

Ariel Practicum

Eugenio Martínez Celdrán

Análisis
espectrográfico
de los sonidos
del habla

Editorial Ariel, S.A.
Barcelona

CAPÍTULO 9 PROSODIA

9.1. Análisis del acento

No parece que el acento tenga un único parámetro acústico al que poder referirse, aunque son tres los que intervienen en la sílaba tónica con diferente grado de importancia. Hay autores que destacan la elevación del F0, otros la duración, aunque es posible que sea una combinación de ambos la que proporciona la prominencia de la sílaba acentuada frente a sus vecinas inacentuadas. La amplitud no parece que juegue un papel demasiado relevante en esta prominencia.

Las lenguas que utilizan el acento fonológicamente suelen distinguir entre tres clases de esquemas acentuales: esdrújulo, llano (o grave) y agudo, según que la sílaba prominente sea la antepenúltima, penúltima o última sílaba de la palabra. Las sílabas inacentuadas se caracterizan por simple oposición con las acentuadas; es decir, tienen los mismos parámetros pero sus valores son menores que en las acentuadas correspondientes. Se forma, pues, una gradación de mayor a menor: 1, 2, 3. Por ejemplo, una esdrújula suele tener el siguiente esquema: 1-3-2, por regla general: o sea, la sílaba más prominente es la antepenúltima y, por ello, lleva el número 1. La posición postónica es bastante débil; por tanto, le asignamos el valor 3. En la evolución de la lengua es una posición que suele cambiar e incluso se pierde la vocal: *rapidu>rabidu>rabdo>raudo* (¹p>²b; ²i>³Ø; ³b>⁴u).

El esquema grave suele tener la siguiente caracterización: 3-1-2/3. En este caso, la prominencia se sitúa en la penúltima sílaba. Las posiciones pre y postónicas suelen ser débiles.

Por último, el esquema agudo se presenta de la siguiente forma: 2-3-1. Lo cual demuestra que la posición más próxima al acento principal suele ser siempre la más débil. Véase en la figura 9.1 la actuación de los

tres parámetros que configuran los tres esquemas. Se ha pronunciado una secuencia sin sentido y con oclusivas para tener perfectamente delimitada la vocal, que es la portadora del acento: *tátata-tatáta-tatata*.

Valores encontrados:

<i>tá</i>	<i>ta</i>	<i>ta</i>	<i>ta</i>	<i>tá</i>	<i>ta</i>	<i>ta</i>	<i>ta</i>	<i>ta</i>	<i>tá</i>	
75	73	72	69	74	74	70	71	75		dB
159	149	156	112	156	154	116	111	156		Hz
90	61	68	63	78	58	66	61	85		ms

Como se ve, coinciden los valores máximos en la sílaba tónica. No obstante, las diferencias son mínimas tanto en las amplitudes (energy), como en el F0 (pitch); sólo el esquema agudo mantiene bastante diferencia entre las dos átonas y su tónica correspondiente. La duración es la que mejor pone de manifiesto los esquemas expuestos con anterioridad.

