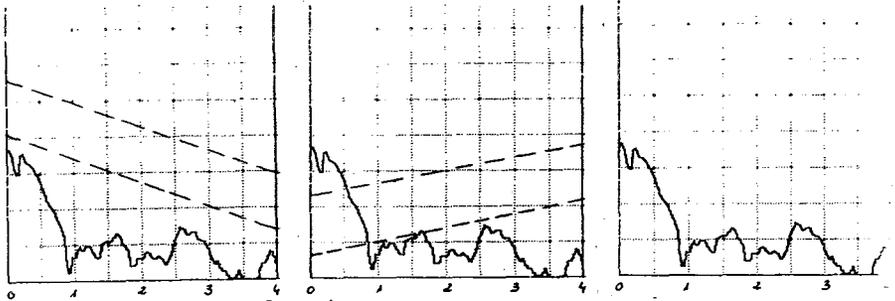


Figura 2

Si analizamos las primeras 26ms. a partir de la explosión de la oclusiva obtenemos un porcentaje menor de labiales sin clasificar - 17,5%- que si estudiamos las primeras 10ms. -48%-[Figura 3].

**LABIAL SIN CLASIFICAR. [pa] (10 ms).****NO DESCENDENTE.****NO ASCENDENTE.****NO COMPACTA.***Figura 3: realización labial*

Las **labiales sonoras** -tanto si se estudian las primeras 26ms. como las primeras 10ms.- se adaptan en mayor porcentaje al patrón difuso-descendente que las labiales sordas. Sin embargo, ni unas ni otras alcanzan el 80% de adaptación a su plantilla difusa-descendente.

Al analizar las primeras 26ms. en las labiales **sonoras**, el 89% acepta correctamente el patrón difuso-descendente. Sin embargo, el 36% de éstas también acepta el patrón difuso-ascendente, por lo que el resultado mencionado no puede ser considerado válido; sólo el 53 % de las labiales sonoras acepta únicamente el patrón difuso- descendente. [Figura 4].

Pese a las confusiones que se dan entre labiales y dentales, y pese al bajo porcentaje de efectividad en la adaptación de las consonantes labiales a su patrón difuso-descendente, cabe señalar que el número de labiales que se adapta solamente a la plantilla difusa-descendente es muchísimo mayor que el número de labiales que se adapta exclusivamente al patrón difuso-ascendente. Esto indica una predisposición de las labiales a adaptarse al patrón que S.E.Blumstein y K.N.Stevens crearon para ellas; sin embargo, es sólo una predisposición, pues no alcanzan el 80% esperado de adaptación y además hay un gran índice de confusión con las dentales y muchas labiales sin clasificar.

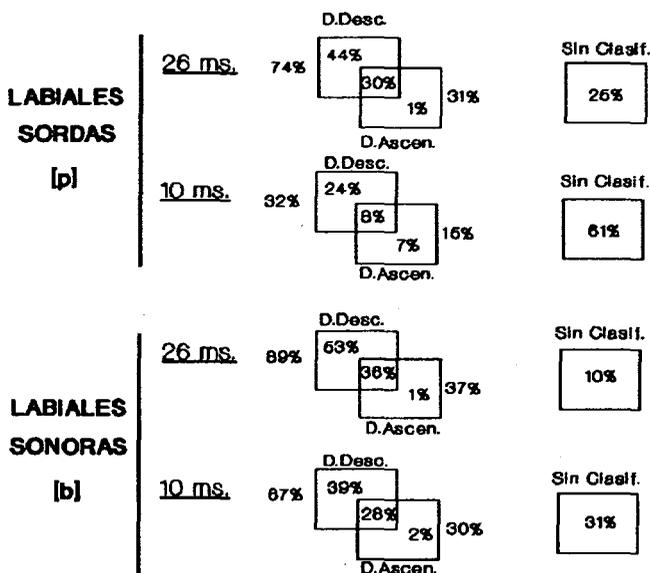


Figura 4

Las labiales sordas en el contexto de las vocales [i e a] son mejor clasificadas que en contexto [o u], tanto si se estudian las primeras 26 ms. como las primeras 10 ms. No obstante, se obtienen mejores resultados al estudiar las primeras 26 ms. ; en el contexto de las vocales [i e a] todas las labiales pueden ser clasificadas por una de las dos plantillas difusas -no queda ninguna labial sin clasificar-. (Figura 5).

