

## LA DURACIÓN VOCÁLICA EN ESPAÑOL

RAFAEL MARÍN GÁLVEZ  
(Universidad Autónoma de Barcelona)

### RESUMEN

In this paper we describe an experimental analysis of vowel duration in Spanish.

Some factors that seem relevant in other studies for characterizing vowel duration were analyzed here using speech data from two speakers. These factors are: vowel identity, phrase boundaries, lexical stress, open versus closed syllables and effect of the postvocalic consonant.

Our results show that Spanish vowels have intrinsic duration and there are some contextual factors, like phrase final boundary, stressed syllables and open syllables, that influence vowel lengthening. For voicing and manner of articulation of postvocalic consonant, significant differences were not found.

*El Grup de Fonètica del Departament de Filologia Espanyola de la Universitat Autònoma de Barcelona* viene desarrollando desde hace unos años una línea de investigación orientada hacia el estudio de la prosodia en español y en catalán, y su aplicación en el desarrollo de sistemas de conversión de texto a habla.

El presente estudio\* se enmarca dentro de esta línea de investigación más general.

### 1. Introducción

El objetivo fundamental de este trabajo consiste en llevar a cabo una caracterización de la duración de las vocales en español. Para ello, resulta

---

\* Presentado en forma de comunicación en el XXIII Simposio de la Sociedad Española de Lingüística, celebrado en Lérida, del 13 al 16 de diciembre de 1993.

Agradezco muy sinceramente la ayuda prestada por Joaquim Llisterra, cuyos comentarios y sugerencias me han sido de gran utilidad.

necesario, en primer lugar, determinar qué factores influyen y en qué medida en la cantidad vocálica.

A fin de establecer cuáles son estos factores, hemos revisado varios trabajos que analizan el fenómeno de la duración, tanto para el español como para otras lenguas<sup>1</sup>. De ellos, hemos extraído una serie de variables que constituyen el punto de partida del estudio que presentamos aquí.

### **1.1. Factores que influyen en la duración vocálica**

Hemos elegido como factores que pueden influir en la duración de las vocales, para su inclusión en nuestro estudio, aquellos que aparecen de forma más recurrente en la bibliografía y cuya pertinencia se ha establecido de forma clara en otras lenguas.

No hemos incluido en el análisis la totalidad de las variables que parecen influir en la duración vocálica en otros estudios, ya que los resultados aquí presentados constituyen una primera aproximación al tema, orientada principalmente a la elaboración de un modelo que permita asignar de forma automática valores de duración a las vocales.

Dejando de lado la duración intrínseca de cada vocal, las variables que hemos tenido en cuenta a la hora de diseñar el experimento son las siguientes:

- Acento
- Estructura silábica
- Sonoridad de la consonante postvocálica
- Modo de articulación de la consonante postvocálica.
- Posición en la frase.

y como variables controladas:

- velocidad de elocución. (5-7 sílabas por segundo)
- número de sílabas de la palabra (2-3 sílabas)

## **2. Protocolo experimental**

### **2.1. Corpus**

Para obtener los datos de esta investigación hemos utilizado un corpus de lectura formado por dos textos extraídos del editorial de un periódico. El

---

<sup>1</sup> Para el español hemos revisado básicamente los trabajos de Navarro Tomás (1916 y 1917) y Borzone y Signorini (1983). Para otras lenguas, como el inglés, el número de trabajos dedicados a este tema es mucho más elevado (véase apartado de bibliografía).

número total de palabras contenidas en el corpus es de 1395.

Con el fin de que la distribución de las variables en el corpus resultara totalmente sistemática, hemos buscado contextos en los que mientras una variable cambia, las otras permanecen constantes. Para ello, ha sido necesario modificar debidamente los dos textos.

La distribución de las variables ha sido la siguiente: cada vocal aparece tres veces representada en cada uno de los contextos que previamente hemos definido, excepto en el caso de [u] que, debido a su menor frecuencia de aparición, solamente figura dos veces en un mismo contexto. De esta forma obtenemos un número total de 491 vocales. El análisis se ha llevado a cabo sobre esta muestra.

