

**LA INVARIACIÓN ACÚSTICA EN LAS FRICATIVAS DEL  
CASTELLANO.  
ESTUDIO PERCEPTIVO**

JAVIER ROMERO Y ANA M<sup>a</sup> FERNÁNDEZ PLANAS  
*Laboratori de Fonètica, Facultat de Filologia  
Universitat de Barcelona*

**RESUMEN**

El presente trabajo estudia a través de dos experimentos si se puede considerar la amplitud como índice acústico invariante en las fricativas del castellano. El método que ha servido de base para llevar a cabo distintas manipulaciones con los estímulos ha consistido en intercambiar la amplitud de [s] y [x] por la de [f] y [θ] y viceversa. Los resultados obtenidos en tests perceptivos constituidos tanto por las sílabas manipuladas como por los ruidos de fricación aislados han demostrado que la amplitud por sí sola no puede ser considerada un índice suficiente para el reconocimiento de las fricativas del castellano.

**ABSTRACT**

The present study examines whether we can regard amplitude as an invariant acoustic cue for fricatives in Castilian. To this aim, we edited a series of stimuli changing the amplitude values of [s] and [x] for those of [f] and [θ] and viceversa. The results obtained from listening tests -made up not only of edited syllables but also of isolated frication noises-, showed us that amplitude on its own cannot be considered a reliable cue for identification of Spanish fricatives.

## 1. INTRODUCCIÓN

El marco teórico en el que se sustenta este trabajo es el de la invariación acústica, es decir, parte de la base según la cual existen en los sonidos índices invariantes a través de hablantes, contextos y lenguas diferentes. Tomamos, concretamente, como modelo un estudio de S. Behrens y S. Blumstein (1988a) acerca de las fricativas sordas del inglés: [f], [ʃ], [s], [θ].

Acústicamente, las fricativas se distinguen entre sí por la duración del ruido de la fricación (Behrens y Blumstein, 1988b), las propiedades espectrales de cada una (Stevens, 1960; Behrens y Blumstein, 1988b; Martínez Celdrán, 1984 y Quilis, 1981) y la amplitud relativa del ruido en relación con la vocal (Stevens, 1960; Stevens y Blumstein, 1988a). Perceptivamente, el peso específico de cada uno de estos parámetros puede diferir mucho. Estudios precedentes han sugerido que el papel de la amplitud es el más relevante (Heinz y Stevens, 1961), con diferencia al resto de parámetros.

El presente trabajo tiene como objetivo estudiar el papel que desempeña la amplitud del ruido de fricación en la percepción de las consonantes fricativas sordas del castellano: [f], [θ], [s], [x].

Nuestro estudio, que emula al hecho por Behrens y Blumstein (1988a) adaptando al castellano las fricativas objeto de observación, centra su interés en comprobar si las fricativas [f] y [θ] serían percibidas por los oyentes como [s] y [x] cuando a aquéllas se les colocase la amplitud de estas últimas, y viceversa. Esto es, delimitar hasta qué punto la amplitud puede ser considerada como un factor decisivo en la percepción del punto de articulación de las consonantes fricativas sordas. Para ello hemos llevado a cabo dos experimentos.

## 2. PRIMER EXPERIMENTO

### 2.1. Método

Los estímulos usados procedían de la voz natural de diez hablantes masculinos sin entrenamiento fonético que pronunciaron el corpus completo en una cabina insonorizada y grabados en un magnetófono. El corpus estaba formado por estructuras CV donde C era una de las cuatro fricativas sordas castellanas y V era una de las vocales extremas del

triángulo vocálico, es decir, el corpus estaba constituido por las siguientes secuencias:

[fa]	[fi]	[fu]
[θa]	[θi]	[θu]
[sa]	[si]	[su]
[xa]	[xi]	[xu]

El siguiente paso consistió en computerizar la totalidad de estímulos mediante el programa SP2 de Philips y, posteriormente, seleccionar de entre las muestras las realmente aptas para nuestros propósitos, es decir, aquellas que se reconocieran perfectamente pues sabida es la confusión incluso en voz natural entre las fricativas, especialmente entre [f] y [θ]. Se buscaba además que fueran válidas las series completas (mismo informante y misma vocal con todas las fricativas) puesto que la manipulación de la amplitud del ruido de fricación debía hacerse sobre los estímulos CV que tuvieran la misma vocal y pertenecieran al mismo informante. Finalmente resultaron aptas 32 muestras divididas en 8 grupos: dos para la vocal media, cuatro para la vocal alta anterior y dos para la alta posterior.

A continuación sintetizamos las 32 muestras resultantes mediante el editor de unidades fonéticas «VOXSY2» de TEVOX. Una vez hecho esto, pudimos manipularlas convenientemente para llevar a cabo el primer experimento con el fin de comprobar si la amplitud era un atributo distintivo a la hora de percibir el punto de articulación de las consonantes fricativas sordas del castellano.

La manipulación se llevó a cabo agrupando las muestras del siguiente modo: [f + V] y [θ + V] frente a [s + V] y [x + V]. Después tomamos la AM (amplitud) del ruido de fricación y del segmento vocálico (puesto que la intensidad que miramos no es absoluta sino relativa ya que lo interesante es la diferencia respecto a la vocal que la acompaña) de las sílabas [s + V] y se la colocamos a las sílabas [f + V] y a las sílabas [θ + V] que pertenecían a la misma serie, o sea, mismo informante y misma vocal. Por ejemplo: tomamos la Am del ruido de fricación y de la vocal de las muestras [fi] y [θi] de un informante determinado y la cambiamos por la AM de la fricativa y la vocal de la muestra [si] del mismo informante.