Las labiales sordas -26 ms.- en el contexto [e a] alcanzan el 100% de adaptación al patrón difuso-descendente, y en el contexto de [i] el 95%. Sin embargo, estos óptimos resultados no son efectivos, puesto que si tenemos en cuenta las labiales que también aceptan el patrón difuso-ascendente, los porcentajes anteriores quedan reducidos a la mitad. En contexto de [e] el 50% de ese 100% se adapta también a la plantilla difusa-ascendente; en contexto de [a] es el 40% del 100% el que se confunde, y en contexto de [i] es el 45% del 95% el que se adapta a ambas plantillas. (Figura 5).

Ya hemos mencionado la predisposición que se observa en las labiales a adaptarse a la plantilla difusa-descendente. Ciertamente existe esta predisposición, puesto que a pesar del gran número de labiales que aceptan ambas plantillas, encontramos que ninguna labial sorda en contexto de las vocales [e a o u] puede ser interpretada exclusivamente como dental -las que pueden ser interpretadas como dentales también se interpretan como labiales-. En contexto de [i] el 5% de las labiales sordas sí se clasifican exclusivamente como dentales. (Figura 5).

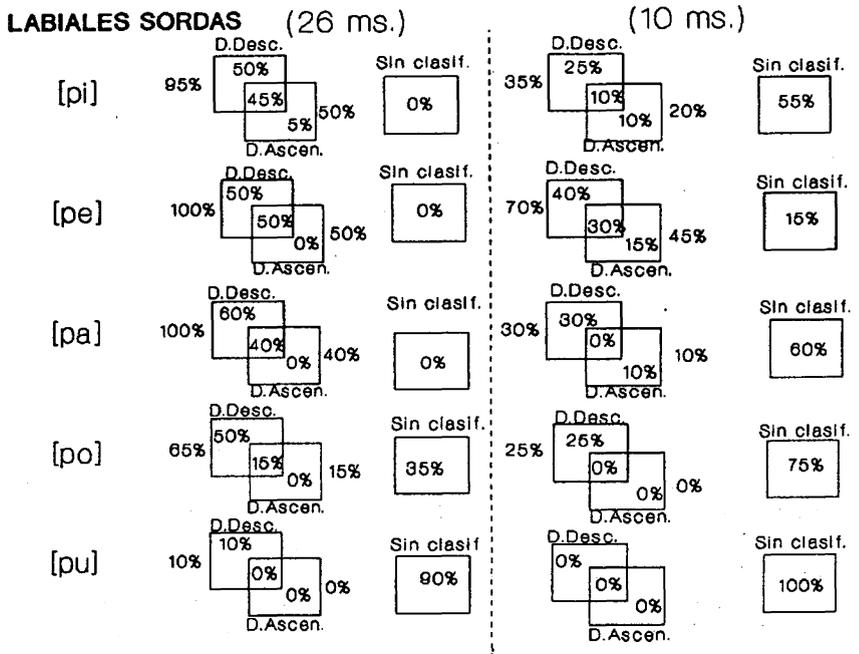


Figura 5

Las labiales **sordas** en contexto de [u] -26 ms.- son las que presentan mayor dificultad para ser interpretadas. Sólo el 10% de ellas aceptan el patrón difuso-descendente, mientras que el 90% restante no puede ser clasificado por ningún tipo de patrón. Si analizamos las primeras 10 ms. no hallamos ninguna labial sorda en contexto de [u] que pueda ser -clasificada. (Figura 5).

Obtenemos similares resultados al estudiar las primeras 26 ms. en las labiales sonoras. En contexto de [e a] las labiales sonoras alcanzan el 100% de aceptación del patrón difuso-descendente, y en contexto de [i] son aceptadas el 95% de ellas. Sin embargo, en el contexto de [e] se confunden con las dentales el 40% de ese 100%, y en el contexto de [a] también el 40% es confundido. El 70% de las labiales sonoras seguidas de [i] se adapta a ambas plantillas. (Figura 6).