La primera división que hemos establecido a la hora de elaborar el corpus tiene que ver con la posición que ocupa una vocal en la frase. Hemos diferenciado, en primer lugar, dos grandes grupos: por un lado, el de las vocales que aparecen en posición no prepausal y, por otro, el de las vocales que aparecen en posición prepausal.

### **2.1.1. Vocales en posición no prepausal**

Por lo que respecta a las vocales en posición no prepausal, hemos distinguido entre vocales acentuadas y no acentuadas, y en cada uno de estos grupos, a su vez, entre las que pertenecen a sílaba trabada o a sílaba libre. Los grupos resultantes han sido divididos nuevamente según el modo de articulación de la consonante postvocálica.

En el caso de las vocales que aparecen en sílaba libre, hemos tenido en cuenta los siguientes modos de articulación de la consonante posterior a la vocal: oclusivo, aproximante, fricativo, nasal y líquido.

Por lo que se refiere a las vocales que pertenecen a sílaba cerrada, sólo hemos incluido los modos de articulación fricativo, nasal y líquido de la consonante posterior a la vocal. No hemos incluido los modos de articulación aproximante y oclusivo, debido a la imposibilidad —en el primer caso— y a la enorme dificultad —en el segundo— de encontrar tales sonidos en un contexto no intervocálico.

### **2.1.2. Vocales en posición prepausal**

En cuanto a las vocales que aparecen en posición prepausal<sup>2</sup> hemos seguido la misma distribución de las variables que en el caso de las vocales en posición no prepausal. Sin embargo, no nos ha sido posible incluir la variable modo de articulación, ya que incorporar de forma sistemática esta variable habría hecho

---

<sup>2</sup> Hemos considerado como posiciones prepausales únicamente aquellas que tienen como correlato ortográfico un punto, una coma o un punto y coma.

necesaria la elaboración de un corpus demasiado extenso y con gran abundancia de pausas.

## **2.2. Locutores**

Hemos entrevistado a dos informantes masculinos hablantes de español peninsular estándar. Ambos poseen un nivel cultural alto y sus edades oscilan entre 35 y 45 años.

## **2.3. Grabación**

### **2.3.1. Instrumental**

La grabación ha sido realizada en la cámara semianecoica del *Laboratori de Fonètica* de la *Universitat Autònoma de Barcelona*, mediante un magnetófono TASCAM, TEAC PROFESSIONAL DIVISION 112, que posee una respuesta de frecuencia de 25 Hz. a 18 KHz. +/- 3 dB, y un micrófono omnidireccional SENHEISER MKH20, con una respuesta de frecuencia de 20 Hz. a 20 KHz.

### **2.3.2. Método**

La lectura de los textos fue realizada a la velocidad de elocución normal del hablante.

Los informantes leyeron el texto varias veces antes de la grabación, ya que el hecho de estar familiarizados con el texto puede contribuir al logro de una mayor uniformidad en el ritmo de lectura.

Hemos comprobado, por un lado, que la velocidad de elocución de los dos informantes fuera similar y, por otro, que se mantuviera uniforme a lo largo de toda la grabación. Para ello, hemos calculado, eligiendo fragmentos de la grabación de forma aleatoria, cuál ha sido esta velocidad de elocución, que se sitúa, en ambos casos, entre 5 y 7 sílabas por segundo.

## **2.4. Análisis acústico**

### **2.4.1. Instrumental**

El estudio acústico ha sido realizado en un ordenador *Macintosh II* mediante el programa *Mac Speech Lab II*.

Para la digitalización (10 KHz.) hemos empleado la tarjeta *Mac Adios II™*.

### **2.4.2. Método**

La segmentación de los sonidos ha sido realizada, básicamente, sobre la representación oscilográfica de la onda sonora. No obstante, en aquellos casos en los que este tipo de representación no nos permitía distinguir claramente los límites entre sonidos, hemos utilizado también la representación espectrográfica.

