

Estruch, M., Garrido, J. M., Llisterri, J., & Riera, M.
(2007). Técnicas y procedimientos para la representación
de las curvas melódicas. *Revista de Lingüística Teórica y
Aplicada*, 45(2), 59-87.

[http://liceu.uab.cat/~joaquim/publicacions/
Estruch_Garrido_Llisterri_Riera_Metodos_
Entonacion_07.pdf](http://liceu.uab.cat/~joaquim/publicacions/Estruch_Garrido_Llisterri_Riera_Metodos_Entonacion_07.pdf)

TECNICAS Y PROCEDIMIENTOS PARA LA REPRESENTACION DE LAS CURVAS MELODICAS

TECHNIQUES AND PROCEDURES FOR REPRESENTING THE MELODIC CONTOUR

MONICA ESTRUCH¹
JUAN MARIA GARRIDO²
JOAQUIM LLISTERRI¹
MONTSERRAT RIERA¹

RESUMEN

Este trabajo pretende presentar la complejidad que supone el análisis de los fenómenos asociados a la variación melódica en el habla. Se propone que es necesario distinguir tres niveles de representación: un nivel fonético, uno fonológico y un nivel que represente la jerarquía de las unidades prosódicas. El nivel fonético debería incluir la curva melódica original, una representación estilizada y una anotación simbólica. El énfasis en los diferentes niveles dependerá de los objetivos de la investigación.

Palabras claves: Prosodia, entonación, niveles de representación, estilización, anotación.

ABSTRACT

This paper aims at discussing the complexity involved in the analysis of the phenomena associated to melodic variation in speech. It argues that it is necessary to distinguish three levels of representation: a phonetic level, a phonological one and a level representing the hierarchy of prosodic units. The phonetic level should include the actual melodic contour, a stylized representation and a symbolic annotation. The emphasis on different levels depends on the objectives of the research.

Keywords: Prosody, intonation, levels of representation, stylization, annotation.

Recibido: 26-06-2007. *Aceptado:* 10-09-2007.

¹Departament de Filologia Espanyola, Universitat Autònoma de Barcelona. monica@emediterrania.com, Joaquim.Llisterri@uab.cat, Montserrat.Riera@uab.cat

²Departament de Traducció i Filologia, Universitat Pompeu Fabra. juanmaria.garrido@upf.edu

1. INTRODUCCION

LA REPRESENTACION de las curvas melódicas es uno de los principales problemas que, tradicionalmente, se han planteado en el estudio de la entonación. La primera dificultad para el análisis prosódico del habla aparece ya, como veremos más adelante, en el propio proceso de detección de la frecuencia fundamental (F_0). Por otra parte, no se dispone aún de un sistema generalizado de representación de los fenómenos y unidades entonativas, a diferencia de lo que ocurre en el plano segmental, para el que existe un alfabeto fonético (el Alfabeto Fonético Internacional) (IPA, 1999) universalmente reconocido como el procedimiento de representación de los sonidos de cualquier lengua.

Los problemas que se acaban de señalar están relacionados, en parte, con la indefinición que caracteriza, en algunos casos, el estudio de los fenómenos entonativos. La entonación se ha abordado tanto desde una perspectiva fonética como desde un punto de vista fonológico, lo que ha tenido como consecuencia la definición de procedimientos de representación diferentes en función del predominio de uno de los dos enfoques¹. Un segundo factor, de fundamental importancia, es la dificultad inherente que conlleva el análisis de un fenómeno de estructura tan compleja como son las curvas melódicas.

La obtención de representaciones válidas y generales de las curvas melódicas puede concebirse como un proceso que parte de la representación de la evolución temporal de la F_0 a lo largo de un enunciado (denominada, en este trabajo, curva melódica) para llegar a una representación simbólica (anotación), en la que se han eliminado las diferentes fuentes de variación que dificultan la interpretación de la curva. Este proceso se realiza, fundamentalmente, en tres etapas:

- a) el cálculo de la curva melódica (estimación de la F_0)
- b) la obtención de la curva estilizada correspondiente
- c) la anotación o transcripción.

Estas tres fases pueden verse como procesos independientes, ya que cada una de ellas nos ofrece un tipo de representación diferente, que puede ser utilizada por sí misma para el análisis de la entonación, o como etapas en un mismo proceso que permite el paso de la forma fonética a la representación fonológica. En este trabajo se presentan los problemas que se plantean y los métodos que suelen emplearse en cada una de estas fases, procurando mostrar, en conjunto, algunos de los principios metodológicos básicos que pueden guiar el estudio de la entonación.

¹ Para una revisión de los diferentes modelos de análisis de la entonación, véase Prieto (Ed.) (2003).

2. DETECCIÓN DE LA F_0

Para obtener una representación acústica de la evolución temporal de la F_0 a lo largo de un enunciado se recurre, normalmente, a algoritmos de detección de la F_0 que actúan directamente sobre la señal acústica para detectar la periodicidad de la misma y la longitud del periodo. A continuación se describen las estrategias más comúnmente empleadas y los procedimientos que permiten eliminar algunos de los errores inherentes a este proceso.

2.1. Técnicas de estimación de la F_0

Existen diferentes técnicas de detección de la F_0 , que pueden agruparse en dos grandes tipos: las que actúan en el dominio temporal y las que lo hacen en el dominio frecuencial². Las técnicas que actúan en el dominio temporal intentan reconocer en la onda sonora formas periódicas recurrentes. Así ocurre, por ejemplo, con los métodos basados en la detección de los picos de la onda sonora (*peak-picking*). Se trata de técnicas relativamente simples que, por tanto, no necesitan mucho tiempo de cálculo, pero cuyos resultados pueden contener fácilmente errores.

Los procedimientos basados en el análisis frecuencial no actúan directamente sobre la forma de la onda, sino que determinan la distancia existente entre los armónicos del espectro. La autocorrelación, la compresión espectral o la comparación de los armónicos son ejemplos de técnicas de este tipo. En general, se trata de técnicas más fiables que las temporales, pero mucho más lentas, dado que requieren mayor tiempo de cálculo.

En la Figura 1 se presenta la curva melódica correspondiente al enunciado “Una vez en el interior, los agentes encontraron a tres de los niños”, obtenida mediante dos técnicas diferentes de estimación de la F_0 : una temporal (el método AMDF o *Average Magnitude Difference Function*) y otra frecuencial (el método Comb), ambas desarrolladas en el *Laboratoire Parole et Langage* de la *Université de Provence*, e integradas en el programa ‘Phonédit’, comercializado por la empresa SQLab³; con ello se pueden apreciar las diferencias de estimación existentes entre uno y otro sistema.

² Una presentación general de los diferentes sistemas de estimación de la F_0 puede encontrarse en Hess (1983) y en Gerhard (2003).

