

**ACTAS DEL V
CONGRESO ANDALUZ
DE LINGÜÍSTICA
GENERAL**

**Homenaje al profesor
José Andrés de Molina Redondo**

Tomo I



**GRANADA LINGVISTICA
SERIE COLLECTAE**

FONÉTICA Y FONOLOGÍA

<i>Lucien Tesnière y la fonología</i>	425
ANTONIO MANJÓN-CABEZA CRUZ	
<i>Ortografía y fonología en Raymond Queneau</i>	433
RODRIGO LÓPEZ CARRILLO	
<i>La aspiración de la «s» implosiva en español. Su influencia en el andaluz</i>	457
ELENA GONZÁLEZ-BLANCO GARCÍA	
<i>Consonantes oclusivas sonoras en inglés y en el español de Jaén, Córdoba y Granada: distinción acústica</i>	473
M. ^a DEL CARMEN RODRÍGUEZ DEL RÍO	
<i>La resolución del elemento sibilante en el andaluz</i>	489
PAUL O'NEILL	
<i>La percepción de la duración vocálica en español</i>	501
ANTONIO PAMIES BERTRÁN, ANA M. ^a FERNÁNDEZ PLANAS	

La percepción de la duración vocálica en español¹

ANTONIO PAMIES BERTRÁN
ANA M.^a FERNÁNDEZ PLANAS

Universidad de Granada y Universidad de Barcelona

1. INTRODUCCIÓN

El hecho de que el factor tiempo pudiera ser relevante en la configuración del contorno tonal de la oración ha sido muy debatido (cf. Toledo 2000) y lo mismo ocurre con respecto a su relación con el acento (Pamies 1996), pero el estudio de dicho factor requiere a su vez disponer de datos más seguros sobre el umbral a partir del cual una diferencia durativa es perceptible, entre otros factores: *si l'on fait abstraction des difficiles problèmes liés à l'évaluation précise des durées phonétiques, les valeurs objectivement mesurées doivent être corrigées en fonction du locuteur, de la vitesse locale d'élocution, du phonème et de son contexte, ainsi que des connaissances sur la perception de ce paramètre* (Méloni & Langlais 1986).

Por otra parte, se ha comprobado que la función que los suprasegmentales tienen en cada lengua influye a su vez sobre la capacidad discriminatoria de sus oyentes (Lehiste 1977, Eek 1978, 1987), por lo que dichos datos deberían establecerse en principio para una lengua en particular.

Se han hecho numerosos estudios sobre la duración en varias lenguas, especialmente en lo que se refiere a la producción, ya sea como realización de la cantidad vocálica en las lenguas donde ésta es relevante, ya sea como posible correlato del acento o de la sonoridad (cf. Durand 1946; Delattre 1966; Bertinetto 1977; Lehiste 1970a, 1976, 1977, 1988, 1997; Carlson & Granstrom 1975; Eck 1979, 1980; Noteboom & Doodeman 1980; Abramson 1993; Lehiste & Ross 1997; Krull, 1998; Pamies, Hagerty & Iñesta 1999; Krull & Traummüller 2000). Junto a los datos obtenidos sobre la ratio temporal entre vocales largas y breves, también se han realizado

¹ Este trabajo se ha beneficiado de una ayuda a proyectos de investigación de la DGSEIC/MEC (PB98-1230), y fue presentado como comunicación en el V Congreso Nacional de Lingüística General (León 2002).

algunos trabajos experimentales sobre la percepción, para hallar las diferencias mínimas perceptibles en materia de duración (*Just Noticeable Differences*), con resultados bastante variados (Delattre 1940, 1967, Denev 1955, Fry 1958, Creelman 1962, Noteboom 1972, 1973, 1979; Rossi 1972; Huggins 1972, Cooper 1975; Lehiste 1975, Fujisaki et al. 1975; Klatt & Cooper 1975, Vaissière 1977, Berinstein 1978; Bovet & Rossi 1978; Eek 1978, 1979; Solé 1984; Fox & Lehiste 1987, 1989; Eek & Meister 1997). Las JNDs obtenidas son muy dispares, y van desde el 3% que obtienen Rhum et al. (1966) con oyentes entrenados ($D=3\text{ms}$, $T=100\text{ms}$)², hasta el 27,4% de Rossi (1972), pasando por el 5,5% de Noteboom & Doodeman, el 7% de Bovet & Rossi (1979), el 20% de Creelman (1962) y de Henry (1948) ($D=16\text{ms}$, $T=77\text{ms}$), el $D=15,8\%$ de Abel (1972) ($D=10\text{ms}$, $T=63\text{ms}$) y de Huggins (1972); por no hablar del abanico obtenido por Eek (1978), que va de 3,7% a 35% dependiendo de la duración de referencia. Está claro que tal diversidad es debida en gran parte a la enorme diferencia metodológica entre estos trabajos.