Después tomamos la AM de la consonante y la vocal de las sílabas [x + V] y se la volvimos a colocar a las sílabas [f + V] y [θ + V] del mismo informante y la misma vocal. Esto es, por ejemplo tomamos la AM de la secuencia CV de las muestras [fi] y [θi] de un informante dado y la

cambiamos por la AM de la consonante y la vocal de la muestra [xi] del mismo sujeto.

Posteriormente, mediante el proceso inverso colocamos la AM del ruido de fricación y de la vocal de las muestras [f+V] a las muestras [s+V] y [x+V]; finalmente, lo mismo hicimos con la AM de la secuencia CV en las muestras [θ+V] sobre las muestras [s+V] y [x+V].

En esquema:

AM de [s+V] sobre [f+V]	AM de [x+V] sobre [f+V]
AM de [s+V] sobre [θ+V]	AM de [x+V] sobre [θ+V]
AM de [f+V] sobre [s+V]	AM de [θ+V] sobre [s+V]
AM de [f+V] sobre [x+V]	AM de [θ+V] sobre [x+V]

Obtuvimos, finalmente, 64 muestras totales, es decir, 8 muestras por cada grupo.

Otra manipulación consistió en eliminar el segmento vocálico de cada una de las muestras resultantes de la prueba anterior. Mediante este procedimiento aislamos el ruido de fricación de cada sílaba para presentarlo sin contexto vocálico. El resultado fue 64 nuevas muestras constituidas únicamente por los ruidos de las consonantes fricativas.

Una vez hecho esto, grabamos en orden aleatorio todas las muestras manipuladas (con estructura CV y con estructura C) para presentarlas en un test de percepción a un grupo de informantes.

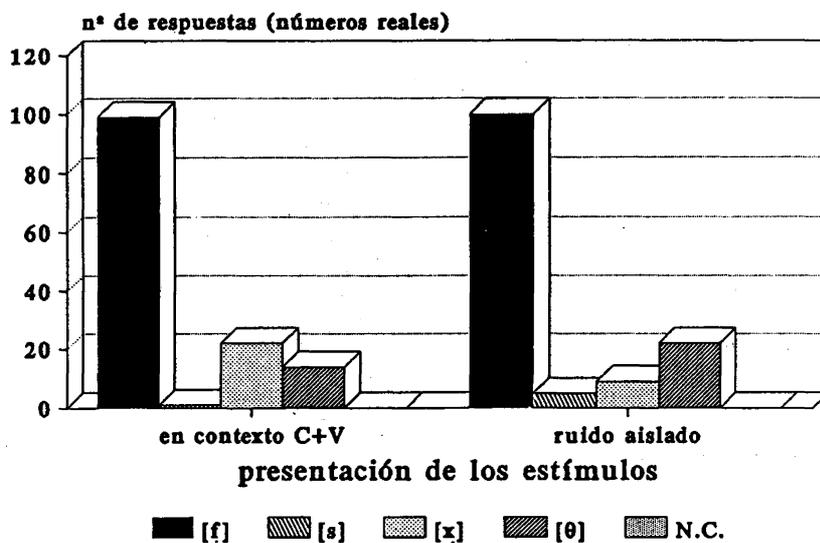
## 2.2. Resultados

Los resultados<sup>1</sup> del test de percepción obtenidos aparecen en las siguientes tablas y gráficos de barras:

---

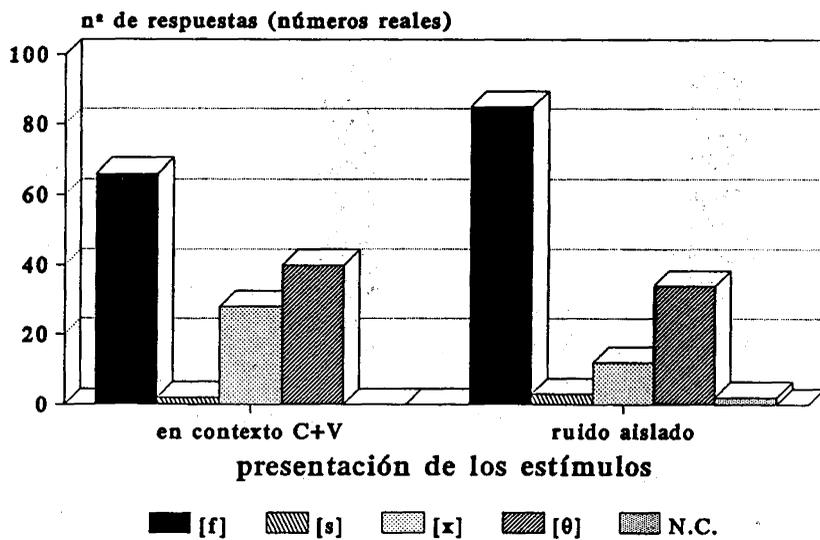
<sup>1</sup> El número de resultados (136) procede de contar cada manipulación con 8 estímulos presentados a 17 informantes.

[f] -> [s] <sup>2</sup>		
	Contexto CV	Ruido aislado
[f]	99	100
[s]	1	5
[x]	22	9
[θ]	14	22