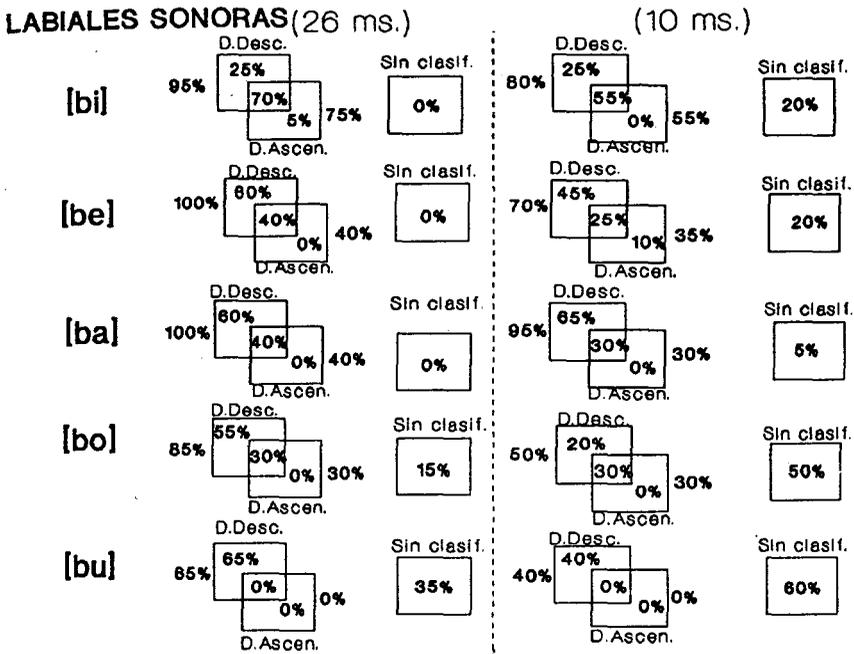


Figura 6

Las labiales sonoras en el contexto de [e a o u] muestran idéntico resultado que las labiales sordas en cuanto al porcentaje de labiales que son interpretadas exclusivamente como dentales: ninguna de ellas se adapta exclusivamente al patrón difuso-ascendente. En el contexto de [i e a] todas las labiales sonoras son clasificadas por algún patrón, al igual que ya hemos visto en las labiales sordas. En contexto de [o] el 15% no

puede ser clasificado, y en contexto de [u] es un 35% el que se queda sin clasificar.

Los resultados obtenidos tras analizar las primeras 10 ms. son aún menos efectivos que los que acabamos de mostrar. [Figura 6].

### 3.2. Dentales

Al estudiar las primeras 26 ms. en los espectros de las dentales, un 36,5% se éstas se adapta correctamente a los patrones de difuso-ascendente y de difuso-descendente. El 26,5% de las dentales se adapta exclusivamente al patrón difuso-ascendente y el 27,5% al de difuso descendente. El 9,5% no puede ser clasificado. [Figura 7].

#### DENTALES

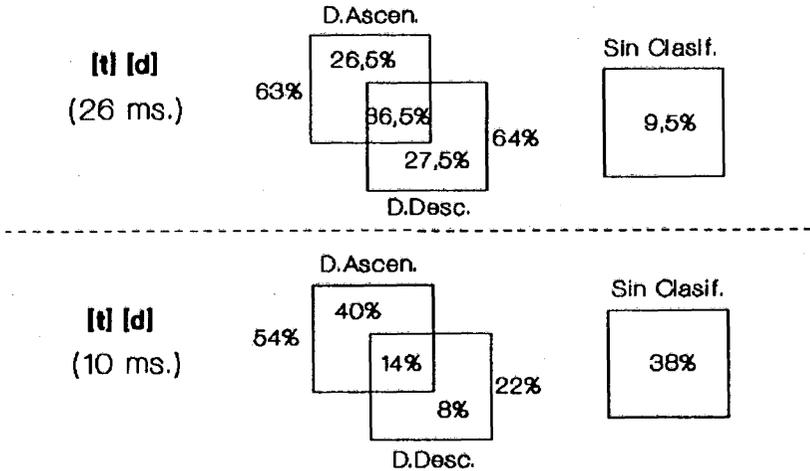


Fig. 7

Por tanto, las consonantes dentales son mejor clasificadas como labiales que como dentales si seguimos la teoría de que las plantillas difusas diferencian unas consonantes de las otras. No obstante, el porcentaje de aceptación de ambas plantillas es muy similar; sólo varía en un 1%. [Figura 7].

