

[r-ɾ]: ¿DOS CLASES DE SONIDOS?

EUGENIO MARTÍNEZ CELDRÁN Y LUCRECIA RALLO
Laboratori de Fonètica, Facultat de Filologia
Universitat de Barcelona

RESUMEN

La presentación en el IPA de los sonidos [ɾ] y [r] en dos modos de articulación diferentes llamados "tap" y "trill", respectivamente, contrasta con el hecho evidente de un comportamiento homogéneo desde un punto de vista fonológico, que lleva a creer en un parentesco fonético; es decir, creemos que no poseen dos modos diferenciados y que constituyen una sola clase. Existen argumentos encontrados entre algunas posiciones articulatorias contra la idea de ese alto parentesco fonético entre ellos (Catford, 1977) y posiciones acústicas que defienden lo contrario (Lindau, 1985). En nuestro estudio, hemos encontrado evidencias a favor de la posición acústica al basarnos en tests perceptivos, después de haber manipulado [ɾ] para convertirla en [r] y viceversa. Un 94.37% de las respuestas aceptan como buenos los sonidos surgidos de las manipulaciones. También hemos comprobado que estos sonidos consisten fundamentalmente en breves oclusiones de 20,08 ms y 15,78 ms respectivamente y que implican una bajada de amplitud de 9 y 7 Db, respectivamente.

ABSTRACT

The IPA classification of [ɾ] and [r] in two different manners of articulation, known as "tap" and "trill", respectively, contrasts with the evidence that the phonological behaviour of these two sounds is homogeneous. This leads us to believe that there exists a phonetical kinship between these two sounds; that is, they are two different varieties that are part of a sole class. Among certain articulatory views, we find arguments which reject the idea of this close phonetic kinship between [ɾ] and [r] (Catford, 1977). However, there are other acoustic standpoints which defend just the opposite idea. In our study we have found evidence in favour of the acoustic view. We relied on the results of various perceptive tests, which we obtained from editing the [ɾ] of a series of words until it sounded [r]. 94.37% of the subjects that listened to the tests accepted as good the sounds that resulted from the editing process. We also proved that these sounds consist basically of short-duration occlusions (20,08 and 15,76 m.s., respectively) with a decrease in amplitude of 9 and 7dB respectively.

1. PRESENTACIÓN

1.1. El problema

El Diccionario de la RAE (1992) define el término **vibrante** del siguiente modo: "dícese del sonido o letra cuya pronunciación se caracteriza por un rápido contacto oclusivo, simple o múltiple, entre los órganos de la articulación. La *r* de *hora* es **vibrante** simple y la de *honra* **vibrante** múltiple". El diccionario no hace sino recoger una larga tradición fonética, no sólo española, que toma esos dos sonidos como pertenecientes a una misma clase. Está claro que la fonética trata con sonidos, no con letras; aunque el diccionario se refiera también a la letra. Es más que probable que la base del concepto definido por el diccionario esté precisamente en la letra: una sola con sonidos diferentes según el contexto, como muestra por los ejemplos aducidos. Pero lo cierto es que la ortografía representa muy bien un hecho muy conocido en fonología: en las posiciones de pertinencia se marca con la escritura *r/rr*, como en *pero/perro*, la diferencia entre dos fonemas; mientras que en las posiciones de no pertinencia se escribe con una sola: *rema/honra/terca/tres*, etc. Por otra parte, la definición del diccionario es incompleta, aunque esté en lo cierto cuando dice que estos sonidos se caracterizan por un "rápido contacto oclusivo, simple o múltiple". Por tanto, es evidente que se considera que se trata de la misma clase de sonidos, lo cual se destaca en el diccionario al colocar en negrita el nombre de la clase: **vibrante** y se deja en letra normal la diferencia entre ambos sonidos: simple/múltiple. Hemos dicho que la definición no es completa, porque las vibrantes no sólo consisten en el rápido contacto oclusivo, sino que también se produce un elemento vocálico muy breve tras la oclusión, por regla general, de modo que podemos hablar de dos fases en una vibración: fase cerrada y fase abierta. En la vibrante simple intervocálica la fase abierta suele coincidir ya con la vocal siguiente, por lo que a veces no se puede segmentar nítidamente. Cuando es final de sílaba el elemento vocálico se manifiesta de forma clara como segunda fase, mientras que si la vibrante sigue a una oclusiva o fricativa formando sílaba con ellas, el elemento vocálico, denominado esvarabático, se produce delante de la oclusión (Cf. R. Cerdà, 1968). En la vibrante múltiple, hay un elemento vocálico tras cada oclusión, aunque el último no se suele apreciar al estar unido con la vocal.