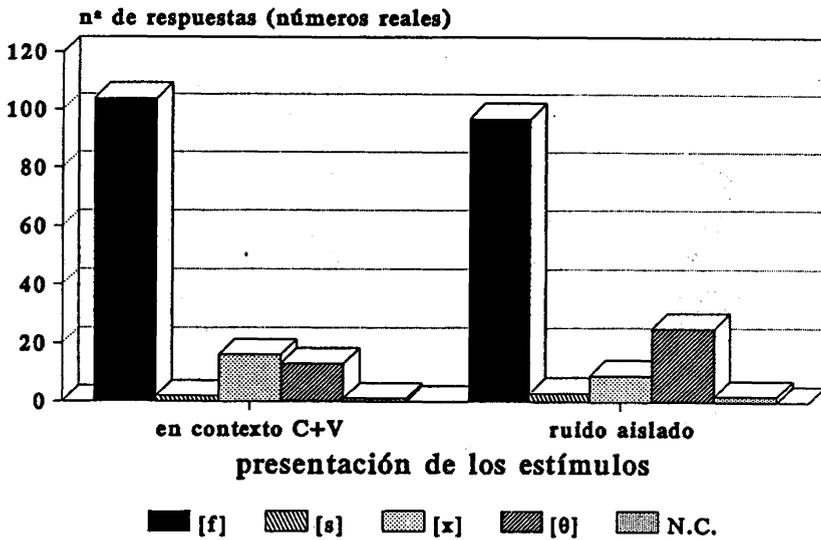


<sup>2</sup> Se lee: AM de [s] colocada en las muestras [f] o AM de [f] modificada hasta convertirla en [s].

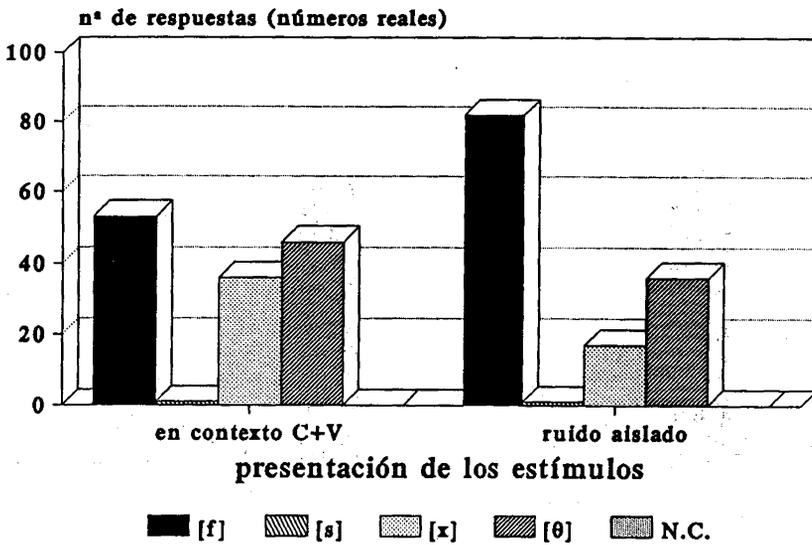
[θ] --> [s]		
	Contexto CV	Ruido aislado
[f]	66	85
[s]	2	3
[x]	28	12
[θ]	40	34
N.C.	0	2



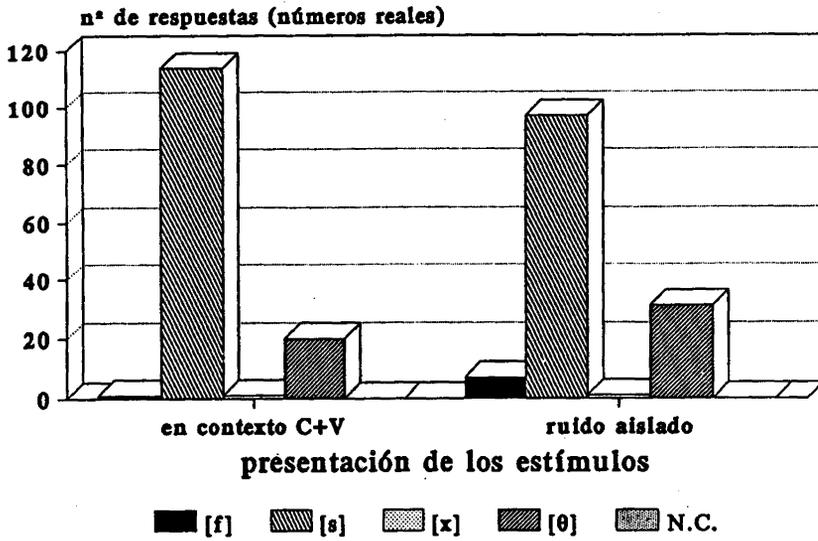
[f] -> [x]		
	Contexto CV	Ruido aislado
[f]	104	97
[s]	2	3
[x]	16	9
[θ]	13	25
N.C.	1	2



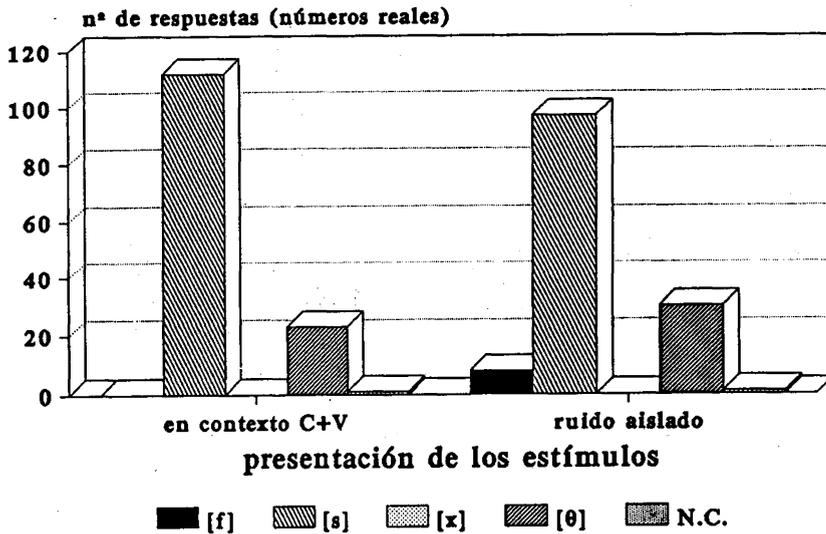
[θ] → [x]		
	Contexto CV	Ruido aislado
[f]	53	82
[s]	1	1
[x]	36	17
[θ]	46	36