³ S.Q.Lab, Aix-en-Provence: <http://aune.lpl.univ-aix.fr/~sqlab/>

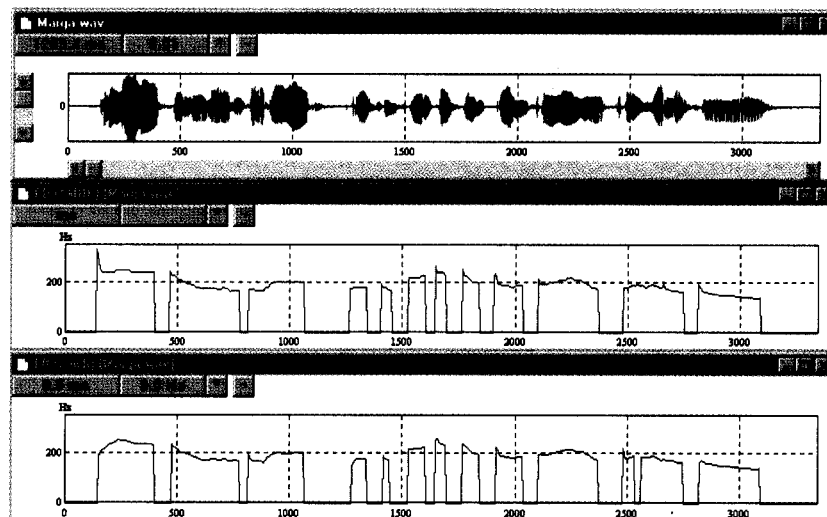


Figura 1. Oscilograma (ventana superior), curva melódica obtenida mediante un sistema de estimación temporal (ventana central) y curva melódica obtenida mediante un sistema de estimación frecuencial (ventana inferior) correspondiente al enunciado “Una vez en el interior, los agentes encontraron a tres de los niños”, pronunciado por un locutor femenino.

Las curvas melódicas que se obtienen tanto con los sistemas temporales como con los sistemas frecuenciales plantean aún algunos problemas si pretendemos utilizarlas como punto de partida para un análisis de la entonación, ya que presentan a menudo errores (la omisión o duplicación de periodos, la consideración de segmentos sordos como sonoros y viceversa, etc.) que es necesario eliminar o minimizar aplicando algunos de los procedimientos que se describen en los siguientes apartados.

2.2. La minimización de los errores de detección

Las dos técnicas más comunes de minimización de los errores de detección son la limitación del rango de búsqueda y el alisado.

2.2.1. La limitación del rango de búsqueda

Una manera de reducir los errores de detección, especialmente los ocasionados por la duplicación y omisión de periodos, es delimitar el rango frecuencial de búsqueda

en función del rango frecuencial del locutor analizado. En el caso de un hablante masculino, por ejemplo, el rango frecuencial suele situarse entre 80 y 300 Hz, mientras que en el caso de un hablante femenino, el rango oscila entre 130 y 525 Hz en el registro modal (Lieberman y Blumstein, 1988; Orlikoff y Kahane, 1996). Es muy probable, por tanto, que los valores situados fuera de este rango sean el resultado de errores cometidos por el detector. Limitando el rango de búsqueda del detector de F_0 , el algoritmo descarta directamente los valores situados por encima o por debajo del umbral mínimo y máximo. Sin embargo, es necesario conocer previamente el rango frecuencial del locutor para ajustar correctamente el sistema.

La Figura 2 presenta un ejemplo de aplicación de esta técnica: puede observarse cómo en la segunda versión de la curva (obtenida delimitando el rango frecuencial entre 130 y 300 Hz) han desaparecido los errores de estimación presentes en la primera estimación (donde el rango frecuencial de detección está comprendido entre 75 y 600 Hz).

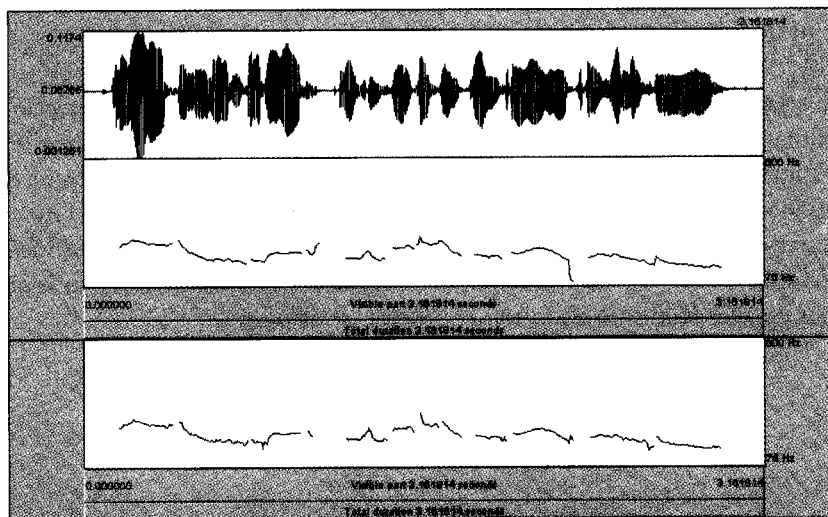


Figura 2. Oscilograma (ventana superior), curva melódica obtenida delimitando el rango entre 75 y 600 Hz (ventana central) y curva melódica obtenida delimitando el rango entre 130 y 300 Hz (ventana inferior) correspondiente al enunciado “Una vez en el interior, los agentes encontraron a tres de los niños”, pronunciado por un locutor femenino.

2.2.2. Las técnicas de alisado

Otro modo de minimizar los errores utilizado por algunos programas es la incorporación de un proceso de alisado (*smoothing*) posterior a la detección de los valores de F_0 . Durante este posproceso se aplican una serie de restricciones para detectar los valores de F_0 que sean resultado de un error de estimación. Este procedimiento consta de dos fases:

a) En primer lugar se compara cada valor de F_0 con los valores anteriores y/o posteriores, y se establece un porcentaje máximo de variación con respecto a los valores adyacentes, considerando así que aquellos valores de F_0 que superan el porcentaje de variación constituyen errores de estimación. El número de valores adyacentes con los que se compara cada valor de F_0 , así como el porcentaje de variación, son específicos de cada programa. Así, por ejemplo, en el programa 'MES' (desarrollado en el *Laboratoire Parole et Langage* de la *Université de Provence*)⁴, cada valor de F_0 se compara con los valores anteriores y posteriores y, si supera un porcentaje de variación predeterminado (el 5% en este caso), es considerado un valor 'erróneo'. En otros algoritmos, sin embargo, los valores se comparan únicamente con el valor anterior de F_0 y el porcentaje de variación puede ser modificado.

b) En segundo lugar, los valores de F_0 considerados como errores de detección se aproximan a los valores adyacentes, se neutralizan o se eliminan de la curva melódica. En la Figura 3 se puede observar un ejemplo de aplicación de esta técnica: en la ventana central se presenta la curva melódica tal como ha sido calculada mediante el sistema de estimación de la F_0 que está integrado en el programa 'Praat' (Boersma y Weenink, 2007), con los correspondientes errores de detección; en la ventana inferior aparece la misma curva tras el proceso de alisado.

⁴ MES Signaux package, Laboratoire Parole et Langage, Université de Provence: http://aune.lpl.univ-aix.fr/projects/mes_signaux.htm/

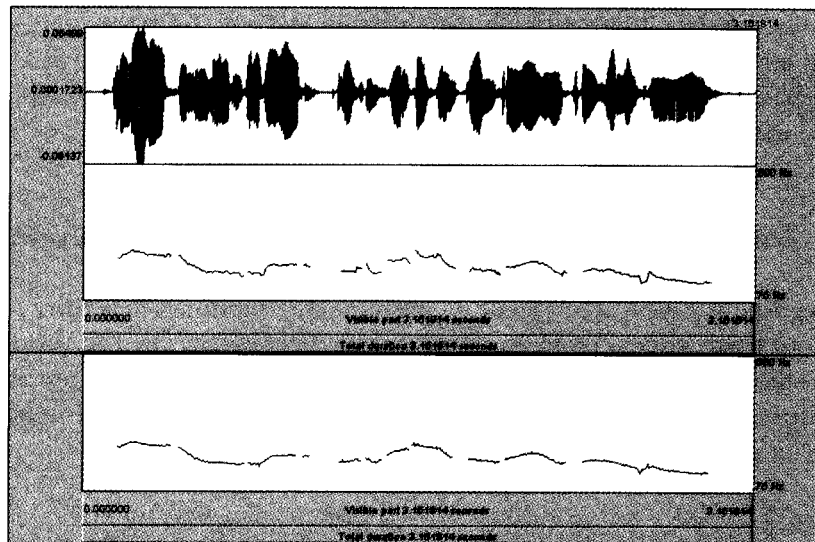


Figura 3. Oscilograma (ventana superior), curva melódica obtenida mediante un sistema de estimación frecuencial sin alisado (ventana superior) y con alisado (ventana central) correspondiente al enunciado “Una vez en el interior, los agentes encontraron a tres de los niños”, pronunciado por un locutor femenino.