Hace ya bastante tiempo que los cuadros del AFI (IPA: International Phonetic Alphabet, cf. 1949, 1989) recogen en modos de articulación diferentes lo que en inglés se denomina **trill** y **tap/flap**, lo cual significaría que son dos clases separadas de sonidos. Y no faltan autores

que defiendan esta postura basándose sobre todo en el gesto articulatorio.

1.2. La posición articulatoria

Catford (1977:130) es uno de los autores que mantienen una postura más firme contra la idea de que los gestos articulatorios que producen esos sonidos sean ni tan siquiera parecidos: "any idea that a trill is a 'rapid series of flaps', or that a flap is just an 'ultrashort trill' is quite wrong". Para justificar este aserto, proporciona el siguiente argumento: "the frequency of alveolar and uvular trills [r] and [R] is of the order of 30 cycles per second. This is much higher than the maximum rate at which one can produce a serie of [r]-flaps (about five or six per second). Once again it is clear that the flap is an essentially and obligatory *momentary gesture*; a trill is an essentially *prolongeable posture*". Otros autores han mantenido posiciones semejantes a la de Catford (cfr. P. Ladefoged, 1975). Incluso en la bibliografía sobre el español parece que se defiende esta postura; por ejemplo M.I. Massone (1988:26) dice: "-percusiva /vibrante simple- se trata de dos sonidos diferentes ya que implican gestos articulatorios diferentes". Parece que esta autora traduce el término *tap* por *percusiva*. Lo curioso es que admita la existencia de esos dos sonidos como diferentes. El mismo P. Ladefoged ha indicado la posibilidad de que existan *trills* con un solo contacto (1975:147). En principio: el *tap* tiene un solo contacto o movimiento, mientras que el *trill* posee más de uno, lo cual admite la mayoría de autores.

T. Navarro Tomás (1918) también destacó diferencias en la articulación de estos dos sonidos en la pronunciación española: "el movimiento de la lengua en [r] es de fuera a adentro, mientras que en [r], la punta de la lengua es empujada repetidamente de dentro a fuera", además añade que la tensión muscular es menor en la primera que en la segunda. No obstante, T. Navarro Tomás sigue considerándolas como dos sonidos de una misma clase: **vibrantes**, destacando que la [r] consta de una sola vibración, mientras que la [r] posee dos o más.

No todas las definiciones articulatorias de autores anglófonos separan de forma tajante esos dos tipos de sonidos. D. Crystal (1987:157) incluye ambos sonidos dentro de la clase que poseen "intermitent closure" y los define de la siguiente manera:

Roll or trill One articulator taps rapidly against another...

Flap A single tap is made by one articulator against another...

Como se ve, D. Crystal utiliza la palabra *tap* en ambos casos con funciones gramaticales diferentes, pero con el mismo significado; por lo que se puede deducir que sigue la idea de que se trata de una diferencia de grado, no de cualidad.

D. Abercrombie (1967:49) ya había indicado con anterioridad que "a trill in which the active articulator strikes only once against the passive articulator is called a *one-tap-trill*". Podemos considerar que la posición de Abercrombie es más parecida, incluso terminológicamente, a la mantenida por la fonética española distinguiendo entre vibrantes simple y múltiple. Esta distinción no es sólo española, la han mantenido también fonetistas de otros muchos sitios: B. Malmberg (1974), J. Morais Barbosa (1983), M. Schubiger (1970), L. Canepari (1977), etc. Por otra parte la postura de Abercrombie y Crystal es la misma que la del más antiguo D. Jones (1917) y los más modernos J. Clark y C. Yallop (1990)

1.3. La posición acústica

Frente a la postura de Catford con base articulatoria principalmente, cabe destacar una posición totalmente contraria basada en datos acústicos procedentes de análisis espectrográficos de varias lenguas que poseen ambos tipos de sonidos. M. Lindau (1985:166) indica que "the taps look very much like the closure phase of a trill. An average apical tap lasts 20 msec. This average was calculated from about 50 speakers of Swedish, Spanish, Housa, Degema, Edo, and Kalabari. Each closure phase of an apical trills also lasts about 20 msec, based on the average from twenty-five speakers of six languages. From an acoustic point of view, a trill can be regarded as series of taps. A tap is also frequently a variant of a trill, particularly in intervocalic position". Está clara la postura de Lindau: opuesta radicalmente a la mantenida por Catford. Lindau, aunque utiliza ambos nombres, los sonidos los incluye dentro de la misma clase bajo la denominación de *rhotics*.