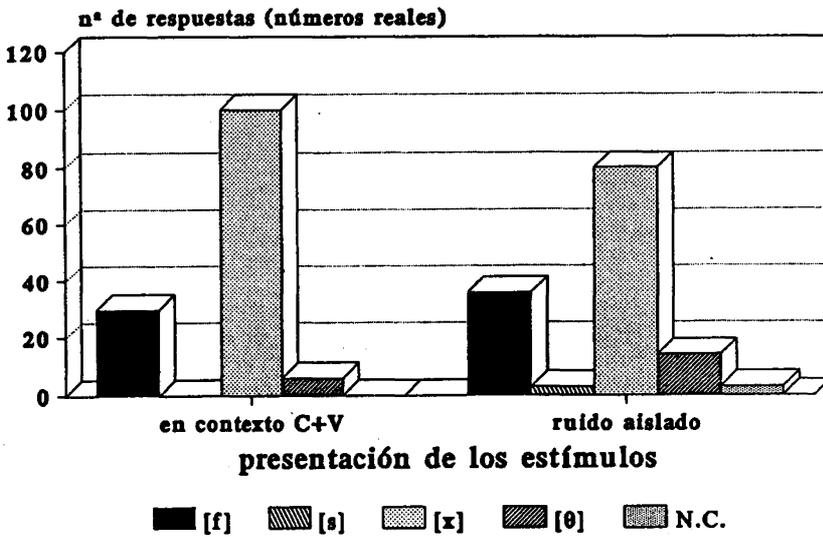
[s] -> [f]		
	Contexto CV	Ruido aislado
[f]	1	7
[s]	114	97
[x]	1	1
[θ]	20	31



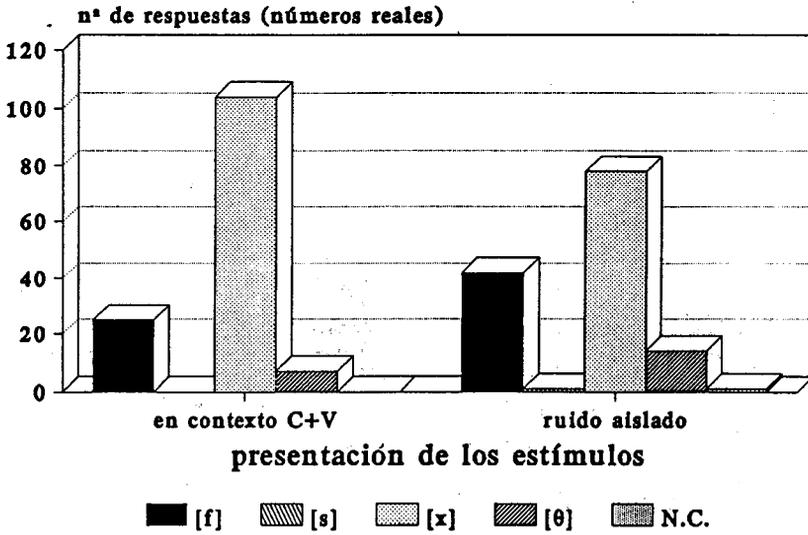
[s] -> [θ]		
	Contexto CV	Ruido aislado
[f]	0	8
[s]	112	97
[x]	0	0
[θ]	23	30
N.C.	1	1



[x] -> [f]		
	Contexto CV	Ruido aislado
[f]	30	36
[s]	0	3
[x]	100	80
[θ]	6	14
N.C.	0	3



[x] --> [θ]		
	Contexto CV	Ruido aislado
[f]	25	42
[s]	0	1
[x]	104	78
[θ]	7	14
N.C.	0	1



Observamos que colocando la amplitud de [s] en las fricativas con menor intensidad [f] y [θ] los resultados son mayoría de respuestas de labiodental en todos los casos y es prácticamente nula la percepción de la sibilante [s] contra lo que cabía esperar en principio. En la manipulación que consistió en colocar la amplitud de [x] en las fricativas de escasa intensidad la mayor parte de las respuestas siguen siendo en ambos contextos (CV y C aislada) [f] y hay que destacar únicamente que cuando la fricativa manipulada fue [θ] hubo también bastantes respuestas [x] y [θ], especialmente en contexto CV, pero siempre en menor medida que [f].

Cambiando la amplitud de [s] por las de las fricativas [f] y [θ], de mucha menor intensidad, hay una mayoría de respuestas con la sibilante en las dos condiciones (con y sin vocal) y en todos los casos un número parecido de respuestas [θ] tanto si la amplitud colocada es la de [f] como la de [θ] y tanto si el contexto es con vocal como si es sin ella. Alterando la amplitud de [x] hasta convertirla en la de [f] y [θ] hay algunas respuestas labiodentales e interdentes, especialmente de las primeras, pero con todo el mayor número de respuestas han sido de [x].

### 3. SEGUNDO EXPERIMENTO

#### 3.1. Método

En este experimento el objetivo seguía siendo operar con la amplitud de las fricativas pero en este caso haciéndolo con las medias de intensidad obtenidas de cada tipo de ellas. Es decir, a partir de los ocho grupos iniciales en el experimento anterior calculamos la AM media de cada serie de muestras [s+V], [x+V], [f+V] y [θ+V], para cada timbre vocálico. Obtuvimos un único valor para cada una de las siguientes secuencias:

[fa]	[fi]	[fu]
[θa]	[θi]	[θu]
[sa]	[si]	[su]
[xa]	[xi]	[xu]

El segundo paso consistió en hacer una media para cada vocal entre [f] y [θ] y otra entre [s] y [x].