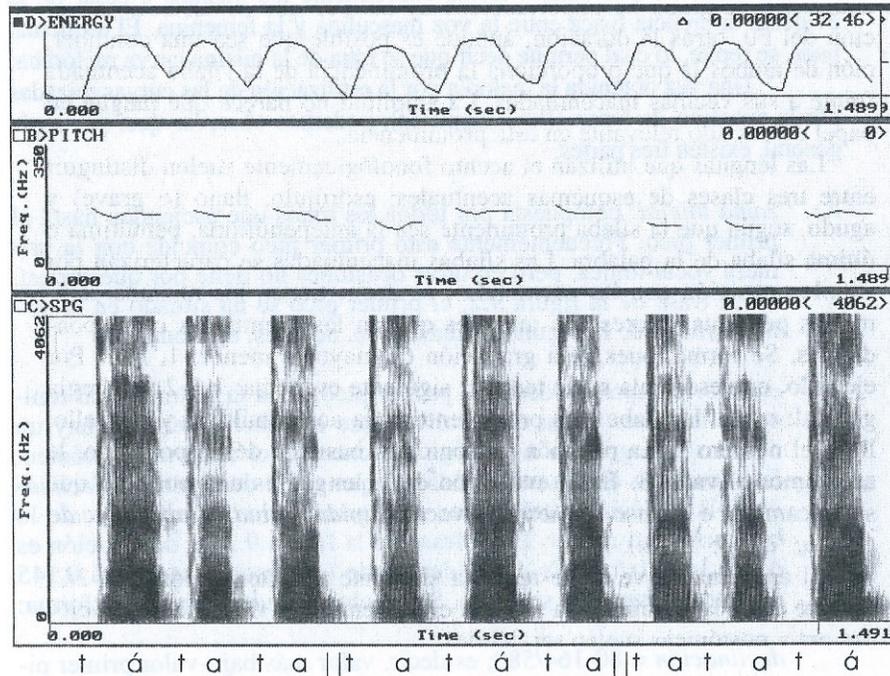


FIG. 9.1. Amplitudes, F0 y espectrograma de la secuencia *tatata* con acento esdrújulo, llano y agudo respectivamente.

Por último, hay que decir que las palabras átonas se integran en los esquemas acentuales de las tónicas; por ej. *Cómelo* = 1-3-2; *la mesa* = 3-1-3; *por amor* = 2-3-1; etc.

9.2. La melodía

La melodía es la contrapartida fonética de la entonación, concepto que debe relegarse únicamente a la fonología, para no confundir los hechos. La melodía consiste acústicamente en la curva que describe el primer armónico o tono fundamental (F0) en el grupo.

Un grupo melódico es el conjunto de tonos que se suceden entre dos pausas. Suele coincidir con una unidad sintáctica: oración, proposición, vocativo, etc. Para el estudio de los contornos melódicos se suele tomar como unidad un grupo fónico.

Existen diversos métodos para la eliminación de la variabilidad individual y toda aquella que no sea relevante. El procedimiento se conoce con el nombre de *estilización* o *normalización* de las curvas (véase Garrido, 1991:20). Por ejemplo, una forma de normalizar dos curvas de la misma frase pronunciadas por un hablante masculino y otro femenino podría ser la siguiente (cfr. Cantero, 1995:411): se toma nota del F0 de las vocales solamente. Una vez analizadas, tomamos el primer valor como referencia; es decir, ese primer valor será el valor cero de un porcentaje; el 100 % será la octava de ese valor (se obtiene multiplicando por dos). Los cálculos posteriores se harán realizando una sencilla regla de tres. Obsérvese lo siguiente: si tenemos dos tonos pronunciados por un hombre cuyos valores son 100 y 150 Hz respectivamente diremos que el segundo es superior al primero en un 50 %, pues $150-100$ (100 representa aquí el valor de la octava) = 50; $50 \cdot 100$ (100 es aquí el valor porcentual) = 5.000 y $5.000/100$ (100 representa aquí de nuevo la octava) = 50; si esos dos tonos son de una mujer, de modo que el primero mida 200 y el segundo 250, ahora la elevación no es del 50 % sino del 25 %, a pesar de que físicamente exista la misma diferencia en hercios; pues $250-200 = 50$; $50 \cdot 100 = 5.000$ y $5.000/200 = 25$. Intuitivamente se ve claro si comparamos las octavas:

- A) Ej. de un hombre:

Hz	100	125	150	175	200:	150 es la mitad de la octava
%	0	25	50	75	100	
- B) Ej. de una mujer:

Hz	200	250	300	350	400:	250 es la cuarta parte de la octava
%	0	25	50	75	100	

Es decir, para que la elevación de la voz femenina fuera equivalente a la masculina tendría que haber subido su fundamental hasta 300 Hz. Esto está en perfecta consonancia con el modo logarítmico con que los humanos percibimos las frecuencias.

Muchos autores tradicionales, como T. Navarro Tomás (1944), utilizaron los semitonos musicales para medir las diferencias melódicas dentro de una octava lo cual se asemeja al sistema musical, pero el sistema de porcentajes establece una medición sin relación con la música y fácil de manejar por todos los que no poseen conocimientos musicales. Ésa es su principal ventaja.