1.4. La posición fonológica

Ya hemos comentado con anterioridad el fenómeno de la neutralización de ambos sonidos que se presenta en español. Unos de los primeros en indicarlo fue E. Alarcos (1950:183) "la oposición privativa bilateral entre los dos fonemas intermitentes o vibrantes /r/ tiene sólo pertinencia fonológica en posición intervocálica: pero/perro, caro/carro, [...] En todas las demás posiciones el número de las vibraciones (vibración simple/múltiple) no es relevante". Por tanto, el autor concluye

que en esas posiciones existirá el archifonema /R/. N. Trubetzkoy (1939:75) había indicado que "dos fonemas que participan en una oposición neutralizable están, aún en la posición de pertinencia, estrechamente emparentados entre sí". De esto se deduce el gran "parentesco" de las vibrantes. No es posible considerarlas como modos de articulación diferenciados. Son dos sonidos que poseen una misma base articulatoria y acústica y sus diferencias son lógicas por la desigualdad de tensión al pronunciarlos. Además, el hecho de la neutralización indica que es una oposición bilateral como decía E. Alarcos, lo cual implica que poseen una base de comparación única en el sistema. Si esa base no está formada por el rasgo **vibrante**, ¿qué rasgo la formaría?

Fonológicamente hay que decir que son parientes no sólo porque pueden neutralizarse, sino porque en muchas lenguas donde existe un solo fonema, estos sonidos pueden ser variantes alofónicas del mismo fonema; por ejemplo en francés (Landercy y Renard, 1977:92): vibrante /R/ "produit avec une ou plusieurs vibrations de la luelette". Si la simple no fuese vibrante sería una contradicción caracterizar el fonema /R/ con el rasgo vibrante. Incluso desde el punto de vista histórico hay intercambios de ambas consonantes sin una explicación lógica: *veruculu* > *berrojo*, *cerrojo*, por ejemplo. Y también existen contradicciones entre las distintas lenguas romances it. *aspárago* esp. *espárrago*; portugués *farrapo*, esp. *harapo*, etc. (Menéndez Pidal, 1904:135).

1.5 Hipótesis experimental

Parece evidente que todos los testimonios apoyan decididamente el hecho de que estos sonidos pertenecen a una sola clase natural, lo cual nos conduce a hipotetizar que a pesar de las diferencias articulatorias que puedan darse la base fonética es compartida ampliamente por ambos sonidos. Como el hablante busca ser entendido cuando habla, los gestos articulatorios están encaminados a producir las perturbaciones en las ondas sonoras que permitan al oyente extraer las pistas acústicas que le ayuden a interpretar el mensaje. Por tanto, el destino final de ese mensaje es el importante (Jakobson y Halle, 1956). Lo cual nos lleva a proponer un experimento de tipo psicoacústico para determinar si existe verdaderamente una diferencia crucial entre ambos sonidos que nos lleve a separarlos en clases diferentes. Si realmente es falsa la consideración de que la diferencia entre ambos es sólo de una vibración frente a varias acabaríamos concluyendo que son dos clases separadas de sonidos; de lo contrario, concluiremos que existe una sola clase: **vibrante**. Nos proponemos reducir la vibrante múltiple a una sola

vibración a ver si los oyentes escuchan [r] y viceversa: repetir dos o tres veces la única vibración de [r] para comprobar si los oyentes la identifican como [r]. Como dicen Lieberman y Blumstein (1988:153): "we never be certain that we have actually isolated the acoustic cues that people use to transmit information to each other unless we run psychoacoustic studies in which human listeners respond to acoustic signals that differ with respect to the acoustic cues that we think are relevant". Esperamos, pues, que el test psicoacústico nos ayude a determinar si la cantidad de las breves oclusiones (una frente a más de una) de los dos sonidos es la pista suficiente que los oyentes necesitan para reconocerlos.