Tras estos pasos previos la manipulación fundamental consistió en colocar para cada vocal la media del grupo [s]-[x] sobre las muestras [f] y [θ] y la media del grupo [f]-[θ] sobre las muestras [s] y [x].

En esquema:

1. AM media de las cuatro muestras [sa]-[xa] colocada en las muestras [fa] y [θa].

2. AM media de las cuatro muestras [fa]-[θa] colocada en las muestras [sa] y [xa].

3. AM media de las ocho muestras [si]-[xi] colocada en las muestras [fi] y [θi].

4. AM media de las ocho muestras [fi]-[θi] colocada en las muestras [si] y [xi].

5. AM media de las cuatro muestras [su]-[xu] colocada en las muestras [fu] y [θu].

6. AM media de las cuatro muestras [fu]-[θu] colocada en las muestras [su] y [xu].

El cambio de valores, igual que en el experimento anterior, lo llevamos a cabo con el editor de unidades fonéticas «VOXSY2» de TEVOX.

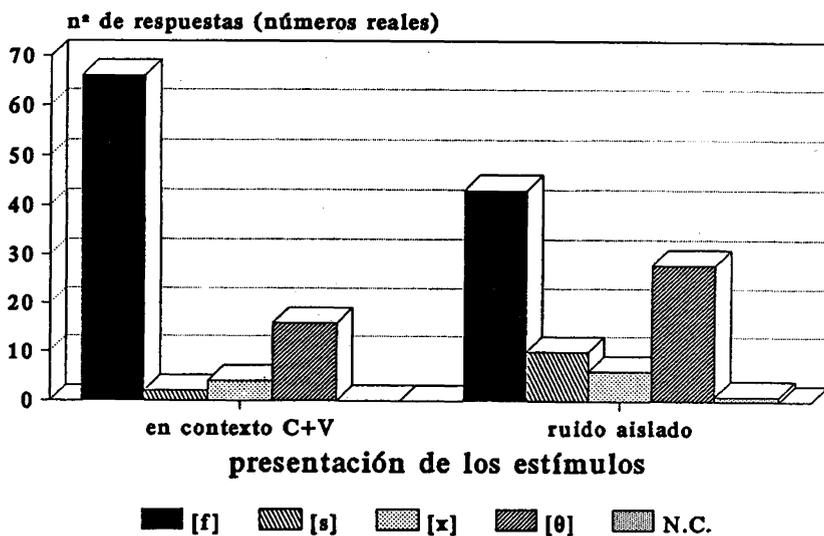
Resultaron 32 muestras manipuladas de fricativa más vocal, cantidad que doblamos posteriormente al volver a operar con estas muestras resultantes para quitarles la vocal y aislar así el ruido de fricación.

Las 64 secuencias totales [32 de estructura CV y 32 de inarmónico aislado] fueron grabadas, como en el primer experimento, de modo aleatorio para presentarlas en un test de percepción a otro grupo de informantes. Los resultados aparecen en el apartado que viene a continuación.

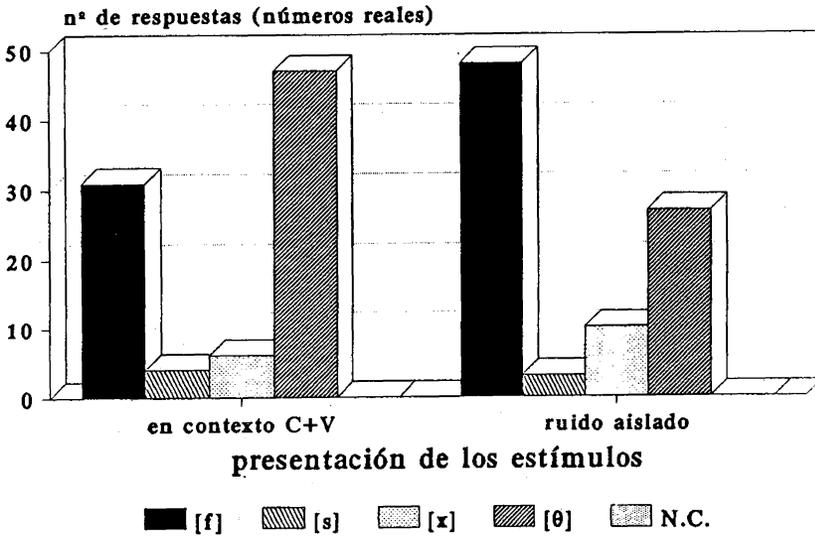
### 3.2. Resultados

Presentamos los resultados del test pasado a 11 de nuevo en tablas y gráficos de barras.

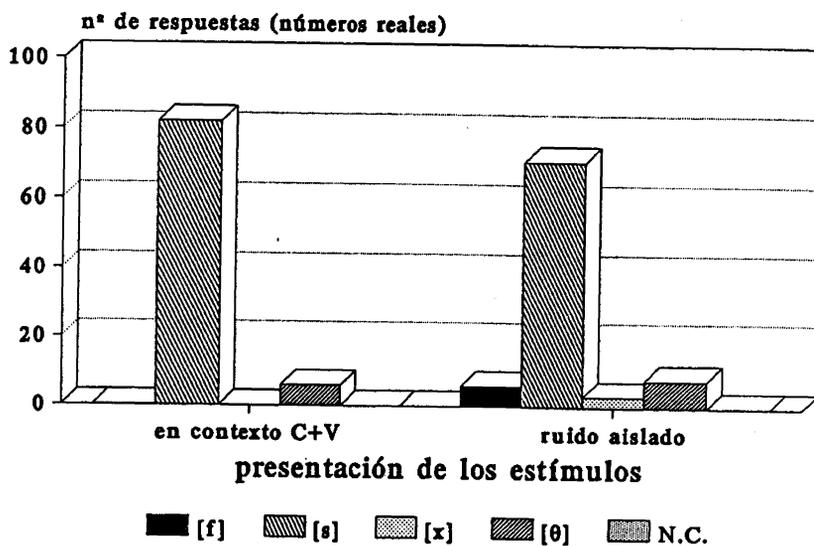
[f] -> media entre [s] y [x]		
	Contexto CV	Ruido aislado
[f]	66	43
[s]	2	10
[x]	4	6
[θ]	16	28
N.C.	0	1