3. ESTILIZACION

Después de eliminar en lo posible los errores de detección, las curvas melódicas contienen aún rupturas y pequeñas variaciones debidas a la propia naturaleza de los elementos segmentales. No deben confundirse con errores de estimación, pero puesto que no se trata de variaciones relevantes desde el punto de vista lingüístico, es conveniente, si se pretende realizar una descripción fonética o fonológica de la melodía, emplear procedimientos que las minimicen.

3.1. Las variaciones micromelódicas

Como se ha indicado anteriormente, las curvas melódicas contienen variaciones y rupturas, relacionadas con las características intrínsecas de los elementos segmentales que componen los enunciados, como son las interrupciones en la curva debidas a la presencia de segmentos sordos, o las pequeñas variaciones en el curso de la curva

melódica debidas a la aparición de determinados alófonos. Estas variaciones, llamadas micromelódicas, no aportan información lingüística relacionada con la interpretación de la curva melódica, aunque pueden constituir un indicio adicional para el reconocimiento del segmento en cuestión.

3.1.1. Interrupciones debidas a la aparición de segmentos sordos a lo largo del enunciado

Los segmentos sordos, al carecer de F_0 , provocan interrupciones en la curva melódica, que pueden causar problemas en el proceso de interpretación. Estas interrupciones no son perceptibles para el oyente, que posiblemente realiza un proceso de reconstrucción del contorno, y no son pertinentes en el estudio de la entonación desde un punto de vista puramente lingüístico.

3.1.2. Variaciones de F_0 debidas a la naturaleza de los elementos segmentales

Las variaciones de F_0 debidas a la propia naturaleza de los elementos segmentales afectan, al igual que las relacionadas con la sonoridad, a la forma de la curva melódica.

Tal y como se describe en Di Cristo (1982), en una curva melódica pueden encontrarse, fundamentalmente, tres tipos de variaciones debidas a la naturaleza de los elementos segmentales:

- El llamado ‘fundamental intrínseco’ de las vocales, que depende del grado de abertura de las mismas, de forma que cuanto más cerrada es la vocal, más alto es el valor de F_0 (Peterson y Barney, 1952, entre otros; Mateo, 1988, para el español).
- Los pequeños descensos en la frecuencia fundamental ocasionados por las características de ciertas clases de consonantes, como las aproximantes y las vibrantes (Di Cristo, 1982; Buenafuentes, Madrigal y Garrido, 2000, para el español).
- Los ascensos y descensos que se observan antes y después de un segmento sordo, debidos a los efectos de la coarticulación (Lehiste y Peterson, 1961, entre otros; Gili Gaya, 1924 y Buenafuentes *et al.*, 2000, para el español).

En la Figura 4 se aprecian algunas muestras de este tipo de variaciones. Obsérvese, por ejemplo, el descenso que provoca la vibrante simple intervocálica que aparece en la palabra ‘interior’, o el pequeño movimiento al inicio de la vocal que sigue inmediatamente a la consonante [t] en esa misma palabra.

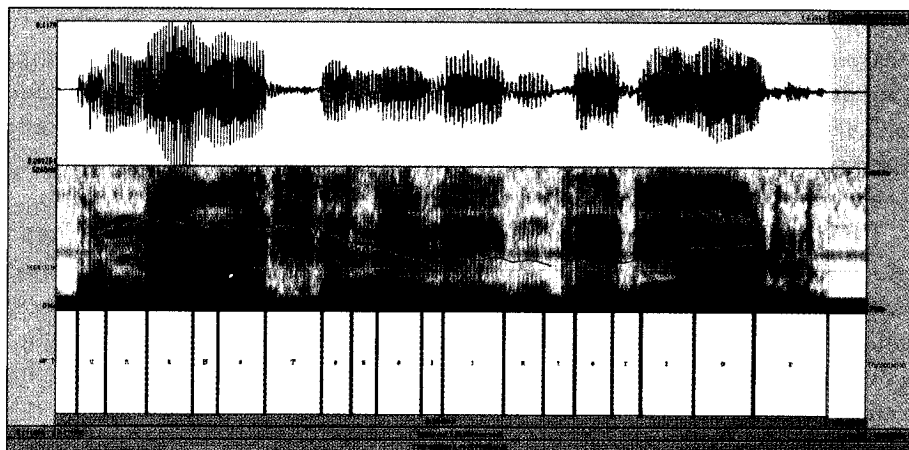


Figura 4. Oscilograma y curva melódica correspondientes al enunciado “Una vez en el interior”, pronunciado por un locutor femenino. Las líneas verticales representan los límites entre segmentos.

Si bien pueden constituir un indicio adicional en el reconocimiento de un sonido, al igual que ocurría en el caso de las interrupciones debidas a la aparición de segmentos sordos en el enunciado, estas variaciones no parecen transmitir ningún tipo de información lingüística.

3.2. El concepto de estilización

Con el fin de eliminar las variaciones a las que hasta ahora se ha hecho referencia, se ha planteado en numerosos estudios la posibilidad de llevar a cabo una estilización, o simplificación de la curva melódica original. En la estilización, las curvas melódicas quedan reducidas a una serie discreta de puntos, considerados como valores de F_0 relevantes para el análisis. Estos puntos pueden, además, estar unidos mediante líneas (rectas o curvas) que nos proporcionan una representación de los movimientos en los que se puede descomponer la curva.

3.3. Tipos de estilización

Según el número de puntos que se utilice para la representación, la estilización será ancha (detección de un número menor de puntos) o estrecha (detección de un número mayor de puntos). En general, una estilización basada en criterios de tipo

lingüístico, en la que se obtiene un valor de F_0 para cada sílaba, por ejemplo, será más estrecha que una estilización basada en criterios acústicos, donde se detecta un valor de F_0 –denominado “punto de inflexión”– al observar un cambio importante en la dirección y la pendiente de la curva. Partiremos de esta distinción para la descripción de los diferentes procedimientos de estilización en las secciones siguientes.

Otra posible clasificación de los sistemas de estilización se basa en la diferencia entre procedimientos manuales y automáticos. Aunque los primeros métodos de estilización se aplicaban manualmente, hoy en día la mayoría de los sistemas que se utilizan son automáticos. La necesidad de manejar cada vez corpus más amplios (por ejemplo, en aplicaciones de conversión de texto en habla) y los avances en las técnicas de procesado han llevado al desarrollo de estos sistemas automáticos.

3.3.1. Métodos basados en criterios lingüísticos

Como ya se ha señalado, existen algunos sistemas de estilización que se caracterizan por el uso de información lingüística para la localización de los puntos relevantes. Esta información lingüística suele consistir en los límites de los segmentos, o de las sílabas, que componen el enunciado.