### 3. Resultados

#### 3.1. Duración intrínseca

Los resultados que hemos obtenido, con respecto a la duración inherente de cada una de las cinco vocales del español, presentan un considerable grado de sistematicidad:

	x (ms)	dt	n
[a]	69,63	17,08	94
[e]	64,95	15,51	94
[i]	60,66	14,96	95
[o]	64,15	15,42	96
[u]	60,93	14,98	61

Tabla 1. *Valores medios de la duración de las vocales en todos sus contextos de aparición.*

Como se puede observar en la tabla 1, la vocal de mayor duración es [a] (69,63 ms.); le siguen [e] (64,95 ms.) y [o] (64,15 ms.), entre las que no hay diferencias significativas, y [u] (60,93 ms.) e [i] (60,66 ms.), que tampoco presentan divergencias entre sí. Por lo tanto, la gradación que se establece, en lo que a la duración vocálica se refiere, es la siguiente: [a] > [e] > [o] > [u] > [i].

De estos datos se puede extraer una conclusión más relevante: podemos constatar la existencia, con respecto a la cantidad, de tres grupos diferenciados en las vocales del español que, ordenados de mayor a menor duración presentan la siguiente gradación: [a] > [e],[o] > [u],[i]. En este aspecto concreto es donde más claramente se percibe la sistematicidad de que hablábamos al comienzo de este apartado: entre [a] y [e],[o] hallamos una diferencia de duración de

aproximadamente 5 ms. (un 7,5% de incremento); entre [e],[o] y [u],[i] la diferencia se sitúa en torno a los 4,5 ms. (un 6,5% de incremento), por lo tanto [a] presenta un incremento del 14% con relación a [u],[i].

La aplicación de un *t-test de Student* muestra que en todos los casos las diferencias son estadísticamente significativas.

En el gráfico que aparece a continuación podemos apreciar más claramente la existencia de los tres grupos de vocales que mencionábamos anteriormente:

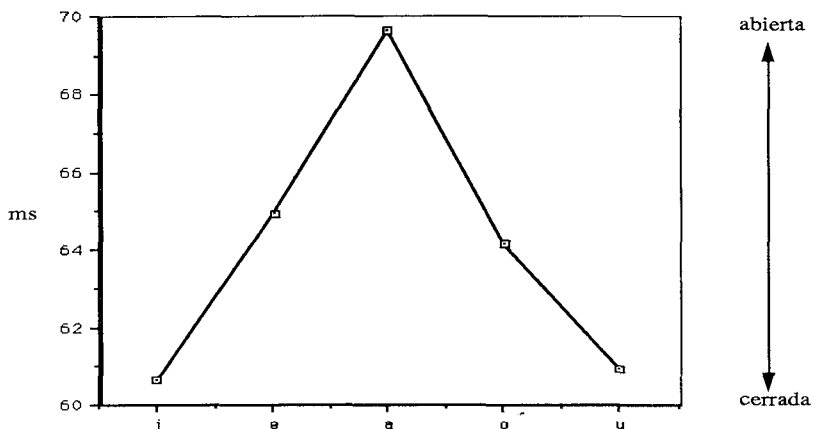


Figura 1. *Valores medios de la duración de las vocales en todos sus contextos de aparición.*

Como se observa en el gráfico anterior, la duración vocálica puede tener un correlato articulatorio. En este sentido, los resultados obtenidos nos permiten extraer la siguiente generalización: en las vocales del español, a una mayor obertura corresponde una mayor duración. Esta afirmación concuerda por completo con las generalizaciones que han sido propuestas por otros autores como, por ejemplo, Navarro Tomás (1916).

### 3.2. Acento

El acento es uno de los factores que influye de forma más clara en la duración de las vocales. En la tabla que aparece a continuación podemos observar que el valor medio de la duración de una vocal acentuada es considerablemente mayor que el valor medio de la duración de una vocal inacentuada.

	x (ms)	dt	n
+ac	70,17	16,11	222
-ac	58,31	13,32	218

Tabla 2. Valores medios de la duración de las vocales en función del acento.

El valor medio de la duración de las vocales acentuadas es de 70,17 ms. mientras que el de las inacentuadas es de 58,31 ms. Esto supone un incremento del 20,3%.