Al analizar las primeras 10 ms. el resultado varía respecto a que no hay tanta semejanza entre las consonantes interpretadas como difusas-ascendentes y las interpretadas como difusas-descendentes. El 40% se adapta exclusivamente a la plantilla difusa-ascendente, mientras que sólo el 8% es interpretado únicamente como difuso-descendente. El 14% se adapta a ambos tipos de plantilla, y el 38% no se puede clasificar. [Figura 7].

Las **dentales sonoras** -26 ms.- se confunden con las labiales, mientras que de las dentales sordas se confunden el 24%. [Figura 8].

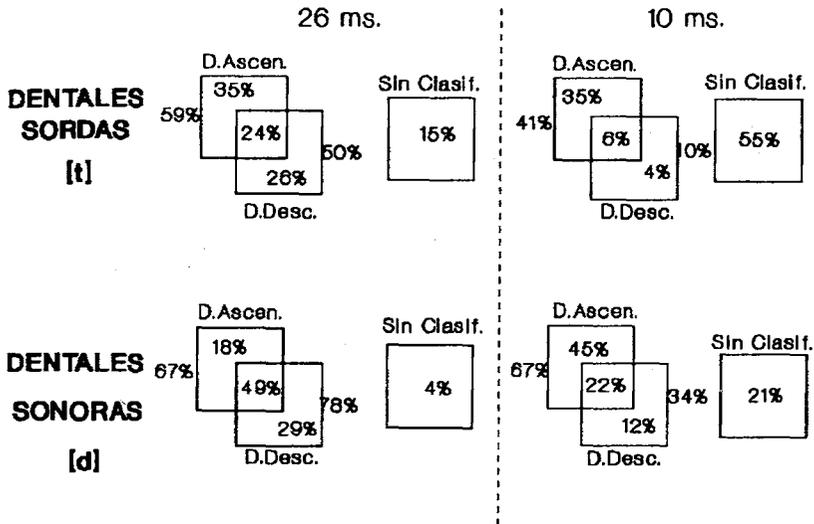


Fig. 8

Por otra parte, las dentales sonoras son mejor interpretadas como labiales que como dentales: el 29% de ellas se interpretan únicamente como labiales, mientras que exclusivamente como dentales sólo son interpretadas el 18%. No ocurre igual si estudiamos las primeras 10 ms., puesto que el 45% de las dentales sonoras son interpretadas únicamente como dentales, mientras que el 12% son interpretadas como labiales. [Figura 8].

En cuanto a las dentales sordas -10 ms.- el porcentaje de consonantes que aceptan únicamente la plantilla difusa-descendente es bastante bajo, puesto que sólo se consideran labiales sin confusión el 4%. Sin embargo, este resultado no es muy significativo, puesto que más de la mitad de las dentales sordas -el 55%- no pueden ser clasificadas por ningún tipo de plantilla. [Figura 8].

En resumen, ni dentales sordas ni sonoras alcanzan el 80% de adaptación a la plantilla difusa-ascendente.

Las **dentales sordas -26 ms.-** no presentan grandes diferencias en sus resultados según el contexto vocálico. Una diferencia significativa es que en contexto de [e a] las dentales sordas son interpretadas mejor como labiales que como dentales, aunque no alcanzan el 80% de aceptación exclusiva del patrón difuso-descendente. El 50% de las dentales sordas en contexto de [e] acepta únicamente la plantilla difusa-descendente. Del 50% restante, el 40% acepta ambas plantillas, y únicamente el 10% es considerado como consonante dental propiamente dicha. [Figura 9].