2. MÉTODO

Un hablante masculino con un español estándar grabó una serie de palabras con vibrantes en todas las posiciones dentro de un frase portadora en una cabina insonorizada. Para ello se utilizó un cassette Marantz, modelo CP430, y un micrófono Shure SM58.

A continuación se procedió a la manipulación de las palabras para el test psicoacústico, mediante el CSL 4300B de Kay y las opciones del menú de edición (Edit: Trim Data y Append Displayed Source) que permiten eliminar, añadir o intercambiar trozos de señal. Por ejemplo, véase en las figuras siguientes cómo a partir del original *parra* se ha obtenido *para* al eliminar los dos elementos vocálicos y las dos oclusiones finales, de modo que nos hemos quedado con una sola oclusión.

De modo semejante, en la figura 2 se ve cómo de la palabra *cero* se ha obtenido *cerro* repitiendo dos veces la oclusión y los tres pulsos glotales iniciales de la vocal siguiente.

Cuando todos los estímulos estuvieron dispuestos, se mezclaron aleatoriamente al grabarlos de nuevo combinando los manipulados y los naturales. A continuación se pasó un test a 18 estudiantes de la Facultad de Filología.

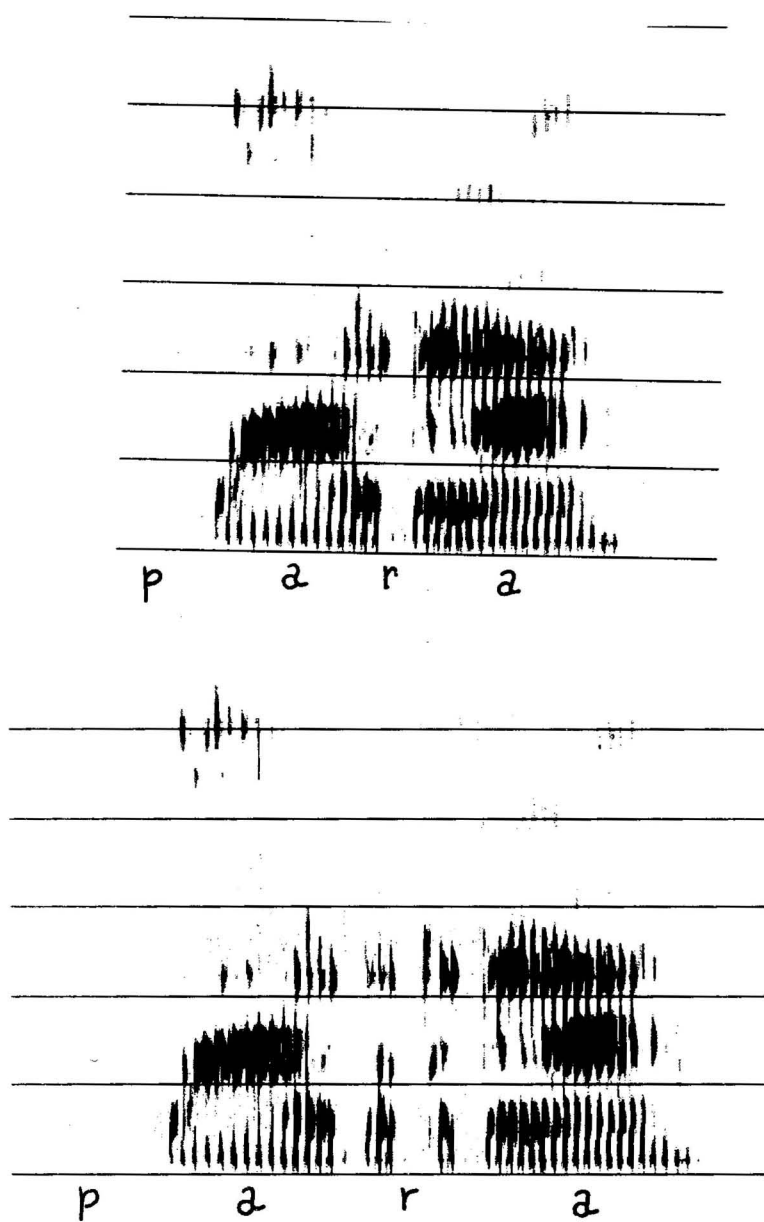


Fig. 1: Sonogramas de "para" natural y "parra" manipulado.