La aplicación de un *t-test de Student* muestra que las diferencias son estadísticamente significativas:

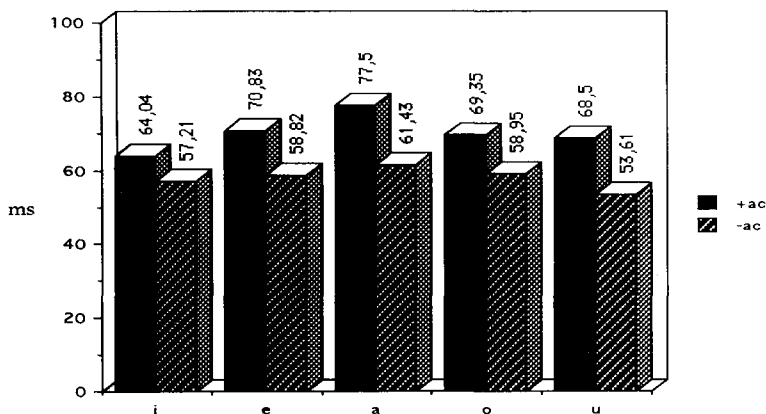


Figura 2. Valores medios de la duración de las vocales acentuadas y no acentuadas.

Como se puede apreciar en el gráfico anterior, al comparar la duración de cada una de las vocales en posición acentuada y no acentuada, observamos que, salvo en el caso de [u], los resultados reflejan un comportamiento muy similar de las vocales en relación con este fenómeno.

### 3.3. Estructura silábica

Si observamos los resultados que aparecen en la tabla 3, podemos constatar que las vocales que se encuentran en sílaba cerrada presentan prácticamente la misma duración que las que aparecen en sílaba abierta:

	x ms)	dt	n
ab	64,17	17,38	269
cer	64,36	15,02	276

Tabla 3. Valores medios de la duración de las vocales en función de la estructura silábica.

A partir de estos datos, debemos señalar que la estructura silábica, al contrario de lo que afirman otros autores, e. g. Navarro Tomás (1916) no parece ser un factor que influya en la duración de las vocales.

Si examinamos la influencia combinada de la estructura silábica y el modo de articulación de la consonante posterior a la vocal, podemos ver que la distribución de los datos dista mucho de ser homogénea:

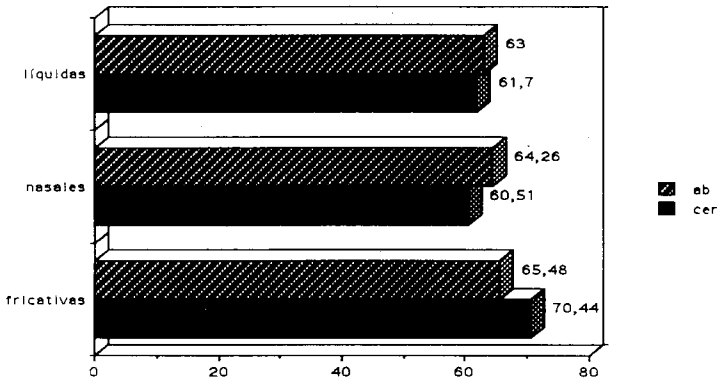


Figura 3. Valores medios de la duración de las vocales en función de la estructura silábica y del modo de articulación de la consonante posterior.



Como muestra el gráfico anterior, la duración de las vocales que aparecen en sílaba abierta es algo mayor que la duración de las que se encuentran en sílaba cerrada si ambas van seguidas de consonantes nasales o líquidas, pero su comportamiento es diferente si preceden a consonantes fricativas. En este último caso, las vocales que aparecen en sílaba abierta tienen una duración menor que las que aparecen en sílaba cerrada.

La aplicación de un *t-test de Student* muestra, además, que las diferencias no son estadísticamente significativas en ningún caso.

### 3.4. Sonoridad de la consonante postvocálica

Como se puede observar en el gráfico que aparece a continuación, una vocal seguida de una consonante sonora es sensiblemente más larga que una vocal seguida de una consonante sorda:

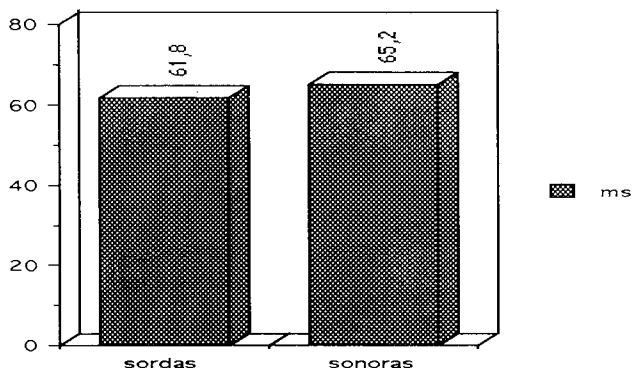


Figura 4. Valores medios de la duración de las vocales en función de la sonoridad de la consonante posterior.