#### DENTALES SORDAS

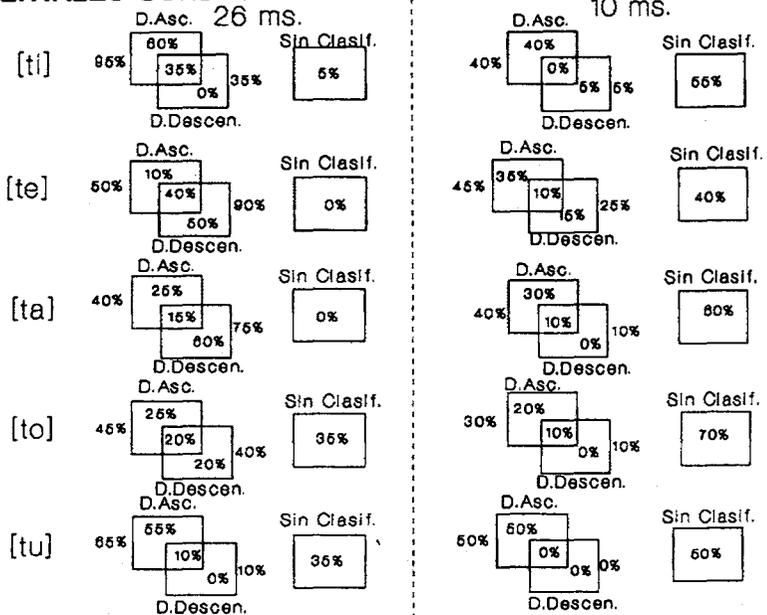


Fig. 9

En contexto de [a] el 60% de las dentales sordas se interpretan sin confusión como consonantes labiales. Del 40% restante, el 15% se adapta a las dos plantillas difusas, y el 25% se considera exclusivamente consonante dental. (Figura 9).

Sin embargo, al analizar las primeras 10 ms. los resultados son muy distintos. Pese a que no se alcanza el 80% de aceptación del patrón difuso-ascendente, hay más predisposición a la aceptación de este patrón que a la del patrón difuso-descendente. Esto ocurre no sólo en el contexto de [e a], sino que ocurre con todas las vocales. También cabe decir que el porcentaje de dentales sordas que queda sin clasificar es muchísimo más elevado que si se estudian las primeras 26 ms., llegando incluso a alcanzar -en contexto de [o]- el 70% de consonantes no clasificadas por ningún patrón. (Figura 9).

En contexto de [u] las dentales sordas están más predispuestas a adaptarse a su patrón que en los demás contextos. Si observamos las primeras 10 ms., el 50% de las dentales sordas en contexto de [u] aceptan solamente el patrón difuso-ascendente y el 50% restante no puede ser clasificado. Al analizar las primeras 26 ms. obtenemos un 55% de aceptación exclusiva del patrón difuso-ascendente y un 35% de consonantes que no se pueden clasificar. El 10% restante se adapta a ambas plantillas difusas. (Figura 9).

Las **dentales sonoras** -26 ms.- en el contexto de las vocales [e a o] también presentan mayor predisposición a interpretarse como labiales que como dentales. (Figura 10).

En contexto de [e] el 30 % se adapta únicamente al patrón difuso-descendente, el 65% acepta ambas plantillas difusas y solamente el 5% se interpreta exclusivamente como dental. Por tanto, el 95% de las dentales sonoras en contexto de [e] se adapta al patrón difuso-descendente de las labiales.

En contexto de [a] los resultados son aún más sorprendentes. El 100% de las dentales sonoras en este contexto acepta el patrón difuso-descendente de las labiales. Sin embargo, el 55% de ese total también acepta la plantilla difusa-ascendente. No hay ninguna dental sonora en contexto de [a] que sea interpretada exclusivamente como dental. (Figura 10).

Si observamos las primeras 10 ms. ya no hallamos esa predisposición de las dentales sonoras a ser mejor interpretadas como labiales que como dentales en el contexto de las vocales [e a o]. Sin

embargo, los resultados tampoco indican una clara preferencia por aceptar el patrón difuso-ascendente, puesto que en esos contextos -como con [i u]- las dentales sonoras no superan ni el 50% de aceptación exclusiva de su patrón difuso-ascendente. (Figura 10).