Teniendo en cuenta que los segmentos vocálicos son las partes que se ven menos afectadas por las variaciones micromelódicas, los valores de F_0 se pueden tomar en algún (o algunos) punto(s) de la curva melódica coincidentes con las vocales del enunciado analizado. Varios sistemas manuales aplicaron esta idea, con variantes en cuanto al número de puntos y la posición dentro de la vocal en la que éstos se tomaban:

- En los estudios sobre el danés, ya clásicos, realizados por Thorsen (1983), la F_0 de cada una de las vocales y de las consonantes silábicas se midió en un punto determinado de éstas (a 2/3 de la distancia temporal desde el inicio del núcleo silábico).
- En Garrido, Llisterri, De la Mota y Ríos (1993, 1995) y en Garrido, Llisterri, Marín, De la Mota y Ríos (1995) los valores se obtuvieron en el centro de las vocales que constituyen núcleo silábico. En estos estudios, en los que se llevó a cabo el análisis de un corpus de lectura de oraciones y párrafos en español, el valor de F_0 se determinó en el punto de inflexión del contorno melódico de la sílaba o en el centro del núcleo silábico en caso de que no se hallara un punto de inflexión.
- En Estruch y Garrido (1995) se tuvo en cuenta la existencia de variaciones de F_0 que se producen dentro de una misma vocal y que, empleando métodos como los anteriores, pueden no quedar reflejadas en la estilización resultante. En este caso,

la representación de las curvas melódicas se obtuvo partiendo del valor de la F_0 en el inicio, centro y final de la vocal, tal como ilustra la Figura 5. También se aplicó este criterio en el análisis del corpus de referencia empleado en el estudio sobre la percepción del acento léxico en español descrito en Llisterri, Machuca, De la Mota, Riera y Ríos (2005a).

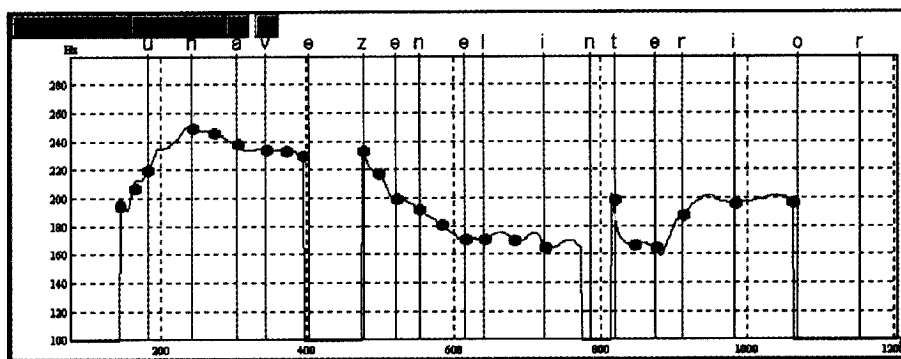


Figura 5. Curva melódica y estilizada correspondientes al enunciado “Una vez en el interior”, pronunciado por un locutor femenino. Los círculos indican los puntos de la curva que se corresponden con el inicio, centro o final de los segmentos vocálicos del enunciado.

Hasta hace poco tiempo, la segmentación de la señal acústica en alófonos o sílabas era un proceso manual y, por tanto, costoso. Sin embargo, la aparición de sistemas de segmentación automática ha permitido también la automatización de este tipo de sistemas de estilización.

Un ejemplo de sistema de estilización automática basada en información lingüística es el prosograma⁵ (*prosogram*), desarrollado por Piet Mertens (D’Alessandro y Mertens, 1995; Mertens, 2004). Este sistema estiliza los contornos a partir de criterios perceptivos, de forma completamente automática, utilizando como base la segmentación en alófonos, y pretende ofrecer contornos estilizados equivalentes a la representación auditiva de los oyentes. El algoritmo está disponible en forma de *script* para el programa ‘Praat’⁶. La Figura 6 muestra un ejemplo de estilización con este método.

⁵ P. Mertens, The Prosogram. Department of Linguistics, K. U. Leuven: <http://bach.arts.kuleuven.be/pmertens/prosogram/>

⁶ Prosogram scripts, P. Mertens, Department of Linguistics, K. U. Leuven: <http://bach.arts.kuleuven.be/pmertens/prosogram/download.html>

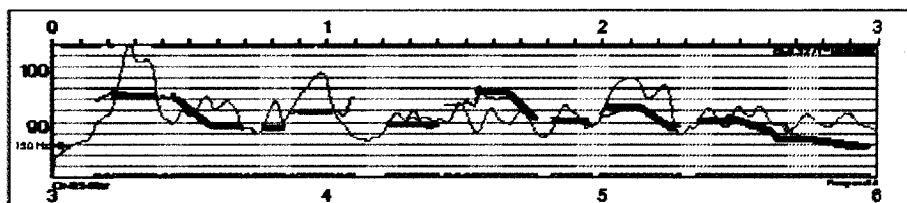


Figura 6. Prosograma correspondiente al enunciado “Una vez en el interior, los agentes encontraron a tres de los niños”, pronunciado por un locutor femenino.

3.3.2. Métodos basados en criterios acústicos

Los métodos basados en criterios acústicos no requieren una segmentación previa de los enunciados en unidades fonéticas o lingüísticas. Estos procedimientos convierten la curva melódica en una serie de puntos de inflexión unidos mediante líneas. Las diferencias entre los distintos métodos de este grupo radican, por una parte, en el procedimiento de obtención de los puntos de inflexión y, por otra, en la interpolación de los puntos, realizada mediante líneas rectas o curvas. El objetivo común a todos ellos, sin embargo, es que la curva estilizada sea perceptivamente equivalente a la original.

El más antiguo de estos sistemas es, probablemente, el desarrollado por los investigadores del IPO en Eindhoven, descrito en T’Hart, Collier y Cohen (1990) y en Garrido (2003). En este método, el objetivo principal es obtener una versión estilizada de la curva original formada por líneas rectas que cumpla dos requisitos: no distinguirse perceptivamente de la curva original y contener el menor número posible de segmentos (líneas rectas). Este tipo de representación recibe el nombre de *close-copy stylization*. Se trataba de un procedimiento manual, que requería un gran esfuerzo para la obtención de las estilizaciones.

Un procedimiento semejante es el presentado en Garrido (1991a, b), aunque en este caso el criterio empleado para la construcción del contorno estilizado no es perceptivo, sino acústico, ya que se tenían en cuenta las diferencias frecuenciales entre los distintos puntos de la curva melódica para la determinación de los puntos de inflexión. Dicho sistema se empleó para el análisis de un corpus de frases aisladas del español que contenían únicamente segmentos sonoros. El proceso de estilización propuesto comprendía un primer paso de extracción de una serie de valores de la curva de F_0 en puntos determinados (valor temporal y de F_0 al inicio de la curva, en cada uno de los puntos de inflexión y al final de la curva) y la posterior unión de los puntos obtenidos mediante líneas rectas.

El sistema MOMEL (*MOdelling MELody*) desarrollado en el *Laboratoire Parole et Langage* de la *Université de Provence* (Hirst y Espesser, 1993; Baqué y Estruch, 2003) ofreció en su momento una alternativa automática a estos sistemas manuales. El algoritmo MOMEL (desarrollado inicialmente como parte del programa 'MES', integrado posteriormente en el programa comercial 'Phonédit', y disponible recientemente como *script* para el programa 'Praat'⁷) determina los puntos de inflexión por medio de un cálculo de regresión que detecta los puntos de inflexión de la curva de F_0 y elimina las variaciones irrelevantes. Una función cuadrática (*quadratic spline function*) une los puntos de inflexión mediante parábolas. Con este método se pretende obtener curvas estilizadas equivalentes perceptivamente a las originales, aunque es necesaria una validación mediante síntesis y, eventualmente, la corrección manual de ciertos puntos de la curva.