No obstante, las diferencias que se aprecian no son estadísticamente significativas. La aplicación de un *t-test de Student* así lo demuestra.

### 3.5. Modo de articulación de la consonante postvocálica

En cuanto a la influencia del modo de articulación de la consonante posterior a la vocal, cabe decir que los datos no son lo suficientemente esclarecedores. La diferencia de duración que se establece en algunos casos —fricativas, nasales y

líquidas— es de tan sólo 2 ms. (véase figura 5)<sup>3</sup>.

La aplicación de un *t-test de Student* demuestra que las diferencias no son significativas en ningún caso, excepto en la comparación del par oclusiva sorda / aproximante:

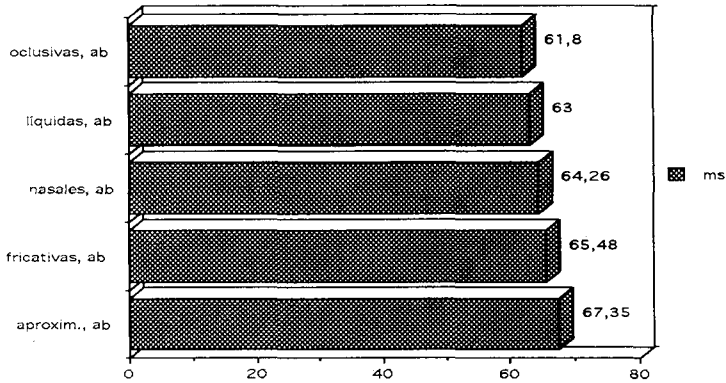


Figura 5. Valores medios de la duración de las vocales en función del modo de articulación de la consonante posterior.

Así pues, debemos concluir que el modo de articulación de la consonante posterior a la vocal no parece ser un factor que influya de forma notable en la duración de las vocales.

De todos modos, la gradación que se puede establecer, de mayor a menor duración, es la siguiente: aproximantes > fricativas > nasales > líquidas > oclusivas sordas.

### 3.6. Posición en la frase

La posición que ocupa un sonido en la frase es un factor que influye en su duración. Los resultados muestran la existencia de una correlación entre posición prepausal y una mayor duración de la vocal.

La media de duración de las vocales en posición prepausal es de 94,51 ms. Este valor medio contrasta con el de las no prepausales, que se sitúa en 64,29 ms. Ello representa un incremento del 47%.

Del total de vocales que se encuentran en posición final absoluta en nuestro

<sup>3</sup> En esta comparación, solo hemos tenido en cuenta las vocales que aparecen en sílaba abierta, ya que para las vocales en sílaba cerrada, no contamos con casos de vocales seguidas de oclusiva sorda o aproximante.

corpus, hemos distinguido entre las pertenecientes a sílaba trabada y a sílaba libre, por un lado, y entre acentuadas e inacentuadas, por otro.

Mediante la aplicación de un ANOVA, hemos podido comprobar que en el caso de la posición prepausal, las variables estructura silábica y acento ejercen una influencia combinada. Las vocales que presentan una mayor duración son aquellas que, además de aparecer en posición prepausal, se encuentran en sílaba libre y en posición acentuada: 121,50 ms. Por el contrario, las vocales de menor duración son las que aparecen en sílaba trabada y en posición inacentuada: 88,33 ms.

Al analizar por separado la influencia del acento y de la estructura silábica obtenemos los resultados que aparecen en la tabla siguiente:

	x (ms)	dt	n
+ac	100,93	27,46	16
-ac	91,57	20,06	35
ab	85,88	15,62	26
cer	103,48	25,74	25

Tabla 4. *Valores medios de la duración de las vocales en posición prepausal en función del acento y de la estructura silábica.*

Estos resultados muestran, por un lado, que una vocal acentuada presenta un valor medio de duración mayor que el de una vocal inacentuada (100,93 ms. y 91,57 ms. respectivamente) y, por otro, que el valor medio de duración de las vocales que se encuentran en sílaba cerrada es menor que el de las vocales que se hallan en sílaba abierta (85,88 ms. y 103,48 ms. respectivamente).