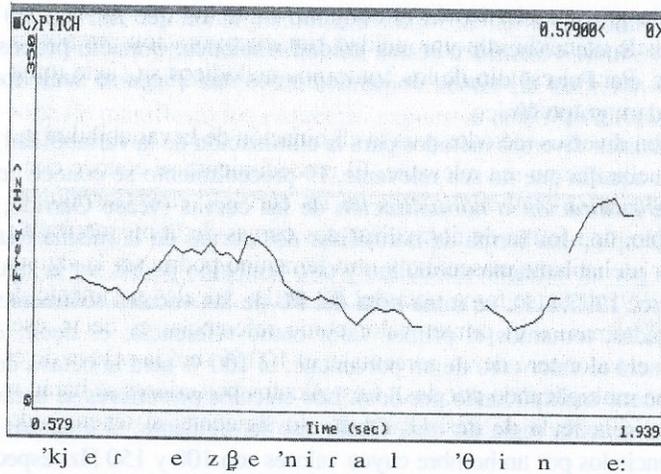


FIG. 9.2. Curva melódica de la frase.

Estilización de la curva:

	<i>kjé-</i>	<i>rez</i>	<i>βe</i>	<i>nír</i>	<i>al</i>	<i>θ</i>	<i>ne</i>
Hz	112	128	164	109	96	80	201
%	0	14	46	-2.6	-14	-28.5	79.46

La estilización¹ nos permite caracterizar en porcentajes algunos hechos, como los siguientes: desde el inicio hasta el primer pico (llamamos

1: Exponemos a continuación un par de ejemplos del cálculo de la estilización: 128-112 = 16; 16*100 = 1.600; 1.600/112 = 14; otro: 164-112 = 52; 52*100 = 5.200; 5.200/112 = 46; etc.

primer pico al punto más alto a partir del inicio) hay una subida del 46 %; desde el primer pico hasta el punto más bajo de la última sílaba acentuada la curva ha ido descendiendo un 74.5 % y, por último, la subida final ha sido de un 107.96 %.

La estilización posee otra ventaja: estandariza las diferencias individuales y, por tanto, las anula. Si una mujer hubiera pronunciado la misma frase, pero con valores de F0 que fuesen una octava más altos; es decir, que fuesen el doble de los pronunciados por la voz masculina tendríamos lo siguiente.

	<i>kjé-</i>	<i>rez</i>	<i>βe</i>	<i>nír</i>	<i>al</i>	<i>θ</i>	<i>ne</i>
Hz	224	256	328	218	192	160	402
%	0	14	46	-2.6	-14	-28.5	79.46

Es decir, los porcentajes serían exactamente los mismos a pesar de la evidente diferencia física entre la voz masculina y la femenina. El esquema, pues, se repite, lo cual permite decir que se trata de la misma curva melódica.

Una vez obtenida la melodía por la estilización de las curvas trazadas por la sucesión de tonos (fundamentales), podemos observar que, por regla general, existen tres partes.

- A) *rama inicial*: compuesta por todos los tonos que ascienden hasta el primer pico. Frecuentemente este primer pico coincide con la primera vocal tónica, pero en otras ocasiones no tiene por qué ser así. En la frase de la figura 9.2, el primer pico se ha situado en la primera sílaba de la segunda palabra que, además, es átona.
- B) *cuerpo*: comprende desde el primer pico hasta la última vocal tónica de la frase. Se ha comprobado que en muchas lenguas hay un descenso progresivo a lo largo del cuerpo del contorno melódico. Ese descenso dibuja una pendiente que se ha denominado *declinación*; la pendiente será mayor si el primer pico es más alto; de lo contrario será menor. En la frase de la figura 9.2, la declinación es de -0,145 Hz/ms. Es decir, desciende (por eso es negativa) 0,145 Hz por milésima de segundo. Se ha obtenido de la siguiente forma:

declinación = 80-164/580; es decir, valor más bajo-valor primer pico, partido por el tiempo existente entre esos dos puntos, expresado en milésimas de segundo.

En la figura 9.4 hemos presentado una voz femenina que pronuncia las cuatro palabras anteriores. Un análisis del F0 nos indica que 5-5 equivale a 250-270 Hz; 3-5 a 205-250 Hz; 2-1-4- a 185-148-205; y 5-1 a 270-140. Para este informante, pues, el 5 está alrededor de 250 o más; 3 y 4, alrededor de 200; 2 alrededor de 175 y 1, por debajo de 150.