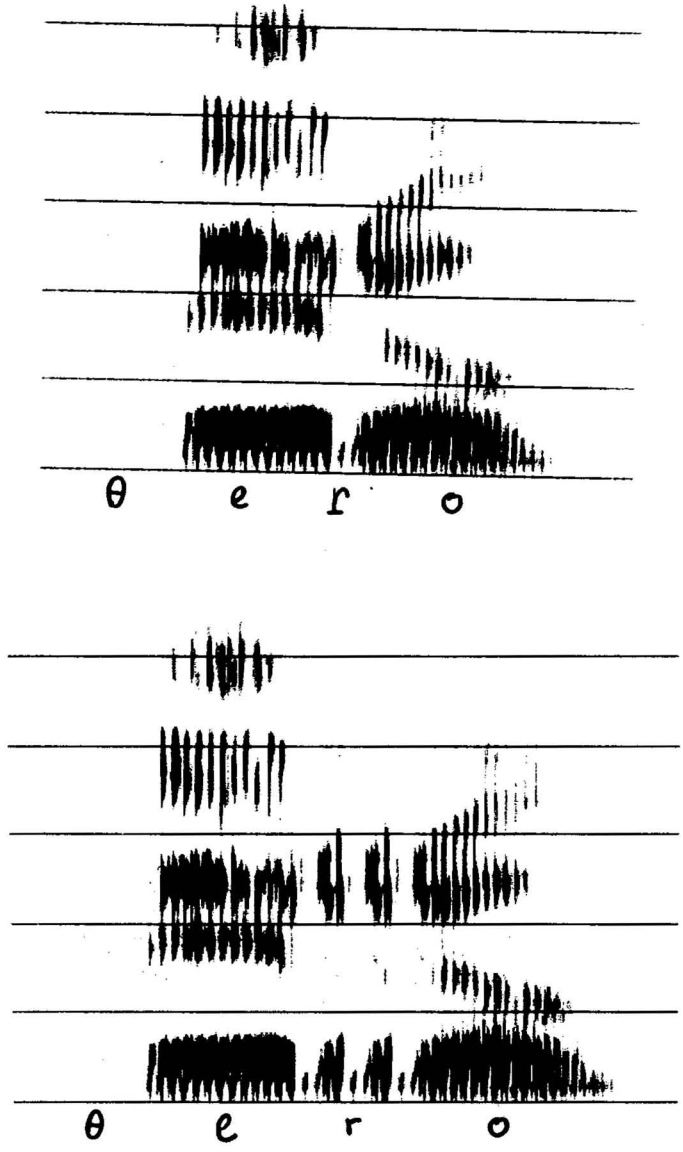


Fig. 2: Sonogramas de "cero" natural y "cerro" manipulado

3. RESULTADOS

3.1. A continuación expondremos los resultados de las parejas mínimas. En primer lugar la palabra natural y en segundo la manipulación resultante. Los datos indican el porcentaje de error en su reconocimiento:

<i>pero</i> -> <i>perro</i> : 5.5	<i>perro</i> -> <i>pero</i> : 0
<i>cero</i> -> <i>cerro</i> : 11.1	<i>cerro</i> -> <i>cero</i> : 5.8
<i>coro</i> -> <i>corra</i> : 0	<i>corro</i> -> <i>coro</i> : 0
<i>caro</i> -> <i>carro</i> : 16.6	<i>carro</i> -> <i>caro</i> : 0
<i>para</i> -> <i>parra</i> : 17.6	<i>parra</i> -> <i>para</i> : 0
<i>quería</i> -> <i>querría</i> : 0	<i>querría</i> -> <i>quería</i> : 11.1

Medias: r -> rr: 8.46; rr -> r: 2.8

Media total: 5.63 (es decir, 94.37% de respuestas a favor de la hipótesis)

3.2. Otra serie de manipulaciones de palabras estaban encaminadas a investigar perceptivamente si las vibrantes implosivas eran simples o múltiples.