Con estos datos, podemos concluir que, si bien el hecho de que una vocal aparezca en posición final absoluta implica un incremento de la duración, éste será mayor o menor dependiendo de otros dos factores: el acento y la estructura silábica.

#### 4. Conclusiones

A partir de los resultados obtenidos en este estudio, hemos extraído las siguientes conclusiones:

Podemos diferenciar tres grupos vocálicos ordenados de mayor a menor duración: [a] > [e]/[o] > [i]/[u]. Ello nos permite relacionar un aspecto acústico con otro articulatorio: a una mayor obertura corresponde una mayor duración de la vocal. A la luz de estos resultados, debemos concluir que las vocales se diferencian acústicamente por su duración.

El acento resulta ser un factor que influye claramente en la duración vocálica: las vocales acentuadas presentan una mayor duración que las vocales inacentuadas, tanto en posición prepausal como en posición no prepausal, aunque en este último caso el incremento es menor.

Hemos podido constatar también una notable influencia de la estructura silábica en la duración de las vocales que aparecen en posición prepausal: las vocales que pertenecen a sílaba abierta presentan una mayor duración que las que pertenecen a sílaba cerrada. Sin embargo, en posición no prepausal no se observan diferencias significativas.

La duración de una vocal que precede a una consonante sonora es mayor que la de una vocal que antecede a una consonante sorda, aunque las diferencias no son estadísticamente significativas.

Los resultados obtenidos en lo que se refiere a la influencia del modo de articulación de la consonante posterior a la vocal no indican que éste sea un factor que afecte a la duración de los sonidos vocálicos. Las diferencias, aunque mínimas, permiten establecer la siguiente gradación: aproximantes > fricativas > nasales > líquidas > oclusivas sordas.

Por último, cabe señalar que las vocales que se encuentran en posición prepausal presentan claramente una mayor duración que las vocales en posición no prepausal.

## 5. Investigaciones futuras

Tomando como punto de partida los resultados de este trabajo, pensamos diseñar un modelo, implementable en un sistema de conversión de texto a habla, que permita asignar de forma automática valores de duración a las vocales.

Una investigación de este tipo puede resultar muy fructífera, ya que al disponer de datos implementados en un conversor de texto a habla, podemos evaluar los resultados mediante tests de percepción.

## BIBLIOGRAFÍA

- BORZONE A. M. y SIGNORINI A., (1983). «Segmental duration and rythm in Spanish», *Journal of Phonetics*, 11, pp. 117-128.
- CAMPBELL, W. N. (1992). «Syllable-based segmental duration», in *Talking Machines: Theories, Models and Designs*, ed. by G. Bailly, C. Benoit and T.