MOMEL fue utilizado en el proyecto MULTEXT (*Multilingual Text Tools and Corpora*) (Hirst, Ide y Véronis, 1994) para realizar la anotación prosódica del corpus EUROM1 en varias lenguas europeas, entre ellas el español y el catalán (Campione y Véronis, 1998, 2000; Campione, Hirst y Véronis, 2000; Estruch, 2000). En el marco de este proyecto se realizó una validación perceptiva del sistema (Astésano, Espesser, Hirst, y Llisterri, 1997; Llisterri (Ed.), 1996), comparando las curvas melódicas de los enunciados de 40 párrafos estilizadas automáticamente con las correspondientes curvas originales. En el caso del corpus en español, el porcentaje de error (curvas obtenidas mediante MOMEL, diferentes perceptivamente del original) fue del 4,41% y en el caso del catalán, del 4,37%.

Un ejemplo de aplicación de MOMEL a un enunciado en español puede observarse en la Figura 7.

⁷ C. Auran, MOMEL-INTSINT, Laboratoire Parole et Langage, Université de Provence: <http://aune.lpl.univ-aix.fr/~auran/english/ressources.html>; G. Rolland, Automatic stylisation of the fundamental frequency (F_0) using MOMEL, Institut de la Communication Parlée, Grenoble: http://www.icp.inpg.fr/~loeven/Praat/momel_english.html

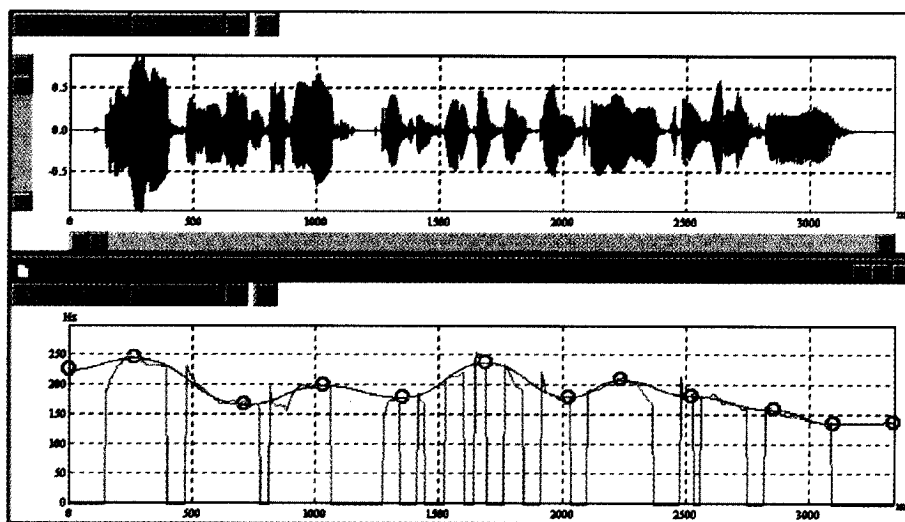


Figura 7. Oscilograma (ventana superior) y curva melódica y estilizada (ventana inferior) correspondientes al enunciado “Una vez en el interior, los agentes encontraron a tres de los niños”, pronunciado por un locutor femenino.

Una de las ventajas de este sistema es que elimina las rupturas de la curva producidas por la aparición de segmentos sordos; es decir, que la curva de F_0 no presenta ningún tipo de interrupción. Sin embargo, lo que por una parte constituye una ventaja, es también un inconveniente en el sentido de que no se encuentran interrupciones en los lugares del enunciado en que el hablante ha realizado una pausa. Es justamente en estas posiciones donde MOMEL efectúa más errores, al no detectar siempre el último punto de inflexión antes de la pausa (especialmente en puntos con valores altos de F_0) y el punto situado justo a continuación de la pausa.

Otro sistema automático de estilización que se ha empleado en algunos estudios es el integrado en el programa ‘Pitch’, una aplicación para Windows desarrollada en *Enginyeria La Salle* de la *Universitat Ramon Llull* (Martínez, 1995) en colaboración con el *Departament de Filologia Espanyola* de la *Universitat Autònoma de Barcelona*, que permitía la obtención, estilización y modificación de la curva melódica de los enunciados, así como su posterior síntesis para la evaluación perceptiva de las estilizaciones obtenidas. El algoritmo de estilización determinaba automáticamente los puntos de inflexión en función de un ‘umbral de semejanza’ establecido previamente por el usuario, de manera que éste podía definir hasta qué punto la curva estilizada obtenida se asemejaría a la curva original. Este sistema fue utilizado, entre otros, por Estruch, Mimó, Renom y Riera (1995) para el catalán y Garrido (1996, 2001) para el español.

Una de las ventajas del sistema de 'Pitch' era la posibilidad de que el usuario determinase el grado de semejanza de la curva estilizada con la original. Para definir el umbral de semejanza que debía emplearse para obtener curvas estilizadas idénticas perceptivamente a las correspondientes curvas originales, se realizaron dos pruebas de percepción, descritas en Jiménez (1994), en Estruch *et al.* (1995) y en Estruch *et al.* (1999). Mediante el umbral definido en estas pruebas es posible obtener curvas equivalentes perceptivamente a las originales sin necesidad de corrección manual.

Un método semejante de estilización se incluye actualmente como parte del programa 'Praat' (Boersma y Weenink, 2007). El sistema permite ajustar el nivel de semejanza entre la curva original y la estilizada gracias a un 'factor de resolución'. También permite determinar el tipo de unidades en que se basará la estilización (hercios o semitonos). La Figura 8 presenta la curva melódica del enunciado de referencia estilizada mediante 'Praat'.

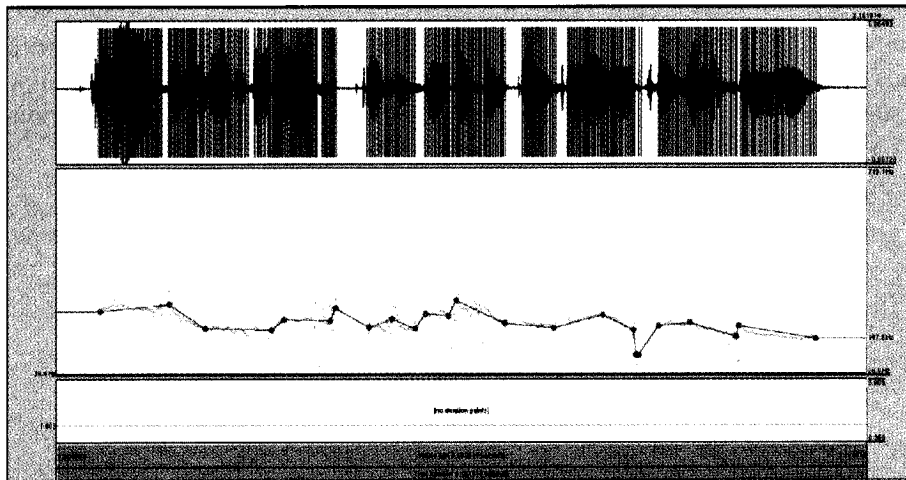


Figura 8. Oscilograma, curva melódica original y contorno estilizado, obtenidos mediante el programa 'Praat', del enunciado "Una vez en el interior, los agentes encontraron a tres de los niños", pronunciado por un locutor femenino.

3.4. La normalización

El proceso de estilización elimina sólo una parte de las variaciones, en principio irrelevantes, que se encuentran presentes en las curvas melódicas. Aun así, se mantienen otras, como las resultantes de la frecuencia fundamental propia del locutor, o las derivadas de la duración del enunciado. En el presente apartado se describen algunos de los procedimientos utilizados para neutralizar estas fuentes de variación.