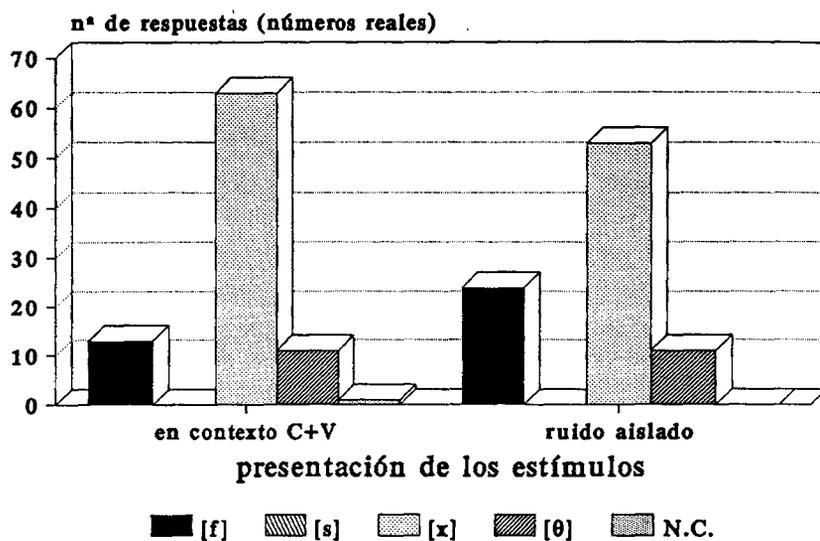
[θ] -> media entre [s] y [x]		
	Contexto CV	Ruido aislado
[f]	31	48
[s]	4	3
[x]	6	10
[θ]	47	27



[s] --> media entre [f] y [θ]		
	Contexto CV	Ruido aislado
[f]	0	6
[s]	82	71
[x]	0	3
[θ]	6	8



[x] --> media entre [f] y [θ]		
	Contexto CV	Ruido aislado
[f]	13	24
[s]	0	0
[x]	63	53
[θ]	11	11
N.C.	1	0