La segunda posibilidad la vamos a ejemplificar con el ibidio (lengua nigeriana):

	Tono alto seguido de	Tono bajo seguido de
alto	ákpá "extensión del océano" (1)	àkpá "primero" (2)
descendente	ákpâ "cesto cuadrado" (3)	àkpô "árbol del caucho" (4)
bajo	ákù "sacerdote" (5)	àkpà "pequeño" (6)

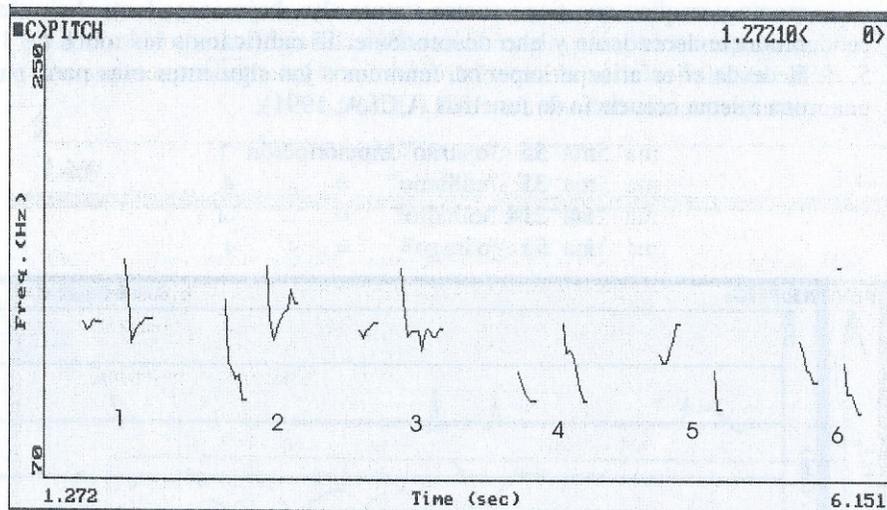
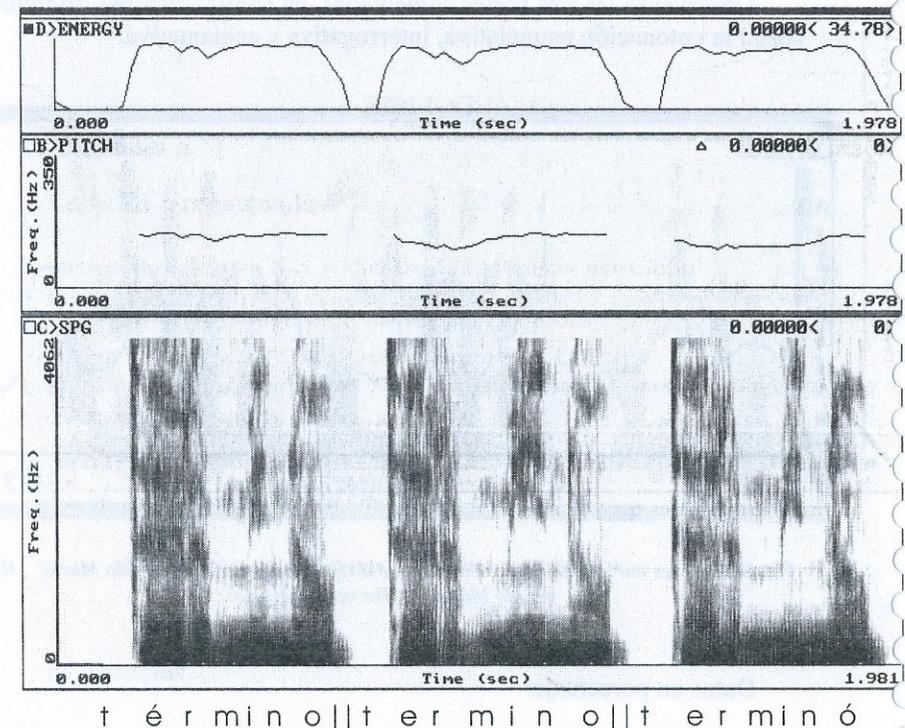