3.2.1. En la primera serie de palabras con vibrante implosiva interior, se procedió a eliminar la consonante siguiente:

<i>parta</i> -> <i>para</i> : 0
<i>parla</i> -> <i>para</i> : 0
<i>corto</i> -> <i>coro</i> : 5.5
<i>Carlos</i> -> <i>caros</i> : 16.6

Media de errores en respuesta de vibrante simple: 5.5%

3.2.2. En la segunda serie con vibrante implosiva final (aunque no era final absoluto; por tanto es equivalente a la anterior). Se trata de añadir o eliminar elementos:

<i>intuir</i> -> <i>intuiremos</i> : 0	<i>intuiremos</i> -> <i>intuir</i> : 5.5
<i>lavar</i> -> <i>lavaremos</i> : 0	<i>lavaremos</i> -> <i>lavar</i> : 5.5
<i>temer</i> -> <i>temeremos</i> : 0	<i>temeremos</i> -> <i>temer</i> : 0
<i>amar</i> -> <i>amaremos</i> : 5.5	<i>amaremos</i> -> <i>amar</i> : 0

Media total de errores en lo esperado: 2%

3.3. Por último, también introdujimos unas pocas palabras para comprobar que la vibrante que va tras una oclusiva formando sílaba con ella es también una vibrante simple:

Teresa -> tres: 0	tres -> Teresa: 0
parado -> prado:29.4	prado -> parado:0
corona -> crónica: 0	

3.4. La duración de las vibrantes en posición intervocálica.

medias:

r: fase cerrada = 20.8 ms	r: fase cerrada = 15.78 ms
fase abierta = 27.66 ms	fase abierta = 24.07 ms

total: 48.46 ms	39.85 ms [por 2 = 79.7 y por 3 = 119.55]
-----------------	---------------------------------------------

Lo cual significa que en un segundo se podrían hacer 20 vibrantes simples. Esto es $1000/48.46 = 20$ o, expresado de otra forma, este sonido se puede realizar a 20 cps. Mientras que partiendo de sólo dos fases de la múltiple tendríamos una frecuencia de 25 cps ($1000/39.85 = 25$).

Las diferencias de duración no son significativas. Realizada la prueba de la t de Student para la fase cerrada se obtiene $t = 1.57$ con $p > 0.1$.

3.5. La intensidad de las vibrantes en posición intervocálica.

medias:

r: fase cerrada = -30 Db	r: fase cerrada = -32 Db
fase abierta = -21 Db	fase abierta = -25 Db

Diferencia:	9 Db	7 Db
-------------	------	------

Las fases abiertas son más intensas que las cerradas. Hay siempre una caída de la amplitud en la fase cerrada y un ascenso de 9 o 7 Db en la fase abierta. Las diferencias entre la amplitud de la simple y la múltiple no son tampoco significativas.

4. DISCUSIÓN

Los datos han validado nuestra hipótesis desde el punto de vista perceptivo, pues en el caso de las parejas mínimas está claro que una vibrante simple se puede convertir en múltiple repitiendo dos o tres

veces las dos partes que la componen y, viceversa: una múltiple se puede convertir en simple reduciéndola a una sola oclusión; lo cual da la razón a M. Lindau cuando decía que "a trill can be regarded as series of taps" y se la quita a Catford que se opone a la idea de Mondau. Creo que la percepción es, en último término, la que debe zanjar la disputa entre el fenómeno articulatorio y el acústico: no importa que los gestos articulatorios sean muy diferenciados si acústica y perceptivamente producen una realidad física y una sensación equivalentes.

Los datos acústicos no parecen dar la razón a Catford cuando indica que la [r] tendría unos 30 ciclos por segundo, mientras que [r] tendría cinco o seis. Como hemos dicho en el apartado 3.4., la vibrante simple española posee 20 cps, mientras que la vibrante múltiple posee 25. La diferencia de 5 cps no es significativa y no sé de donde saca Catford el dato de 5 o 6 cps. Nuestros datos, por otra parte, coinciden sensiblemente con los aportados por M. Lindau al indicar que, analizadas varias lenguas, se obtiene una media de 20 ms para la fase cerrada de la vibrante simple (Cf. De la Mota, 1991). La diferencia sigue estando en el hecho de existir una sola oclusión [articulatoriamente] o una fase de silencio [acústicamente] en la vibrante simple (*tap*) y más de una en la múltiple (*trill*), lo cual implica una caída de la intensidad seguida de un subida rápida.