Por citar tan sólo algunos ejemplos representativos de esta disparidad metodológica, Creelman (1962) realizó un test perceptivo con sonidos puros, a partir de unos estímulos base de distinta duración ($T=120\text{ms}$, $T=180\text{ms}$ y $T=265\text{ms}$) contrastados a unas variantes a las que añadía o restaba fragmentos (D) de 10ms (o múltiplos). Los oyentes (sobreetrenados) debían responder de forma binaria y discreta igual/desigual (*binary forced choice*). Obtuvo un umbral de $D=20\%$ para los estímulos de partida situados entre $T=100$ y $T=300\text{ms}$. Rossi (1972) realizó en cambio un test con la vocal [a] aislada (voz natural manipulada), presentado a 23 informantes franceses no entrenados. El umbral no resultó ser el mismo para todas las duraciones de referencia. Partiendo de unos estímulos de $T=120\text{ms}$, $T=180\text{ms}$ y $T=260\text{ms}$, las variables manipuladas añadían o restaban fragmentos de 20ms (o múltiplos) hasta cubrir todo el abanico de duraciones que se suelen dar en la producción. La pregunta admitía tres respuestas acerca del segundo estímulo: *más largo / más breve / igual*. Los umbrales obtenidos son $D=27,4\%$, $D=21,3\%$ y $D=24\%$ respectivamente para cada serie, por lo que no se cumpliría totalmente la ley de Weber, ya que ciertas duraciones de partida conllevan unas JNDs más altas que otras. En el análisis separado de los estímulos de referencia comprendidos entre 60ms y 130ms, el umbral estaría aproximadamente en $D=30\%$. También concluyó que el orden de la sucesión entre larga y breve influye en la percepción. Sometió el mismo test a un grupo de checos y eslovacos (lenguas con cantidad vocálica fonológica) y, contrariamente a lo esperable, no halló un umbral más bajo que el de los informantes franceses.

Huggins (1972) utilizó en cambio frases marco donde se insertaban las vocales estudiadas y propuso otro tipo de respuestas (normal vs. breve, normal vs. larga), hallando un umbral en torno a $D=15\%$. Klatt & Cooper (1975) investigaron tanto voca-

les como consonantes insertadas en palabras y con duraciones de referencia y resultados variados dependiendo del tipo de sonido. Bovet y Rossi (1978) compararon la percepción de los sonidos puros con la de las vocales. Los estímulos eran de $T=200\text{ms}$ y las variantes (D) añadían o restaban fragmentos de 20% o de 15% (y múltiplos). Se exigía al oyente una elección binaria entre breve/largo tras la audición de un sólo estímulo, sin comparación con otro. Había un número muy reducido de informantes y un alto número de pruebas. Se concluyó que el umbral estaría en torno a $D=7\%$, y que no había diferencia relevante entre sonidos puros y vocales en lo que se refiere a la percepción de su duración.

Eek (1978) estudió las JNDs durativas en la percepción de cantidad vocal en estonio, empleó sólo 8 informantes pero un alto número de estímulos realizado en un aula sin auriculares. Dos de los informantes eran rusos (lengua sin cantidad vocálica). La secuencia era la vocal [a] aislada, con una gradación por tramos de 2ms, añadidos a distintas duraciones de base (desde 40 a 320ms), y preguntaba a los informantes si la segunda [a] era más larga o más breve que la primera. Evaluando el umbral en un 75% de respuestas correctas, halló distintos valores de D en función de cada valor inicial de T. La ratio obtenida para los 6 informantes estonios difería de los dos rusos, ya que los segundos requirieron unas JNDs bastante mayores que las de los primeros, contrariamente a la comparación que hizo Rossi entre oyentes franceses y checos y eslovacos. Los datos de Eek coinciden claramente con los de Rossi en el hecho de que la proporcionalidad no es lineal: la ratio T/D es más alta en las secuencias breves, disminuye en las largas, y vuelve a aumentar ligeramente en las extralargas, tanto en secuencias crecientes (A) como decrecientes (B). Este cuadro refleja la importancia del valor de referencia sobre las JND, así como el distinto resultado obtenido por estonios y rusos:

(Eek 978)	oyentes estonios		oyentes rusos	
	A	B	A	B
T=40ms	D= 8ms (20%)	D= 8 ms (20%)	D= 14 ms (35%)	D= 12 ms (30%)
T=80ms	D= 12ms (15%)	D= 6 ms (7,5%)	D= 20 ms (25%)	D= 16 ms (20%)
T=120ms	D= 18ms (15%)	D= 8 ms (6,7%)	D= 16 ms (13,3%)	D= 18 ms (15%)
T=160ms	D= 12ms (7,5%)	D= 8 ms (5%)	D= 10 ms (6,2%)	D= 14 ms (8,7%)
T=200ms	D= 14ms (7%)	D= 12ms (6%)	D= 12 ms (6%)	D= 18 ms (9%)
T=240ms	D= 16ms (6,7%)	D= 12ms (5%)	D= 12 ms (7,5%)	D= 24 ms (10%)
T=280ms	D= 22ms (7,8%)	D= 20ms (7,1%)	D= 20 ms (7,1%)	D= 24 ms (8,5%)
T=320ms	D= 30ms (9,4%)	D= 20ms (6,2%)	D= 28 ms (8,7%)	D= 34 ms (10,6%)

Noteboom & Doodeman (1980) estudiaron la oposición fonológica neerlandesa entre /a/ y /a:/ en pares mínimos de palabras insertadas en una frase (5 posiciones), grabadas con voz natural posteriormente manipuladas para obtener duraciones de 66, 75, 84, 93, 102, 110 y 120ms. El test perceptivo fue sometido a 12 oyentes holandeses (10 no entrenados, más los propios autores), que debían distinguir los pares míni-

mos *tak* («rama») vs. *taak* («tarea»), insertados en 5 parejas de frases que escucharon individualmente con auriculares, al ritmo que ellos querían y midiéndose el tiempo de respuesta. Como en el interior de una frase no se puede aplicar el contraste directo entre los elementos a comparar, supusieron la existencia de un «criterio interno» similar al que propone por la *teoría de detección de señales*³ y aplicaron por tanto la fórmula ($d_0 = (2/\sqrt{2}) \cdot d_{10} = \sqrt{2} \cdot d_{10}$), hallando un umbral diferencial muy bajo ($D=5\text{ms}$ [5,5%] para $T=90\text{ms}$). Concluían que el umbral diferencial durativo en las frases sería incluso inferior al que se obtendría comparando dos sonidos consecutivos, y que se ajustaría al tipo de consonante que sigue inmediatamente a la vocal usada como variable. En un trabajo sobre la cantidad silábica en estonio, Krull & Traunmüller (2000) estudiaron por su parte el límite fonológico en la percepción de breves, largas y extralargas en estonio, demostrando que, en la cadena hablada, el punto de referencia para establecer la comparación entre largas y breves no sería la vocal de la sílaba vecina, sino el *tempo* de la frase, de forma que habría que normalizar las duraciones absolutas, p.ej., con la duración media por fonema de cada frase.

Todos estos resultados son no sólo divergentes, sino incomparables entre sí, dado que no coinciden ni en el método empleado ni en el tipo de material analizado. Varía el tipo de sonido utilizado (sonidos puros, vocales aisladas, vocales insertadas en frases, consonantes, vocales a las que se ha suprimido las transiciones iniciales), la forma de obtenerlo (voz natural, síntesis), el tipo de oyente (no entrenado, sobreentrenado, hablante nativo de una lengua con o sin cantidad vocálica fonológica), el número de oyentes empleados (generalmente demasiado bajo), la duración de referencia (a veces demasiado elevada), las condiciones de realización del test (audición individual en cabina anecoica con auriculares, audición colectiva en sala normal sin auriculares), sin contar el punto en el que se establece la medición así como el tipo de respuesta que el test permite (no es lo mismo decir si son distintos/iguales que decir además cuál es más largo).

2. DISEÑO EXPERIMENTAL

Nos hemos planteado un diseño experimental en el cual se procura:

- a) un máximo aislamiento de la variable tiempo, neutralizando los demás factores, gracias a la manipulación por resíntesis;
- b) unos estímulos de partida representativos de los valores más habituales en la producción por estar en la zona central (y más frecuente) del abanico de duraciones vocálicas que se dan en la producción (entre 100ms y 200ms);

³ The sensitivity (d') —expressed as the separation means of two stimulus strength distributions divided by the common standard deviation— is $\sqrt{2}$ smaller than the corresponding measure of the sensitivity in a forced choice two interval task, expressed as two times the separation of means of two stimulus strength distribution divided by the common standard deviation.

- c) unos informantes numerosos y sin ningún tipo de habilidad especial (lenguas sin cantidad vocálica, excluyendo los resultados de los que tienen formación musical);
- d) unas condiciones «naturales» de audición que no superen excesivamente las más habituales en la comunicación, para no favorecer una percepción artificialmente precisa que sería de escasa aplicabilidad;
- e) una fórmula sencilla para el cálculo del punto de inflexión, como es el cruce del nivel de error con el de aciertos;
- f) una pregunta que corresponda al objetivo, que es sólo la audibilidad de una diferencia durativa, sin que por ello se exija saber decir cuál es la vocal más larga.