- R. Sawallis (North-Holland, Amsterdam), pp. 211-224.
- CARLSON, R. and GRANSTRÖM, B. (1986). «A search for durational rules in a real-speech database», *Phonetica*, 43, 140-154.
- CRYSTAL, T. H., y HOUSE, A. S. (1988a). «Segmental durations in connected speech signals: Current results», *JASA*, 83(4), pp. 1553-1573.
- CRYSTAL, T. H., y HOUSE, A. S. (1988b). «Segmentals durations in connected speech signals: Syllabic stress», *JASA*, 83(4), pp. 1574-85.
- CRYSTAL, T. H., y HOUSE, A. S. (1988c). «The duration of American English vowels: An Overview», *Journal of Phonetics*, 16, pp. 263-284.
- CHEN, M., (1970). «Vowel length variation as a function of the voicing of the consonant environment», *Phonetica*, 22, pp. 129-159.
- DELATTRE, P. (1962). «Some factors of vowel duration and their cross-linguistic validity», *JASA*, 34, pp. 1141-1143.
- FARNETANI, E. y KORI, S. (1986). «Effects of syllable and word structure on segmental durations in spoken italian», *Speech Communication*, 5, pp. 17-34.
- GILI y GAYA (1940). «La cantidad silábica en la frase», *Castilla* (Valladolid), 1, pp. 287-298.
- GOPAL, H. S. (1990). «Effects of speaking rate on the behavior of tense and lax vowel durations», *Journal of Phonetics*, 18, pp. 497-518.
- HOUSE, A. S. (1961). «On vowel duration in English», *JASA*, 33, pp. 1174-1178.
- HOUSE, A. S. y FAIRBANKS, G. (1953). «The influence of consonant environment upon the secondary acoustical characteristics of vowels», *JASA*, 25, pp. 105-113.
- JANSON, T. (1979). «Vowel duration, vowel quality, and perceptual compensation», *Journal of Phonetics*, 7, pp. 93-103.
- KLATT, D. H. (1973). «Interaction between two factors that influence vowel duration. *JASA*, 54, pp. 1102-1104.
- KLATT, D. H. (1975). «Vowel lengthening is syntactically determined in a connected discourse», *Journal of Phonetics*, 3, pp. 129-140.
- KLATT, D. H. (1976). «Linguistic uses of segmental duration in English: Acoustic and perceptual evidence», *JASA*, 59, pp. 1208-1222.
- LEHISTE, I. (1970). *Suprasegmentals*, The M.I.T. Press, Cambridge, Massachusetts and London.
- LEHISTE, I. (1973). «Production and Perception of Vowel Duration. A Study of Durational Properties of Vowels in Dutch», *Journal of Phonetics*, 1, pp. 285-287.
- LUCE, P. A., y CHARLES LUCE, J. (1985). «Contextual effects on vowel duration, cosure duration, and the consonant/vowel ratio in speech production», *JASA*, 78, pp. 1949-1957.
- MACARRÓN, A. *et alia*. (1991). «Generation of duration rules for a Spanish text-to-speech synthesizer, *Proceedings of Eurospeech 91*, pp. 617-620.

- MILLER J. L. (1981). «Effects of Speaking Rate on Segmental Distinctions», en EINAS, P. D. y MILLER, J. L. (eds), *Perspectives on the Study of speech*, Lawrence Erlbaum Assoc. Hilydale, pp. 40-74.
- NAVARRO TOMÁS, T. (1916). «Cantidad de las vocales acentuadas», *RFE*, III, pp. 387-407.
- NAVARRO TOMÁS, T. (1917). «Cantidad de las vocales inacentuadas», *RFE*, IV, pp. 371-388.
- O'SAUGHNESSY, D. (1981). «A Study of French vowel and consonant durations», *Journal of Phonetics*, 9, pp. 385-406.
- O'SAUGHNESSY, D. (1974). «A multispeaker analysis of durations in read French paragraphs», *JASA*, 76, n.º 6, pp. 1664-1672.
- OLLER, D. K. (1973). «The effect of position in utterance on speech segment duration in english», *JASA*, pp. 1235-1247.
- PETERSON, G. E. y LEHISTE, I. (1960). «Duration of syllable nuclei in English», *JASA*, 32, pp. 693-703.
- RAPHAEL, J. L. (1975). «The physiological control of durational differences between vowels preceding voiced and voiceless consonants in English», *Journal of Phonetics*, 3, pp. 25-33.
- SANTOS, A. *et alia*. (1988). Diseño y evaluación de reglas de duración en la conversión de texto a voz, *SEPLN*, boletín nº 6, pp. 71-92.
- SHARF, D. (1962). «Duration of post-stress intervocalic stops and preceding vowels», *Language and Speech*, 5, pp. 26-30.
- SUMMERS, W. V. (1987). «Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses», *JASA*, 82, pp. 847-863.
- UMEDA, N. (1975). «Vowel duration in American English», *JASA*, 58, n.º 2, pp. 434-445.
- VAN SANTEN, J. P. H. (1992a). «Contextual effects on vowel duration», *Speech Communication*, 11, pp. 513-546.
- VAN SANTEN, J. P. H. and OLIVE, J. P. (1990). «The analysis of contextual effects on segmental duration», *Computer Speech and Language*, Vol. 4, pp. 359-391.
- ZIMMERMAN, S. A. y SAPON, S. M. (1958). «Note on vowel duration seen cross-linguistically», *JASA*, 30, pp. 152-153.