### DENTALES SONORAS

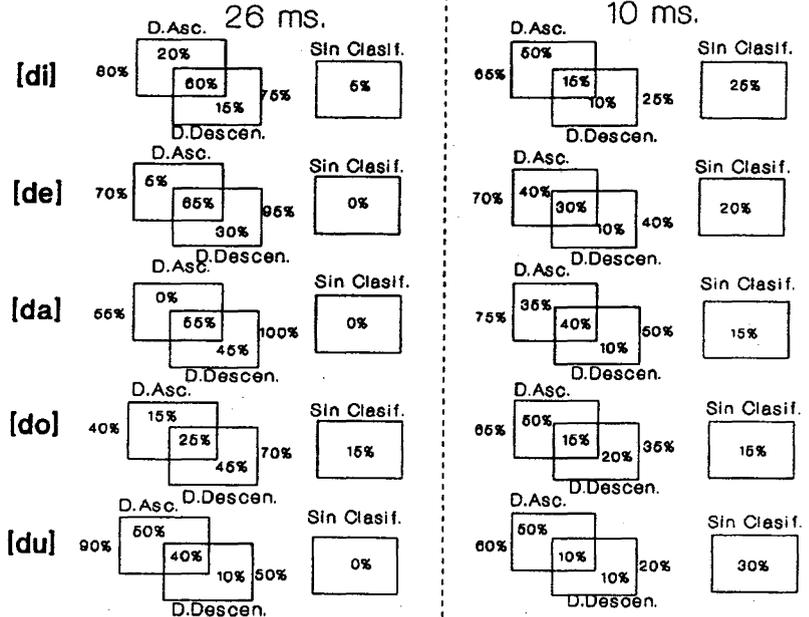


Fig. 10

### 3.3.-Velares

Las consonantes velares son las únicas oclusivas que han dado los resultados esperados.

Para su estudio han sido analizadas sólo las primeras 26 ms. a partir de la primera barra de explosión.

El 100% de las velares, tanto sordas como sonoras y en el contexto de todas las vocales, son bien clasificadas como compactas. (Figura 11)

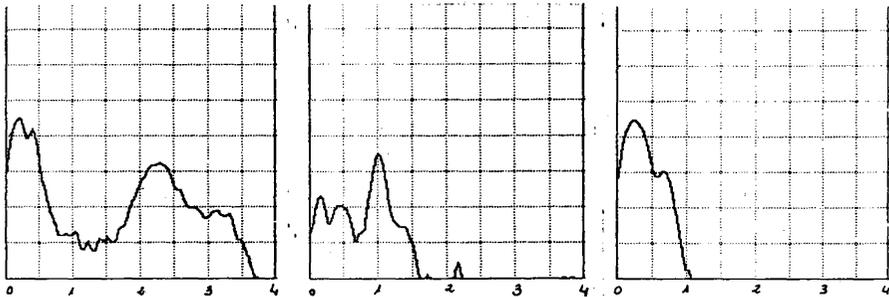
**VELARES SORDAS Y SONORAS**



Fig. 11

En el contexto de las vocales [i e a], el 100% son compactas, ya que se observa un pico prominente entre los 1.200 Hz. y los 3.500 Hz. [Figura 12].

Sin embargo, en el contexto de las vocales [o u], el 100% de las velares son también compactas, pero el pico prominente se halla antes de los 1.200 Hz. El hecho de que estos espectros aparezcan más graves en el contexto [o u] tiene su explicación si tenemos en cuenta que con [o u] el locus de las velares se sitúa hacia los 800 Hz. o 1.000 Hz.



[ge] (26 ms)

[ko] (26 ms)

[ku] (26 ms)

Figs. 12, 13 y 14

También cabe destacar que las velares en contexto de [o] presentan el pico prominente antes de los 1.200 Hz. (Figura 13), mientras que en contexto de [u] se encuentra antes de los 1.000 Hz. (Figura 14).

#### 4. CONCLUSIÓN

S.E.Blumstein y K.N.Stevens basaron su estudio en los espectros que mostraban las primeras 10 o 20 ms. a partir de la relajación de la consonante, y tomaban, dentro de ese corto espacio de tiempo, aquellos espectros que mejor encajasen en las plantillas elaboradas.