Nuestra intuición inicial se basaba en un paralelismo con la música, ya que la escritura musical refleja explícitamente determinados umbrales durativos diferenciales, con un sistema preciso y eficaz de segmentación temporal basado en las proporciones y nunca en valores absolutos (cf. Rossi 1972:436, Eek 1980:27; Krull & Trau Müller 2000, etc). En la música popular más «rudimentaria», es decir, cuya apreciación y memorización están al alcance de personas sin ninguna aptitud especial ni formación musical alguna⁴, las notas mantienen diferencias durativas indiscutiblemente pertinentes: la más evidente es la oposición entre negra y blanca o entre corchea y negra (+100% si tomamos la más breve como valor de referencia). En géneros musicales que requieren cierto aprendizaje⁵, suelen añadirse contrastes durativos más sutiles, por ejemplo, la diferencia entre una corchea de tresillo y una corchea normal (+33% si tomamos la más breve como referencia⁶). Se trata de diferencias proporcionales, nunca absolutas. Dicho umbral es curiosamente superior a todos los mencionados en la literatura sobre percepción durativa en el habla, tan sólo el 27,4% señalado por Rossi (1972) se acerca a este valor.

3. EXPERIMENTO I

3.1. Planteamiento

Nuestro corpus de estudio se compone por tanto de unos estímulos-base de 100ms, y variantes con incrementos temporales de +10%, +20%, +25%, +33%, +40%, +50%, +66%, +75% y +100%. De esta forma se controlaban proporciones de $\frac{1}{10}$, $\frac{1}{5}$, $\frac{1}{4}$, $\frac{1}{3}$, $\frac{2}{5}$, $\frac{1}{2}$, $\frac{2}{3}$ y $\frac{3}{4}$. La secuencia con +40% no corresponde a una proporción musicalmente relevante pero sirve para no dejar sin medición intermedia el largo tramo que va desde $+\frac{1}{3}$ a $+\frac{1}{2}$.

⁴ P. ej. canciones infantiles, himnos, marchas militares.

⁵ P. ej. la llamada música «clásica» o el jazz.

⁶ La música toma la negra como valor de referencia, pero en el lenguaje, la breve es siempre el miembro no-marcado de la oposición durativa; Trubetsky habla incluso de *breve* vs. *no-breve*.

La secuencia [ma] se grabó con voz natural femenina, y se uniformizaron los valores decibélicos, tonales y formánticos, por re-síntesis mediante el programa ASL de KAY-Elementrics, con el que se obtuvieron las variantes manipuladas. Se duplicó la sílaba base y se alteró una de las sílabas, siempre restándole al material inicialmente grabado, obteniendo parejas de 200-100ms, 175-100ms, 166-100ms, 150-100ms, 140-100ms, 133-100ms, 125-100ms, 120-100ms y 110-100ms. El corpus del test es por tanto una serie de secuencias bislabas con una diferencia durativa entre sus dos vocales, desordenadas y mezcladas con otras secuencias en las que ambas vocales duraban lo mismo (100-100ms). Los oyentes (no entrenados) tenían que dar solamente una respuesta binaria discreta (iguales /desiguales). La audición se realizó de forma colectiva y sin auriculares. Los datos de cada informante fueron almacenados en una hoja de cálculo electrónica. Cada estímulo fue oído dos veces con un intervalo entre ambos de 1 segundo, y una pausa de 3 segundos entre estímulos diferentes. Los estímulos anisocrónicos fueron mezclados aleatoriamente con otros rigurosamente isocrónicos.

3.2. Resultados

Se hizo un experimento previo de calibración, los informantes fueron 50 estudiantes de la Facultad de Letras de la Universidad de Granada (grupo A), obteniéndose el siguiente resultado:

+ ms	fallos grupo A
+10%	36/50
+20%	29/50
+25%	22/50
+33%	17/50
+40%	18/50
+50%	16/50
+66%	10/50
+75%	9/50
+100%	8/50

El grupo A muestra un punto de inflexión del número de errores en torno a un incremento temporal de entre 20% y 25%, resultado bastante cercano a los de Creelman y el de Rossi. Pero como también hemos anotado los errores cometidos en los impulsos «de relleno», donde ambas sílabas eran iguales, nos sorprendió que se detectara un nivel medio de error en igualdad sospechosamente elevado (28%), lo cual hace pensar que algo podría estar equivocado en el propio experimento: ¿cómo es posible que los oyentes hayan sido tan sensibles para captar diferencias pero tan «torpes» como para creer oír diferencias donde no las había? Lo normal sería que la media de error en desigualdad superase ampliamente la media de error en igualdad. Un músico (no incluido en los resultados) realizó el test con 100% de éxito en las

desigualdades —captando incluso las de tan sólo +10%— y, sin embargo, también creyó que secuencia iguales eran distintas, con 100% de error en igualdad, lo que confirmaba que el estímulo empleado podía favorecer la opción «desigual» en las respuestas de los informantes.