En Garrido (1991a, b) se proponía un sistema para representar los valores frecuenciales de la curva estilizada de manera independiente del locutor. Este sistema consistía en normalizar los valores de los diferentes puntos de inflexión de la curva estilizada, restándole el valor de F_0 del primer punto que se encuentra en la curva. De esta manera, el valor inicial se convierte en punto de referencia (valor 0) y los valores de los demás puntos se expresan en término de diferencias frecuenciales con respecto a este punto.

En el procedimiento descrito en Cantero (1995, 2002), la altura (ascenso o descenso) se cuantifica por medio de porcentajes con respecto al total del rango del locutor. Un sistema similar fue adoptado por De la Mota (1995) para cuantificar la pendiente de las líneas que unen los diferentes puntos de inflexión, y se aplicó en el análisis de un corpus del catalán procedente de los medios de comunicación en Font (2007).

Las diferencias en la F_0 inherentes a cada hablante pueden neutralizarse también mediante el uso de una escala logarítmica para la representación de los valores frecuenciales. Así, por ejemplo, el rango frecuencial de la curva correspondiente a un mismo enunciado realizado por dos locutores distintos (uno masculino y otro femenino) se normaliza si en lugar de emplear una escala lineal en Hz se recurre a una escala logarítmica, tal como se observa en la Figura 9. Por otro lado, el uso de escalas logarítmicas refleja más fielmente el proceso de percepción de la melodía (T'Hart *et al.*, 1990).

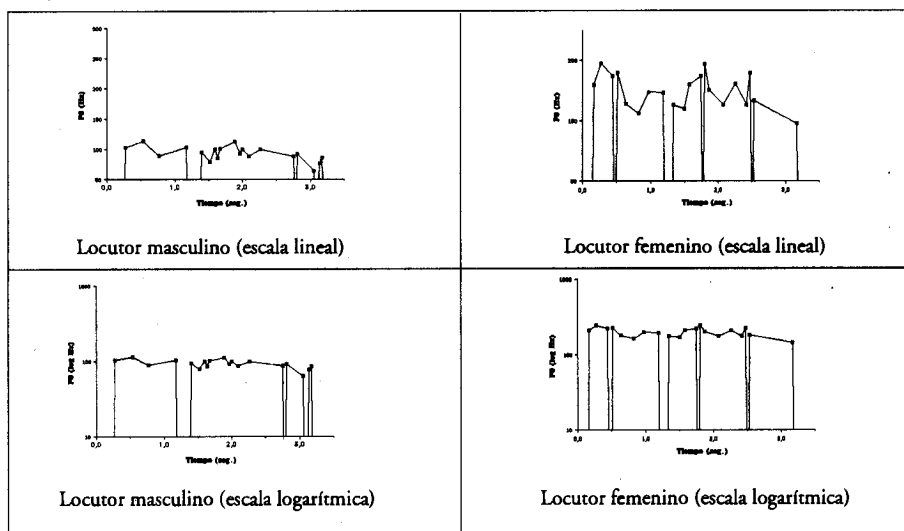


Figura 9. Curvas estilizadas correspondientes al enunciado “Una vez en el interior, los agentes encontraron a tres de los niños” pronunciadas por un locutor masculino (izquierda) y otro femenino (derecha), representadas utilizando una escala frecuencial lineal (arriba) y logarítmica (abajo).

Existen distintos sistemas que utilizan escalas logarítmicas como, por ejemplo, la representación en términos de octavas (Navarro, 1944) o en términos de semitonos (T'Hart *et al.*, 1990).

4. REPRESENTACION SIMBOLICA DE LAS CURVAS MELODICAS: LA ANOTACION⁸

Si por medio de la estilización se obtiene una versión simplificada de las curvas melódicas en la que se mantiene la información lingüísticamente pertinente, la anotación permite la representación de los movimientos de las mismas mediante un conjunto de símbolos convencionales. La anotación constituye una alternativa a la estilización, puesto que puede realizarse directamente sobre las curvas melódicas originales, o incluso sin necesidad de llevar a cabo un análisis acústico. Sin embargo, también puede concebirse como un paso más, posterior a la estilización, en el proceso de abstracción desde la señal acústica hasta la representación fonológica de la entonación, en el que se eliminan los últimos restos de información irrelevante desde el punto de vista lingüístico.

Existe actualmente una gran cantidad de sistemas de anotación, desarrollados con fines muy diversos y desde perspectivas teóricas muy diferentes⁹, por lo que resulta imposible realizar aquí un análisis exhaustivo de todos ellos. Por ejemplo, no se abordan en este trabajo los procedimientos basados en el uso de símbolos fonéticos como SAMPROSA (Wells, 1995) y el propio Alfabeto Fonético Internacional (Bruce, 1988; IPA, 1999) ni tampoco los sistemas de anotación prosódica propios del análisis del discurso y de la conversación (Edwards y Lampert (Eds.), 1993; Payrató, 1995) o de la lingüística de corpus (Leech, Myers y Thomas (Eds.), 1995). Una presentación más detallada de los principales métodos de anotación se encuentra, entre otros, en Llisterri (1994, 1996, 1999), Klein *et al.* (1998), Dybkjaer *et al.* (2001) y en Cosi (2002). Llisterri, Machuca, De la Mota, Riera y Ríos (2005b) y Albelda (2005) se centran específicamente en el español.

Los diferentes procedimientos pueden clasificarse en dos grandes grupos: los que pretenden representar la forma fonética de la curva melódica (que denominaremos aquí 'sistemas de anotación fonética'), y los que pretenden representar la forma fonológica subyacente a la curva melódica (que denominaremos, en el presente trabajo, 'sistemas de anotación fonológica').

⁸ 'Anotación' y 'transcripción' son dos términos que tradicionalmente se han usado de manera indiferenciada para referirse al mismo concepto. Sin embargo, pueden utilizarse para distinguir las representaciones simbólicas que se usan para etiquetar una curva melódica ('anotación') de las que se utilizan para representar la forma sonora de la lengua ('transcripción'). En este trabajo utilizaremos al término 'anotación' para referirnos indistintamente a ambos tipos de etiquetado.

⁹ Véase, por ejemplo, Prieto (Ed.) (2003).

4.1. Anotación fonética de la melodía

Como ya se ha señalado en trabajos anteriores (Quilis 1981, 1993, por ejemplo) existen dos aproximaciones fundamentales al estudio de la entonación: la primera considera los contornos melódicos como una serie de niveles tonales (análisis por niveles) (Martínez Celdrán, 2003), en tanto que la segunda considera las curvas melódicas como el resultado de la concatenación de una serie de segmentos (análisis por contornos) (García Lecumberri, 2003). Teniendo en cuenta esta clasificación, los diferentes sistemas de anotación fonética pueden dividirse en dos grupos: los que intentan representar simbólicamente la altura tonal en determinados puntos de la curva –niveles tonales, puntos de inflexión– y los que tratan de representar mediante símbolos los segmentos o contornos que componen la curva melódica. En los apartados siguientes presentamos por separado ambos tipos de anotación.

4.1.1. Anotación de segmentos

Ciertos sistemas intentan representar mediante símbolos la forma de los contornos que presentan las curvas melódicas. En la tradición de los estudios entonativos españoles, la anotación de contornos se ha limitado normalmente a los movimientos finales ('tonemas' o 'junturas terminales', según los diferentes autores). Ejemplos de este tipo de representación son los que emplean Navarro (1944) y Quilis (1981, 1993), que se ilustran en la Figura 10.