Al comienzo de este trabajo también fueron tomados aquellos espectros que mejor se adaptaban a las plantillas, sin tener en cuenta cuántas milésimas de segundo -entre 10 y 26 ms.- se habían acotado con los cursores del sonógrafo. Por tanto, si ante una consonante labial obteníamos un espectro difuso-descendente con 18 ms. ya no era necesario seguir buscando, pues ya se había conseguido el resultado esperado.

No obstante, muchos espectros aceptaban ambas plantillas difusas, cosa que no ocurría en el estudio del inglés. Por ello fue conveniente cambiar el método de trabajo, puesto que si ante una consonante labial nos proponíamos buscar un espectro difuso-ascendente también era posible encontrarlo. Por tanto, debido a la confusión existente entre consonantes labiales y dentales se estimó más oportuno realizar dos análisis paralelos: uno, tomando las primeras 10 ms. a partir de la explosión de la consonante y otro, tomando las primeras 26 ms. -excepto en las oclusivas velares, que sólo han sido analizados los espectros que muestran las primeras 26 ms. a partir de la primera barra de explosión-.

El análisis de las primeras 26 ms. ha resultado ser más efectivo que el de las primeras 10 ms., pues ofrece un menor porcentaje de consonantes sin clasificar. En cambio, el porcentaje de consonantes que encajan en ambas plantillas difusas es mayor que el obtenido al estudiar las primeras 10 ms.

De todas formas, tanto si se estudian las primeras 10 ms. como las primeras 26 ms., los resultados no superan el 80 % de aceptación exclusiva de la plantilla correspondiente a cada consonante -labiales y dentales-.

Con los resultados obtenidos podemos afirmar que la teoría de la invariación acústica, tal y como ha sido estudiada para el inglés, no se

cumple para la lengua castellana. Las plantillas que se han establecido para realizar esta investigación han sido elaboradas para alcanzar el mayor grado de efectividad posible; sin embargo, un porcentaje muy elevado de los espectros de las labiales y de las dentales, se adapta indistintamente a las plantillas de difuso-ascendente y de difuso-descendente.

De ser modificadas las plantillas para no dar lugar a tanta confusión entre oclusivas labiales y dentales, el resultado final era aún menos efectivo que el obtenido, pues aumentaba el porcentaje de oclusivas sin clasificar.

De hecho, estos resultados no sorprenden demasiado, puesto que Lahiri, Gewirth y Blumstein ya obtuvieron unos resultados similares al realizar un estudio comparativo entre el malayalam y el francés. El porcentaje tan bajo de oclusivas dentales clasificadas correctamente por el patrón difuso-ascendente les indicó que la teoría que se basa en las características del espectro en la zona próxima a la explosión no puede distinguir de forma adecuada las oclusivas labiales de las dentales, ni tampoco puede agrupar las dentales con las alveolares ni en malayalam ni en francés. [Ver Villalba ]

En su estudio no hallan problemas al estudiar las labiales; son las dentales las que no aceptan mayoritariamente el patrón difuso-ascendente. En cambio, en el estudio del castellano, los problemas aparecen tanto con las dentales como con las labiales; tanto se confunden unas como las otras.

No obstante, los resultados obtenidos no indican que la teoría de la invariación acústica no sea una teoría válida, puesto que nuestro oído percibe claramente las diferencias existentes entre las diferentes consonantes. Está claro que hay algún indicio invariante que distingue concretamente las consonantes labiales de las dentales, pero aún hay que descubrirlo.

**5. REFERENCIAS BIBLIOGRÁFICAS**

BLUMSTEIN, S.E. y STEVENS, K.N. (1979): "Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants", *JASA*, 66(4), 1001-1007.

LAHIRI, A., GEWIRTH, L. y BLUMSTEIN, S.E. (1984): "A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study", *JASA*, 76 (2), 391-404.

VILLALBA, X. (1995): "Los invariantes acústicos y el punto de articulación de las oclusivas en español: Una revisión de Lahiri, Gewirth y Blumstein", en este volumen.