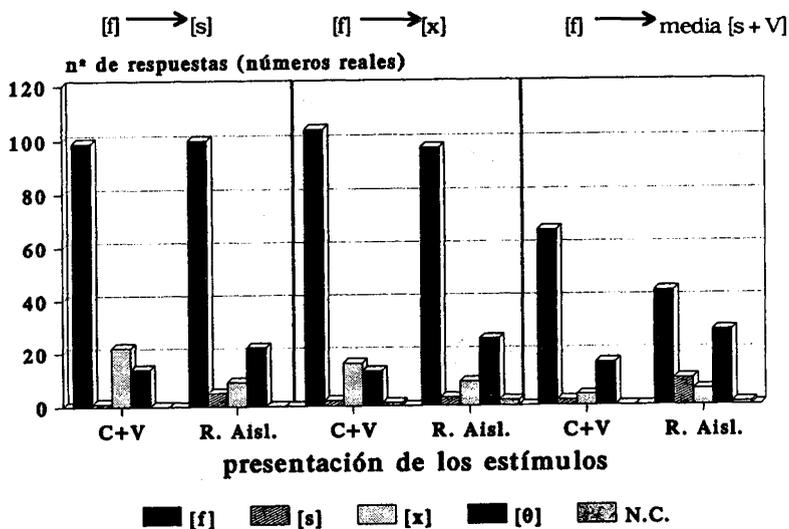


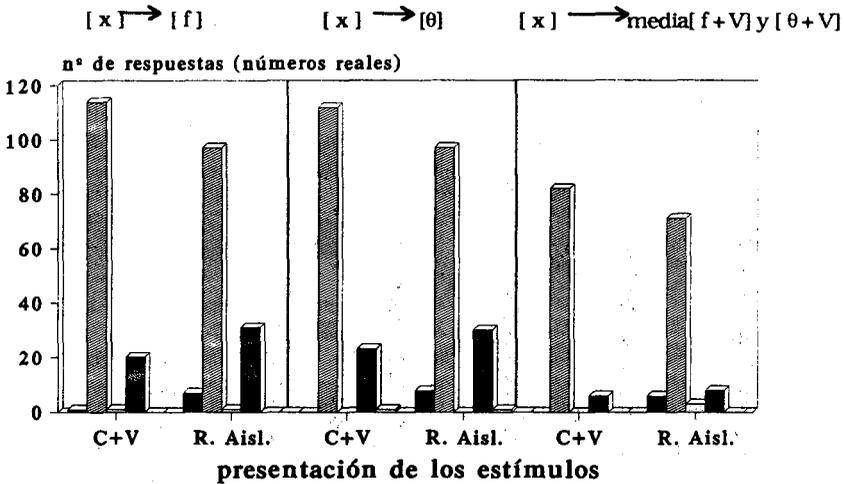
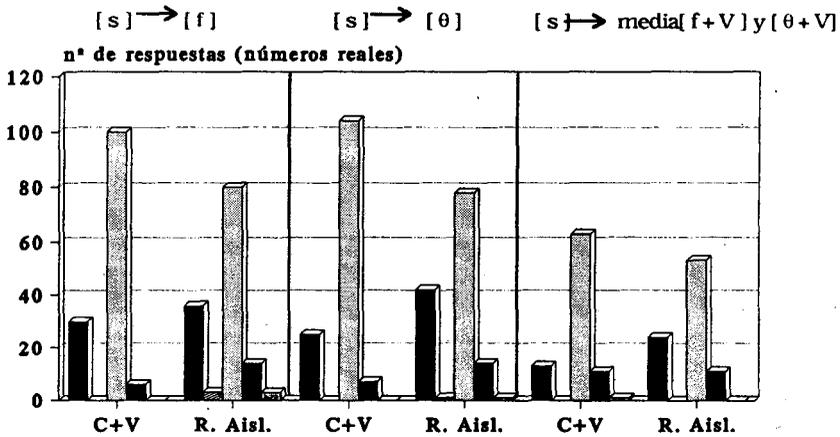
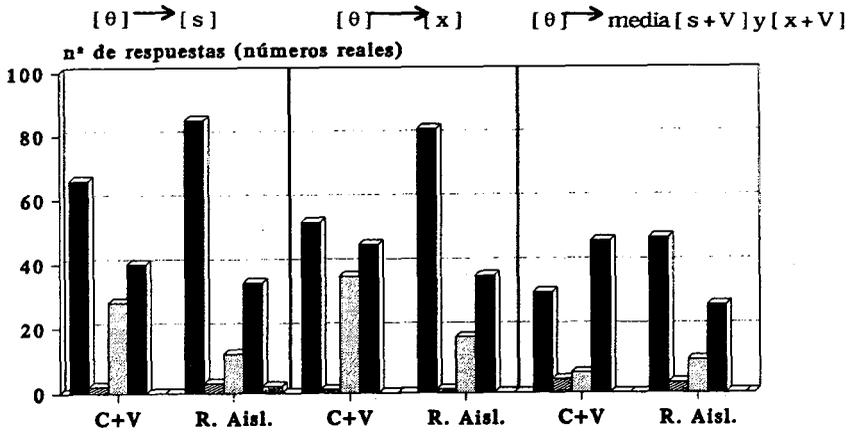
Sustituyendo la amplitud de [f] por la media de las amplitudes de las fricativas más intensas el resultado refleja un espectacular número de respuestas [f] en contexto CV; al tratarse de la fricativa aislada sigue predominando la respuesta [f] pero su número ha disminuido y ha aumentado el resultado [θ]. Al cambiar la amplitud de [θ] por la misma media los resultados nos dicen que el mayor número de respuestas ha recaído en [θ] seguido de [f] en contexto CV, proporción que se invierte al presentar la fricativa aislada, lo cual resulta sorprendente puesto que parámetros del estímulo [f] no han intervenido para nada en esta manipulación.

El hecho de poner la amplitud media entre [f] y [θ] a [s] no ha hecho que la percepción del sonido resultante dejara de ser [s] en ninguno de los contextos. El poner dicha media a [x] ha provocado que el resultado se percibiera mayoritariamente como [x] pero con un destacable papel de [f] y [θ] sobre todo en la condición de ruido aislado.

#### 4. DISCUSIÓN

Una vez analizados los resultados del primer y del segundo experimento comprobamos que tanto en uno como en otro siguen percibiéndose los sonidos [f], [s] y [x], pese a haberse manipulado, en las dos condiciones de estudio, es decir, tanto en estructura CV como en el ruido de fricación aislado. En el caso de [θ] podemos decir que en el primer experimento la respuesta mayoritaria fue [f] en ambos contextos; en el segundo, la mayoría de respuestas [f] se dio en condición aislada, en la estructura CV fue la fricativa interdental sobre la que se había realizado la manipulación de la amplitud percibida principal. Lo comprobamos en los gráficos que aparecen a continuación.





■ [f]    ▨ [s]    □ [x]    ■ [θ]    □ N.C.

Así pues, el hecho de colocar la amplitud de una consonante fricativa sorda a otra distinta no da lugar a que los oyentes perciban esta última. Por el contrario, siguen percibiendo, de hecho, la misma fricativa sobre la que se ha llevado a cabo la modificación. Por lo tanto, a la vista de los resultados de los experimentos realizados cuyas pautas metodológicas han quedado expuestas en páginas precedentes, parece que la amplitud no es un parámetro lo suficientemente distintivo por sí solo como para que pueda ser considerado atributo distintivo en la percepción de las consonantes fricativas sordas del castellano, por un lado [f] y [θ] y, por otro [s] y [x], como sugiere McCasland (1979 a, b). Nuestros resultados son similares a los obtenidos por Behrens y Blumstein (1988a) y difieren de estudios previos en los cuales los efectos perceptivos de manipulaciones de amplitud son muy grandes incluso cuando el ruido de fricación y las transiciones formánticas hacia la vocal siguiente son poco compatibles (Guerlekian, 1981).

## **5. POSIBLES FACTORES DISTINTIVOS EN LA PERCEPCIÓN DEL PUNTO DE ARTICULACIÓN DE LAS FRICATIVAS SORDAS.**

### **5.1. Método**

Una vez comprobado, a través de los dos experimentos efectuados, que la amplitud en castellano no es el factor decisivo, por lo menos por sí solo, a la hora de percibir de forma distinta los cuatro puntos de articulación de los diferentes sonidos llevamos a cabo la revisión de otros parámetros tales como la frecuencia media de los formantes consonánticos y de los anchos de banda que afectan en nuestro sistema, fundamentalmente, a la amplitud del formante al que se refieren (ej. B1 hace referencia a F1). Con esta revisión llevada a cabo con un programa estadístico pretendíamos extraer algún dato que realmente pudiera constituir un factor distintivo en la percepción de las fricativas sordas castellanas.