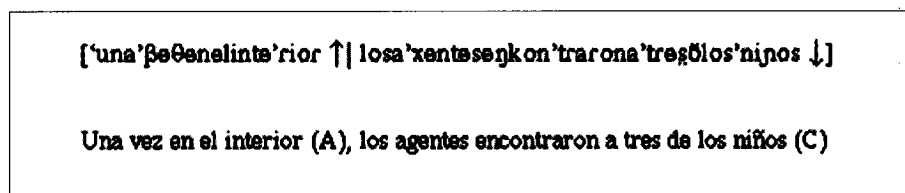


Figura 10. Representación simbólica según el sistema de Quilis (1981,1993) (parte superior) y de Navarro (1944) (parte inferior) de los movimientos finales de la entonación del enunciado “Una vez en el interior, los agentes encontraron a tres de los niños”.

Aunque no se ha aplicado al análisis de la entonación del español, el sistema de anotación desarrollado por el IPO para la descripción de los contornos melódicos del holandés (T’Hart *et al.*, 1990; Garrido, 2003) es quizá el mejor ejemplo del tipo de anotación al que nos estamos refiriendo. Este método se emplea para eti-

quetar los diferentes segmentos (líneas rectas que unen dos puntos de inflexión) de una curva previamente estilizada mediante el procedimiento de la *close-copy* al que nos hemos referido en el apartado 3.3.2. En la Figura 11 puede observarse el enunciado de referencia transcrito mediante este sistema.

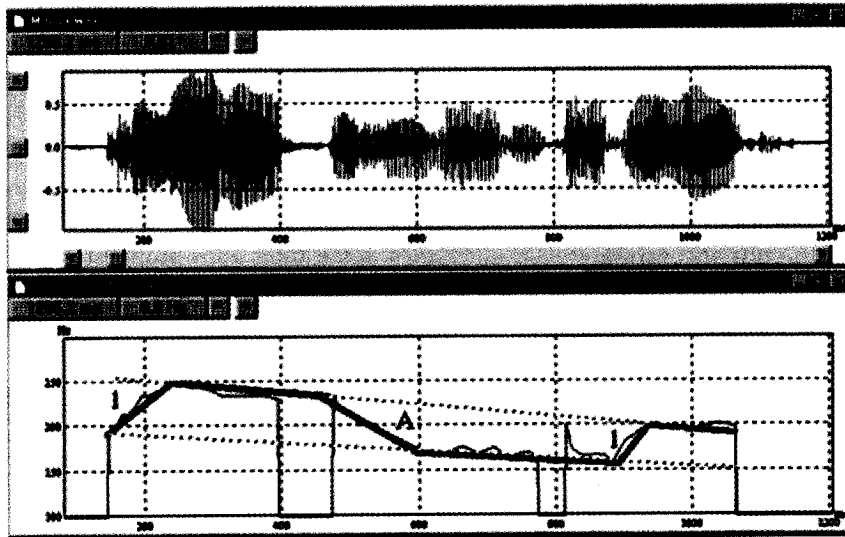


Figura 11. Anotación de la curva estandarizada correspondiente al enunciado “Una vez en el interior”, según el método descrito en T’Hart *et al.* (1990).

El sistema del IPO tiene en cuenta no sólo la altura tonal que alcanza el movimiento (longitud media o completa), sino también su dirección (ascendente o descendente), el alineamiento temporal con el enunciado y la velocidad del cambio (rápido o lento). Se asume que los segmentos se superponen a una trama de tres líneas de declinación paralelas.

4.1.2. Anotación de la altura tonal

A diferencia de los anteriores, los sistemas de anotación que intentan codificar los niveles de altura tonal de la curva melódica a lo largo de un enunciado tienen ya cierta tradición en los estudios entonativos del español. Así, Quilis (1981, 1993) emplea un procedimiento de anotación por niveles –complementario al sistema de anotación por segmentos descrito en el apartado anterior– que considera la existencia de tres niveles tonales: alto, medio y bajo. Cada nivel está representado por

un número diferente (3, 2 y 1, respectivamente) asociado a cada una de las sílabas que componen el enunciado. Un sistema parecido se utiliza en Alcina y Blecua (1975).

Por su parte, Fant (1984) utiliza un método semejante, que emplea en este caso un conjunto de letras: A+ (muy alto), A (alto), M (medio) y B (bajo) para codificar cuatro niveles tonales distintos. En este caso, a diferencia del sistema empleado por Quilis, los niveles tonales no se asocian a la sílaba entera, sino a un punto de inflexión, que normalmente suele coincidir con el final de una sílaba.

Más recientemente, se han desarrollado otros métodos para la anotación en niveles de las curvas del español. Así, por ejemplo, en Garrido (1996, 2001) se propone un procedimiento de anotación de las curvas melódicas que utiliza tres símbolos diferentes, P (Pico), M (Medio) y V (Valle). El sistema asume que el rango tonal de la curva melódica a lo largo de un grupo entonativo puede describirse mediante tres niveles tonales diferentes, semejantes a los propuestos por el IPO en su descripción del holandés (T'Hart *et al.*, 1990), y que los puntos de inflexión relevantes en un contorno melódico se sitúan en uno de estos niveles. La representación de cada uno de estos tres niveles teóricos a lo largo del grupo entonativo se lleva a cabo por medio de tres líneas de declinación (superior, media e inferior), que tienen en consideración la tendencia observada en las curvas melódicas al descenso gradual de la F_0 a lo largo de las mismas, como ilustra la Figura 12. De esta manera se consigue representar la altura de cada punto no ya con un valor determinado de F_0 , sino con un sistema que, en principio, es independiente del locutor. Sin embargo, el método propuesto no pretende en absoluto representar la estructura fonológica subyacente a la curva entonativa.

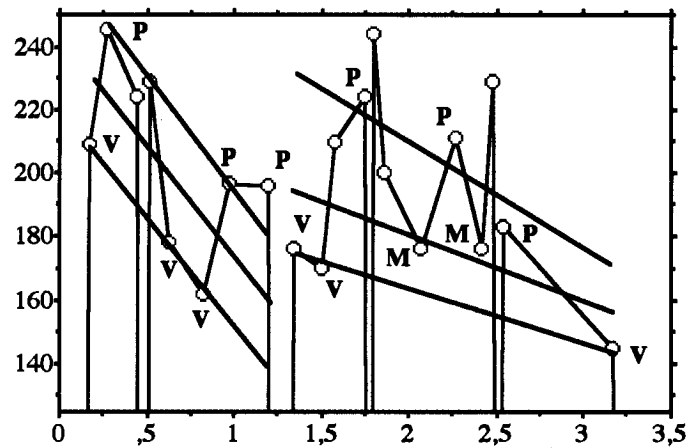


Figura 12. Curva estilizada y anotada según el sistema de Garrido (1996, 2001) correspondiente a la oración “Una vez en el interior, los agentes encontraron a tres de los niños”, pronunciada por un locutor femenino.

El sistema de anotación INTSINT (*IN*ternational *T*ranscription *S*ystem for *IN*Tonation), desarrollado en la *Université de Provence* (Hirst y Di Cristo, 1998; Campione *et al.*, 2000; Hirst, Di Cristo y Espesser, 2000; Di Cristo, Hirst, Boudouresques, y Louis, 2002; Baqué y Estruch, 2003; Hirst, 2005), toma como punto de partida la curva estilizada, obtenida con la aplicación de MOMEL (descrito en el apartado 2.3.2), y etiqueta cada uno de los puntos de inflexión establecidos mediante un símbolo. Las etiquetas pueden asignarse automáticamente a partir de la curva estilizada por medio del programa ‘MES’ o, como se ha señalado anteriormente, a partir de los ‘scripts’ para ‘Praat’ mencionados en la nota 7. Los símbolos empleados se clasifican en absolutos –*Top*, *Mid* y *Bottom*, definidos de manera global con respecto al rango tonal del locutor analizado– y relativos –*High*, *Low*, *Upstepped*, *Downstepped* y *Same*, definidos localmente, en función de los puntos anteriores y posteriores–. A su vez, los símbolos relativos se dividen en iterativos (pueden ir seguidos de un tono idéntico) y no iterativos (siempre van seguidos de un tono diferente). La Figura 13 presenta un ejemplo de curva anotada mediante este sistema.