El estímulo era la secuencia [mama], en la cual se había realizado un límite intersilábico lo más discreto posible, «cosiendo» dos cadenas [mam] por la [m] central que luego se acortaba en su parte central para simular una palabra bisílaba; el fallo podía residir entonces en que la primera [a] tenía por tanto las transiciones formánticas de salida propias de una [a] seguida de [m], mientras que la segunda, obviamente, no las tenía. Por lo tanto las dos vocales no eran totalmente iguales: el final suavizado de la primera [a] cuyas transiciones formánticas eulazaban perfectamente con la [m] siguiente contrastaban con el corte abrupto de la segunda vocal. Había que modificar el diseño de los estímulos para el experimento definitivo.

4. EXPERIMENTO 2

4.1. Planteamiento

Considerando que el primer test era poco fiable, el experimento definitivo exigió modificar el diseño inicial, usando ahora sucesiones de dos sílabas ([mam] [mam]) en vez de la secuencia [mama] empleada anteriormente, así ambas vocales tenían las mismas transiciones formánticas de salida. Ambas sílabas estaban separadas por un silencio brevísimo. El test fue sometido a un total de 100 oyentes no-entrenados, todos ellos españoles, estudiantes universitarios de Barcelona (B) y de Granada (C). Al procesar los resultados, se descartaron los de algunos alumnos, por ser éstos extranjeros o por tener una formación musical.

El grupo B lo forman 50 estudiantes de Filología de la Universidad de Barcelona; el grupo C lo conforman 50 estudiantes de la Facultad de Traducción e Interpretación de Granada.

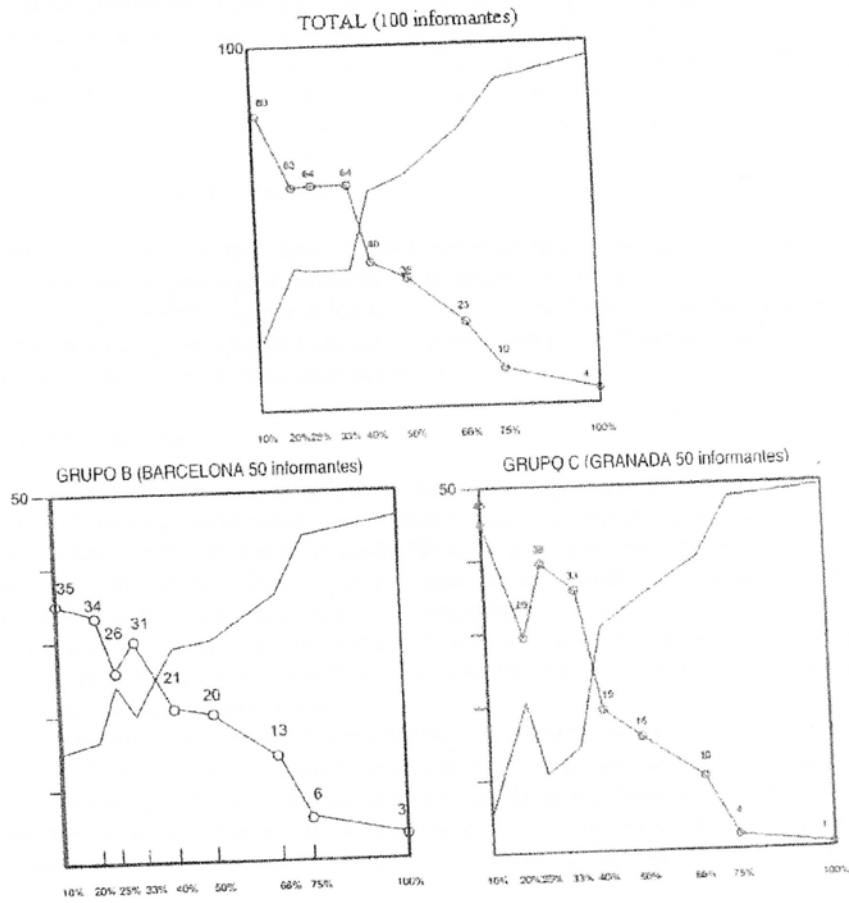
4.2. Resultados

En ambos casos, el nivel medio de error en igualdad bajó (20% y 14%) mientras aumentaba el error en desigualdad, como era previsible y deseable. La diferencia con respecto a la prueba de calibración nos confirmaba dos cosas: a) un pequeño cambio metodológico altera considerablemente los resultados; b) el nivel de error en igualdad, no utilizado por otros autores, es un factor que se debe evaluar porque puede informarnos acerca de la fiabilidad del propio experimento.