Revisamos por un lado aquellas muestras manipuladas de fricativa-vocal en las que el cambio de amplitud había supuesto que los informantes percibieran mayoritariamente, en los test de audición, la fricativa cuya amplitud había sido colocada. Por ejemplo, la percepción de [sa] habiendo colocado la AM de [s] sobre la [f] de la sílaba inicial [fa]. Por otro lado, también aquellas muestras en las cuales la manipulación no supuso que los informantes percibieran otra fricativa distinta a aquella sobre la que se había realizado la manipulación, por ejemplo la percepción de [fa] habiendo colocado la AM de [s] a la sílaba inicial [fa].

Pretendíamos comprobar si había algún factor común entre las muestras percibidas como [fa], y entre las muestras percibidas como [sa] (siguiendo el ej. anterior) que fuera distinto entre ambas percepciones.

## 5.2 Resultados

Presentamos a continuación los resultados de la revisión efectuada en la frecuencia media de los formantes, de sus anchos de banda y de la AM.

[x + V] con AM de [f + V]		
	Mayoría respuestas [x]	Mayoría respuestas [f] <sup>3</sup>
B1	560	493/508
B3	518	424/442
B4	279	513/373
B5	614	667/442
AM	5	4/4

[x + V] con AM de [θ + V]		
	Mayoría respuestas [x]	Distribución de respuestas
B3	800	424/673/442
B4	663	513/151/373
AM	3	4/3/4

<sup>3</sup> Se lee: En el grupo cuya manipulación consistió en colocar la AM de las muestras [f + V] o las muestras [x + V] revisamos, en este caso, 3 estímulos: uno en el que la mayoría de respuestas fue [x] y dos en los que hubo mayoría de respuestas [f].

[s + V] con AM de [f + V]		
	Mayoría respuestas [s]	Distribución de respuestas
F2	2144	1880/1946
B1	760	534/546
B4	268	417/365
AM	4	5/5

[s + V] con AM de [θ + V]		
	Mayoría respuestas [θ]	Mayoría respuestas [s]
F2	1868	1950/1946
F3	2678	2810/2995
B2	830	655/559
B3	356	572/444
B5	600	415/331
AM	3	4/6

[f + V] con AM de [x + V]		
	Mayoría respuestas [f]	Mayoría respuestas [x]
F1	303	966
B3	410	215
B5	641	772
AM	5	6

[f + V] con AM de [s + V]		
	Mayoría respuestas [f]	Distribución de respuestas
F1	892	603
B2	252	402
B5	800	406
AM	3	6

[θ + V] con AM de [x + V]		
	Mayoría respuestas [θ]	Mayoría respuestas [x]
F2	1324	1796/1904
F3	1903	2566/2719
F4	2766	3529/3665
F5	3900	4600/4600
B1	303	426/565
B3	3000	497/190
B5	335	412/498
AM	3	6/6

En el apartado [θ + V] con AM de [s + V] no pudimos revisar ningún estímulo puesto que en ninguna de las muestras en las que efectuamos la manipulación se dio ninguna respuesta [s], con lo cual no podíamos contrastar datos del mismo modo que lo habíamos hecho anteriormente.

A la vista de los resultados obtenidos en este análisis parece que podría ser relevante centrar la atención en los anchos de banda de los picos espectrales, probablemente en unión con los datos de intensidad,

para hallar los índices invariantes que distinguieran los puntos de articulación de las fricativas.

## 6. CONCLUSIÓN

Con este trabajo hemos pretendido comprobar si la amplitud podría constituir un factor decisivo en la percepción del punto de articulación de las consonantes fricativas. Llevamos a cabo dos experimentos que consistían en modificar las amplitudes de los grupos *fricativa + vocal* intercambiando las intensidades de [f] y [θ], por un lado, y de [s] y [x], por otro, bien en relación una a una, bien en relación a una media de un grupo de ellas. La presentación del test de percepción a los informantes se llevó a cabo a través de estas sílabas manipuladas y a través de los ruidos de fricación aislados. Los resultados, similares a los hallados por Behrens y Blumstein (1988a), demostraron que la amplitud por sí sola no puede ser considerada factor distintivo y decisivo en la percepción y reconocimiento de las consonantes fricativas sordas. Sin embargo, la revisión de otros parámetros como los anchos de banda de los picos espectrales en consonancia con la amplitud, especialmente de los picos de mayor frecuencia, nos permiten intuir que realmente los índices invariantes deben apuntar en esa dirección.

## 7. REFERENCIAS BIBLIOGRÁFICAS

- BEHRENS,S. y BLUMSTEIN,S.E. (1988a) «On the role of the amplitude of the fricative noise in the perception of place of articulation in voiceless fricative consonants», *Journal of the Acoustical Society of America*, 84, pp.861-867.
- BEHRENS,S. y BLUMSTEIN,S.E. (1988b) «Acoustic characteristics of English voiceless fricatives: A descriptive analysis», *Journal of Phonetics*, 16, pp.295-298.
- GUERLEKIAN,J.A. (1981) «Recognition of the Spanish fricatives /s/ and /θ/», *Journal of the Acoustical Society of America*, 70, pp.1624-1627.
- HEINZ,J.M. y STEVENS,K.N. (1961) «On the properties of voiceless fricative consonants», *Journal of the Acoustical Society of America*,33, pp.589-596.
- MARTÍNEZ CELDRÁN,E. (1984), *Fonética*, Teide, Barcelona.

- McCASLAND,G.P. (1979a) «Noise intensity and spectrum cues for spoken fricatives», *Journal of the Acoustical Society of America*, Suppl.1, 65, pp.78-79.
- McCASLAND,G.P. (1979b) «Noise intensity cues of spoken fricatives», *Journal of the Acoustical Society of America*, Supp.1, 66, pp.88.
- QUILIS, A (1981): *Fonética acústica de la lengua española*, Madrid, Gredos.
- STEVENS,P. (1960) «Spectra of fricative noise in human speech», *Language and Speech*, 3, pp.32-49.