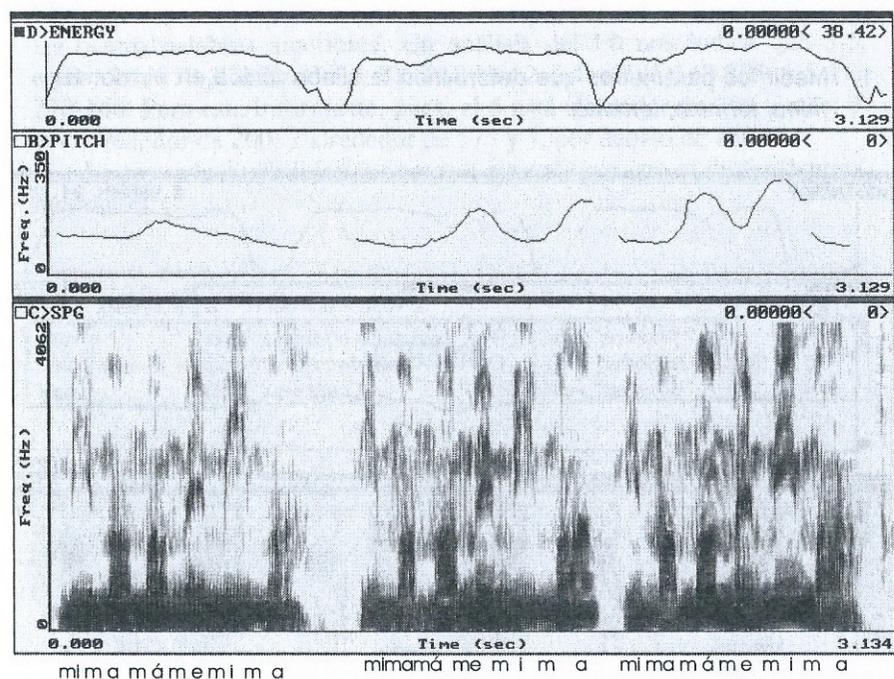
FIG. 9.5. Estructura tonal de las seis palabras del ibidio.

Ejercicios

1. Medir los parámetros que determinan la sílaba tónica en el trío: *término, termino, terminó*.



2. Estandarícense y díganse en porcentaje las diferencias entre las curvas enunciativa, interrogativa y exclamativa de la frase "Mi mamá me mima".
3. Averígüese la pendiente de la enunciativa y de la interrogativa.
4. Determinénse las tres ramas en cada curva y las diferencias que existen entre las tres modalidades basándose en las diferencias de las tres ramas.



BIBLIOGRAFÍA

Lecturas recomendadas

- BORDEN, G. J.; HARRIS, K. S. y RAPHAEL, L. J. (1980) (en referencias).
- BORZONE DE MANRIQUE, A. M. (1980): *Manual de fonética acústica*, Buenos Aires, Hachette.
- FRY, D. B. (1979): *The Physics of Speech*, Cambridge, CUP.
- KENT, R. D. y READ, CH. (1992): *The Acoustic Analysis of Speech*, San Diego, Singular Publishing Group.
- *LADEFOGED, P. (1962) (en referencias).
- *MARTÍNEZ CELDRÁN, E. (1996): *El sonido en la comunicación humana*, Barcelona, Octaedro.

(Los señalados con un asterisco representan lecturas más básicas.)

Referencias

- BORDEN, G. J.; HARRIS, K. S. y RAPHAEL, L. J. (1980): *Speech Science Primer*, Baltimore, Williams & Wilkins, 3.ª ed., 1994.
- CANTERO, F. J. (1995): *Estructura de los modelos entonativos: interpretación fonológica del acento y la entonación en castellano*, tesis doctoral, Universidad de Barcelona.
- CASTAÑEDA, M.ª L. (1986): "El VOT de las oclusivas sordas y sonoras españolas", *Estudios de Fonética Experimental*, II, Barcelona, PPU, pp. 91-110.
- FISHER-JØRGENSEN, E. (1968): "Voicing, Tenseness, and Aspiration in Stop Consonants, with Special Reference to French and Danish", *Annual Report of the Institute of Phonetics of the University of Copenhagen*, 3, pp. 136-164.
- GARRIDO, J. M. (1991): *Modelización de patrones melódicos del español para la síntesis y el reconocimiento de habla*, Bellaterra, Universitat Autònoma de Barcelona.
- LADEFOGED, P. (1962): *Elements of Acoustics Phonetics*, Chicago, The University of Chicago Press (1996, 2.ª ed.).