Al haber 50 oyentes en cada grupo, el total no requiere ni hacer la media ni convertir a porcentajes, por lo que no es indispensable —en principio— la ponderación estadística. La simple suma de todos los resultados individuales nos da este esquema donde datos absolutos y datos relativos son idénticos porque todos ellos están en una escala entre 0 y 100:

+ ms	fallos B	fallos grupo C	total fallos B+C
+10	35	45	80
+20	34	29	63
+25	26	38	64
+33	31	33	64
+40	21	19	40
+50	20	16	36
+66	13	10	23
+75	6	4	10
+100	3	1	4

Éstas serían las representaciones gráficas de dicha tabla numérica:



5. DISCUSIÓN

La interpretación de los datos obtenidos también es objeto de divergencias metodológicas que influyen de forma decisiva las posibles conclusiones, ya que la curva nos muestra una correlación entre los aciertos perceptivos y cada incremento durativo, pero no en qué punto de la misma se sitúa el umbral. Algunos autores eligieron el punto en que la curva de aciertos alcanza el 75%, decisión relativamente arbitraria que debería ser argumentada y que los lleva a un umbral muy distinto que si hubieran elegido otro punto, no menos peligroso.

Dentro de la prudencia a la que obliga tan arriesgada decisión, optamos por el punto en que los aciertos empiezan a superar los errores, que corresponde —por definición— al 50% de aciertos (cf. *phoneme boundary of length perception*, en Noteboom & Doodeman 1980). Este punto coincide además con el tramo de máxima pendiente de la curva, situado a su vez entre dos tramos de pendiente más moderada, lo cual es también un indicio de representatividad. Este punto está entre el 33% y el 40% de incremento durativo, tanto en el gráfico que representa el total de los hablantes, como en los que representan por separado al grupo barcelonés y granadino. Para calcular el valor exacto del umbral, aplicamos una interpolación lineal a partir de los dos valores conocidos a izquierda y derecha de la incógnita en la curva:

$$D = x_1 + [(x_2 - x_1) / (y_2 - y_1)] \cdot (y_3 - y_1)$$

siendo el valor de y_3 el 50% (cruce entre la curva de fallos y la curva de aciertos). El valor del umbral perceptivo diferencial resultante es un incremento temporal de $D=35,9\%$. Este valor, superior a los que se citan en la literatura, es muy cercano al umbral mínimo musicalmente relevante, representado por la diferencia entre una corchea de tresillo y una corchea normal (+33,333 %).

6. CONCLUSIONES

1) La metodología experimental es tan decisiva a la hora de abordar la percepción de la duración en el habla, que el umbral resultante puede ser ínfimo o considerable, dependiendo del diseño aplicado. No es lo mismo un estímulo base de 300ms que uno de 100ms, ni oírlo en un aula grande con un altavoz o mediante auriculares en una cabina anecoica. No es lo mismo preguntar si dos secuencias son iguales que preguntar cuál es más larga; tampoco es lo mismo evaluar los datos obtenidos para un 75% de aciertos que para un 50% o un 25%. De ahí que resultados muy dispares no se contradigan necesariamente.

2) Nuestro experimento pretende obtener un umbral representativo de la media. Por ello ha optado por condiciones de percepción que no sean mejores de lo normal, una duración de partida y unos incrementos similares a la horquilla durativa de las vocales tónicas en el habla real, un porcentaje de éxito que también sea conforme a la media, así como un número de oyentes lo suficientemente alto como representarla.

3) Partiendo de la metodología y presupuestos aquí explicados, identificamos un umbral perceptivo medio de las duraciones vocálicas de +36%, valor casi idéntico al de la mínima diferencia temporal relevante en música (+33%).