Por otra parte, parece claro que son dos sonidos que pertenecen a la misma categoría y, por tanto, es interesante mantenerlos bajo la denominación tradicional de vibrantes. Alguna vez se ha argumentado contra el mantenimiento de ambos sonidos en la clase de las vibrantes diciendo que la simple no es una verdadera vibrante pues la vibración implica más de una oclusión. Los diccionarios suelen decir que *vibrante* se aplica a lo que vibra y *vibrar* se define como "moverse un cuerpo o una parte de él entre dos posiciones..." (M^a Moliner). No dice que ese movimiento tenga que ser repetido, aunque se sobreentienda. La misma autora define *vibración* como cada movimiento completo de vaivén de un cuerpo que vibra. En la medida en que la vibrante simple es un solo movimiento de los dos o tres que posee la múltiple, parece evidente que se pueda incluir en esta denominación. Por otra parte, la palabra *conjunto* implica una "reunión de varios" y, sin embargo, los matemáticos pueden hablar de conjunto unitario y vacío. Definiendo los elementos de forma clara y rigurosa, no vemos por qué no llamar vibrante a la simple.

5. CONCLUSIONES

Desde el punto de vista de la invariación está claro que la breve oclusión o, si se prefiere, una disminución brusca seguida de un aumento de la intensidad por un breve período o períodos repetidos es lo que caracteriza a esta clase de sonidos y, además, se demuestra por los tests perceptivos que se trata de una misma clase.

Por último, hay que señalar que los datos confirman la hipótesis que indica que la vibrante implosiva es simple. También lo es la que aparece tras la oclusiva formando sílaba con ella.

6. REFERENCIAS

- ABERCROMBIE, D. (1967): *Elements of general phonetics*, Edimburgo, Edinburgh University Press.
- ALARCOS, E. (1950): *Fonología española*, Madrid, Gredos, [1983].
- CANEPARI, L. (1979): *Introduzione alla fonetica*, Turín, Giulio Einaudi.
- CATFORD, C. J. (1977): *Fundamental problems in phonetics*, Edimburgo, Edinburgh University Press.
- CERDÀ, R. (1968): "Algunas observaciones en torno a la definición de "R" en español", *Boletín de Filología Española*, vol VII, nº 26-27, 19-24.
- CLARK, J. Y YALLOP, C. (1990): *An Introduction to Phonetics & Phonology*, Oxford, Basil Blackwell.
- CRYSTAL, D. (1987): *The Cambridge encyclopedia of language*, Cambridge, Cambridge University Press.
- IPA, (1989): "The IPA 1989 Kiel Convention", *Journal of the International Phonetic Association*, 19:2, 81-82.

- JAKOBSON, R. Y HALLE, M. (1956): *Fundamentos del lenguaje*, Madrid, Ayuso, 1980.
- JONES, D. (1918): *An Outline of English Phonetics*, Cambridge, Cambridge University Press.
- LADEFOGED, P. (1975): *A Course in Phonetics*, Nueva York, Harcourt Brace Jovanovich.
- LANDERCY, A. Y RENARD, R. (1977): *Éléments de phonétique*, Bruselas, Didier.
- LIEBERMAN, P. Y BLUMSTEIN, S. E. (1988): *Speech physiology, speech perception, and acoustic phonetics*, Cambridge, Cambridge University Press.
- LINDAU, M. (1985): "The story of /r/", en V.A. Fromkin (Ed.), *Phonetic linguistics: essays in honor of Peter Ladefoged*, Orlando, Academic Press, 157-168.
- MALMBERG, B. (1974): *Manuel de phonétique générale*, Paris, Éditions Picard.
- MASSONE, M. I. (1988): "Estudio acústico y perceptivo de las consonantes nasales y líquidas del Español", en *Estudios de Fonética Experimental*, vol. III, Barcelona, PPU, 13-34.
- MENÉNDEZ PIDAL, R. (1904): *Manual de gramática histórica española*, Madrid, Espasa Calpe, (1977).
- MOLINER, M. (1977): *Diccionario de uso del español*, Madrid, Gredos.
- Morais Barbosa, J. (1983): *Etudes de phonologie portugaise*, Universidade de Évora.
- MOTA GORRIZ, C. DE LA (1991): "A study of [r] and [ɾ] in spontaneous speech", *Proceedings of the XIIIth International Congress of phonetics Sciences*, Vol. IV, Aix-en-Provence, 386-389.
- NAVARRO TOMÁS, T. (1918): *Manual de pronunciación española*, Madrid, CSIC, (1971).
- RAE, (1992): *Diccionario de la lengua española*, Madrid, Espasa Calpe.

SCHUBIGER, M. (1970): *Introducción a la fonética*, Universidad de Valladolid (1989).

TRUBETZKOY, N. S. (1939): *Principios de fonología*, Madrid, Cincel, 1973.