7. REFERENCIAS BIBLIOGRÁFICAS

- ABEL, S.M. (1972): «Duration discrimination of noise and tone bursts», *Journal of the Acoustical Society of America*, 51, 1219-1223.
- ABRAMSON, A.S. (1993): «The stability of distinctive vowel length in Thai», in *Haskins Laboratories Report on Speech Research*, Sr 114; 91-100.
- BERINSTEIN, A. (1978): «A cross-linguistic study on the perception of stress», *Journal of the Acoustical Society of America*, 63, 55-56.
- BERTINETTO, P.M. (1977): «A contrastive study on the production and perception of stress by English and Italian speakers», in Conte & Ramat (eds.): *Sprache im Kontext, Akten des 12. Linguistischen Kolloquiums*, Pavia 2, 19.
- BOVET P. & M. ROSSI (1978): «Étude comparée de la sensibilité différentielle à la durée avec un son pur et avec une voyelle», *IXes JEP*, Lannion, 43-55.
- CARLSON, R. & B. GRANSTROM (1975): «Perception of Segmental Duration», in Cohen & Nootboom (eds.), *Structure and Process in Speech Perception*, Springer-Verlag, 90-106.
- CREELMAN C.D. (1962): «Human discrimination on auditory duration», *Journal of the Acoustic Society of America*, 34/5, 582-593.
- DELATTRE, P. (1940): *La durée des voyelles en français*. Paris: D'Artrey.
- DELATTRE, P. (1966): «A Comparison of Syllable Length Conditioning among Languages» *IRAL* 4, 183-198.
- DELATTRE, P. (1967): «From acoustic cues to linguistic features», *6è Congrès International des Sciences Phonétiques*. Praguc: Académie des Sciences.
- DENEV, P. (1955): «Effect of duration on the perception of voicing», *Journal of the Acoustic Society of America*, 27/4, 761-764.
- EEK, A. (1978): «Just noticeable differences of duration and language type: some preliminary notes», *Estonian Papers in Phonetics*. Tallinn: Academy of Sciences of the Estonian S.S.R. 21-24.
- EEK A. (1979): «Estonian quantity: notes on the perception of duration». *Estonian Papers in Phonetics*. Tallinn: Academy of Sciences of the Estonian S.S.R. 529.
- EEK, A. (1980): *Studies in Quantity and Stress in Estonian*. Tartu: Dissertationes Philologiae Estonicae Universitatis Tartuensis, 4.
- EEK, A. (1987): «The perception of word stress: a comparison of Estonian and Russian», in Channon & Shokey (eds.): *In Honour of Ilse Lehiste*. The Netherlands Phonetic Archives, VI. Dordrecht: Foris 19-32.
- EEK, A. & MEISTER, E. (1997): «Simple perception experiments on Estonian word prosody: foot structure vs. segmental quantity», in Lehiste & Ross (eds.): *Estonian Prosody: Papers from a Symposium*, Tallinn: Institute of Estonian Language. 71-99.
- FOX, R.A. & LEHISTE, I. (1987): «Discrimination of duration ratios in bisyllabic tokens by native English and Estonian listeners, Part I», *Journal of Phonetics*, 15: 349-363.
- FOX, R.A. & LEHISTE, I. (1989): «Discrimination of duration ratios in bisyllabic tokens by native English and Estonian listeners, Part II», *Journal of Phonetics*, 157: 167-174.
- FRY, D.B. (1958) «Experiments in the perception of stress», *Language and Speech* 1, 126-152

- FUJISAKI, H., K. NAKAMURA & T. IMOTO (1975): «Auditory perception of duration of speech and non-speech stimuli», in Fant, G. & Tatham, A. (eds.) *Auditory analysis and perception of speech*. New York: Academy.
- HENRY, F. (1948): «Discrimination of the duration of sound», *J. Exp. Psychology*, 38, 734-743.
- HOUSE, D. (1999): «Perception of pitch and tonal timing: implications for mechanisms of tonogenesis», San Francisco: *ICPhS-99*, 1823-1826.
- HUGGINS, A. (1972): «Just noticeable differences for segment duration in natural speech», *Journal of the Acoustic Society of America*, 51, 1270-1278.
- KLATT, D.H. (1976): «Linguistic uses of segmental duration in English: acoustic and perceptual evidence», *Journal of the Acoustic Society of America*, 59, 1208-1221.
- KLATT, D.H.; W.E. COOPER (1975): «Perception of segment duration in sentence contexts», in A. Cohen & S.G. Noteboom (eds.) *Structure and process in speech perception*. Berlin: Springer, 69-89.
- KRULL, D. (1998) «Perception of quantity in Estonian: a preliminary study of words excised from conversational speech», contribution to the 11th Swedish Phonetics Conference. Abstract available at www.ling.su.se/FON/publications/fonctik98
- KRULL, D.; H. TRAUNMÜLLER (2000): «Perception of quantity in Estonian», contribution to *Fonetik 2000*. Electronic version in http://www.ling.su.se/staff/hartmut/Eesti_Q.htm
- LEHISTE, I. (1970b): *Suprasegmentals*. Cambridge Mass. M.I.T.
- LEHISTE I. (1975): «Experiments with synthetic speech concerning quantity in Estonian», in *Congressus Tertius Internationalis Fenno-Ugristarum*. Tallinn: Part I, 254-269.
- LEHISTE, I. (1976): «Suprasegmental Features of Speech» in: N. Lass: *Contemporary Issues in Experimental Phonetics*. Academic Press. N.Y. (225-239).
- LEHISTE, I. (1977) «Variability in the production of suprasegmental patterns», in *Studies in Finno-Ugric Linguistics*, 131 (Bloomington: University of Indiana), 131-139.
- LEHISTE, I. (1988): «Current debates concerning Estonian quantity», *FUSAC'88. Proceedings of the 6th annual meeting of the Fenno-Ugric Studies Association of Canada*. New York: Lanham. 77-86
- LEHISTE, I. (1997): «Search for phonetic correlates in Estonian Prosody», in Lehiste & Ross (eds.): *Estonian Prosody: Papers from a Symposium*, Tallinn: Institute of Estonian Language. 11-35.
- LEHISTE I. & ROSS J. (eds.) (1997): *Estonian Prosody: Papers from a Symposium*. Tallinn: Institute of Estonian Language.
- MÉLONI, H. & PH. LANGLAIS (1996): «Prosodie et reconnaissance de la parole», in Méloni, H.: *Perspectives et Fondements en Traitement Automatique de la Parole*: Éditions de l'AUPELF. reed. <http://www.bibliotheque.refer.org/parole/meloni/meloni.htm>
- NOOTEBOOM, S.G. (1972). *Production and Perception of Vowel Duration: a Study of Durational Properties of vowels in Dutch*. Ph.D. Thesis, Utrecht University.
- NOOTEBOOM, S.G. (1973): «The perceptual reality of some prosodic durations», *Journal of Phonetics*, 1, 25-45.
- NOOTEBOOM, S.G. (1979) Complex control of simple decisions in the perception of vowel length, in *Proceedings of the Ninth International Congress of Phonetic Sciences*, 6-11 August 1979, Institute of Phonetics, University of Copenhagen, Vol. II, 298-304.
- NOOTEBOOM S.G. & J.N. DOODEMAN (1980): «Production & perception of vowel length in spoken sentences», *Journal of the Acoustic Society of America*, 67, 276-287.
- PAMIES, A.: 1996 «Consideraciones sobre la marca acústica del acento fonológico», *Estudios de Fonética Experimental*, VIII, 11-49.

- PAMIES, A.; A. FERNÁNDEZ PLANAS; E. MARTÍNEZ CELDRÁN; A. ORTEGA ESCANDELL; M.C. AMORÓS CÉSPEDES (2001) «Umbrales tonales en español peninsular», II Congreso Nacional de Fonética Experimental, Universidad de Sevilla (en prensa).
- PAMIES, A.; E. IÑESTA; M. HAGERTY (1999): «Observaciones experimentales sobre la cantidad vocálica», en E. Martínez Celdrán, S. Planas Morales & J. Moreno Gallego (eds.): *Primer congreso Nacional de fonética experimental*. Tarragona: Universitat y Universitat Rovira Virgili.
- RHUM, H.B., E.O. MENCKE, B. MILBURN, W.A. COOPER, D.E. ROSE (1966): «Differential sensitivity to acoustic signals», *J. Speech, Hear. Res.* 9. 371-384.
- ROSSI, M. (1972): «Le seuil différentiel de durée», in A. Valdman (ed.) *Papers in Linguistics and Phonetics to the Memory of P. Delattre*, Mouton The Hague, 64-94.
- SOLÉ, M.J. (1984): «Experimentos sobre la percepción del acento», *Estudios de Fonética Experimental*, I. 135-242.
- TOLEDO, G. A. (2000): «Taxonomía tonal en español», *Language Design*, 3.
- VAISSIÈRE, J. (1977): «Quelques expériences d'analyse perceptuelle en français», in *Villes JEP*, 183-189.

- PAMIES, A.; A. FERNÁNDEZ PLANAS; E. MARTÍNEZ CELDRÁN; A. ORTEGA ESCANDELL; M.C. AMORÓS CÉSPEDES (2001) «Umbrals tonals en espanyol peninsular», II Congreso Nacional de Fonética Experimental, Universidad de Sevilla (en prensa).
- PAMIES, A.; E. IÑESTA; M. HAGERTY (1999): «Observaciones experimentales sobre la cantidad vocálica», en E. Martínez Celdrán, S. Planas Morales & J. Moreno Gallego (eds.): *Primer congreso Nacional de fonética experimental*. Tarragona: Universitat y Universitat Rovira Virgili.
- RHUM, H.B., E.O. MENCKE, B. MILBURN, W.A. COOPER, D.E. ROSE (1966): «Differential sensitivity to acoustic signals», *J. Speech, Hear. Res.* 9. 371-384.
- ROSSI, M. (1972): «Le seuil différentiel de durée», in A. Valdman (ed.) *Papers in Linguistics and Phonetics to the Memory of P. Delattre*, Mouton The Hague, 64-94.
- SOLÉ, M.J. (1984): «Experimentos sobre la percepción del acento», *Estudios de Fonética Experimental*, I. 135-242.
- TOLEDO, G.A. (2000): «Taxonomía tonal en español», *Language Design*, 3.
- VAISSIÈRE, J. (1977): «Quelques expériences d'analyse perceptuelle en français», in *Villes JEP*, 183-189.