INTSINT se ha utilizado, como ya hemos señalado, para la anotación prosódica del corpus EUROM1 en español (Campione y Véronis, 1998, 2000; Campione *et al.*, 2000) y de la versión correspondiente en catalán (Estruch, 2000).

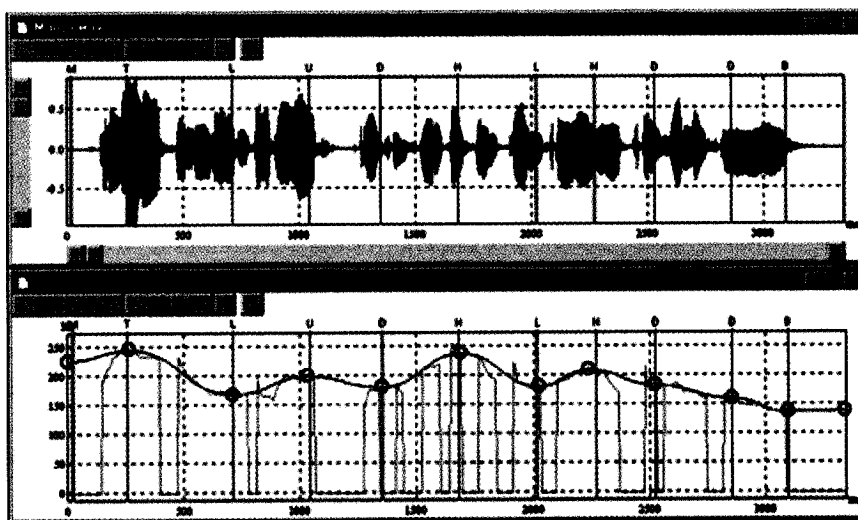


Figura 13. Oscilograma (ventana superior), curva melódica original, curva estilizada (sistema MOMEL) y anotación INTSINT (ventana inferior) correspondientes al enunciado “Una vez en el interior, los agentes encontraron a tres de los niños”, pronunciado por un locutor femenino.

Al igual que el sistema de Garrido (1996, 2001), INTSINT no pretende representar la forma fonológica de la entonación, sino que se concibe más bien como un procedimiento de representación formal de las curvas melódicas. Sin embargo, a diferencia de los métodos anteriores, INTSINT no determina los niveles tomando como unidad de base el grupo entonativo, sino la totalidad de un enunciado.

Independientemente del sistema utilizado, los diversos procedimientos de anotación tienen como objetivo común la representación por medio de símbolos de la altura tonal de determinados puntos de la curva melódica. Las diferencias estriban, esencialmente, en el establecimiento del número de niveles y en la determinación del ámbito adecuado –párrafos, oraciones, grupos entonativos– para la definición de los mismos.

4.2. Anotación fonológica

Actualmente, uno de los principales procedimientos de anotación de curvas melódicas con una orientación claramente fonológica es el conocido como ToBI (*TONes and Break Indices*) (Silverman *et al.*, 1992; Beckman y Ayers, 1997; Beckman, Hirschberg y Shattuck-Hufnagel, 2005). Este sistema, desarrollado inicialmente para el inglés americano, tiene la particularidad de que no aspira a describir la forma fonética de la entonación, sino su representación subyacente o fonológica. ToBI incluye dos tipos de símbolos: los que representan la estructura tonal subyacente (*Tones*), y los que marcan los límites entre unidades prosódicas (*Break Indices*). Este procedimiento se ha aplicado también a la descripción de los contornos melódicos del español (Sosa, 2003) y del catalán (Prieto *et al.*, 2007).

El sistema recoge el inventario de unidades entonativas propuestas en el modelo de Pierrehumbert (1980), también conocido como modelo métrico y autosegmental (Hualde, 2003), y considera tres tipos diferentes de tonos: los llamados ‘acentos tonales’ (*Pitch Accents*) o tonos asociados con sílabas acentuadas, los ‘tonos de límite’ (*Boundary Tones*) o tonos que marcarían el inicio o el final de una frase entonativa, y los denominados ‘acentos de frase’ (*Phrase Accents*) o tonos que aparecerían justo antes de un tono de límite final y que señalan, de acuerdo con el modelo de Pierrehumbert, el final de una frase intermedia. El sistema incluye también símbolos para la anotación de los límites de unidades prosódicas o ‘indicadores de límite’ (*Break Indices*). Se incluyen símbolos para cuatro tipos de unidades: el grupo clítico, la palabra (fonológica), la frase intermedia y la frase entonativa. La Figura 14 presenta un ejemplo de curva anotada mediante ToBI.

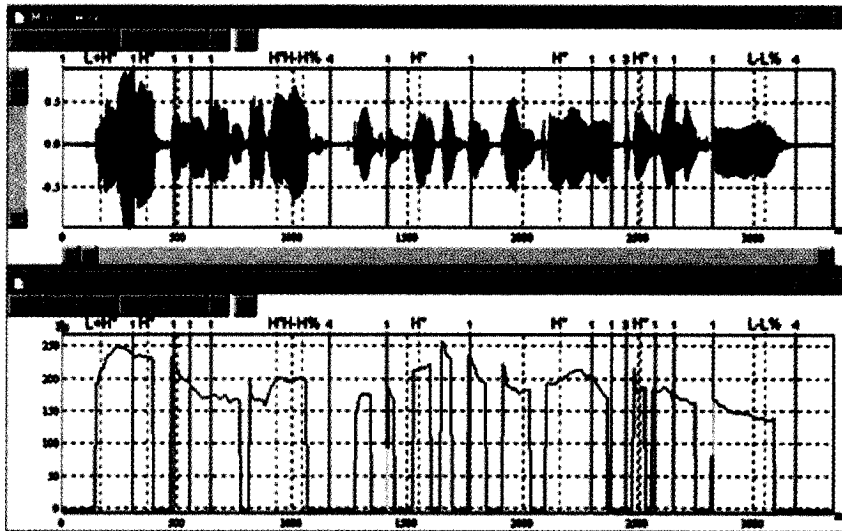


Figura 14. Oscilograma, curva melódica y anotación mediante el sistema ToBI del enunciado “Una vez en el interior, los agentes encontraron a tres de los niños”, pronunciado por un locutor femenino.

Independientemente del hecho de que ToBI sea un sistema fuertemente influido por la teoría a partir de la que se ha desarrollado, uno de sus aciertos es ofrecer un procedimiento separado para la anotación de dos aspectos diferentes de la representación fonológica de la entonación: por un lado, las unidades prosódicas en que se organizan los enunciados y, por otro, la representación fonológica del fenómeno de la entonación, en este caso en términos de tonos, que enlaza con los trabajos en fonología prosódica.

5. CONCLUSION

A lo largo de este trabajo, se han revisado algunos de los diferentes sistemas de representación de la entonación, más o menos ligados a la forma fonética de la misma, que tiene su manifestación en la curva melódica, siguiendo la aproximación esbozada por Estruch, Garrido, Llisterri y Riera (1996). De esta presentación se deduce que no parece existir hoy en día un procedimiento que pueda considerarse estándar –equivalente al Alfabeto Fonético Internacional (IPA, 1999) o a SAMPA (Wells, 2003) en lo que se refiere a la transcripción de los elementos segmentales– que permita representar mediante un conjunto discreto de símbolos

la forma de las curvas melódicas. En todo caso, parece claro que ésta puede realizarse a distintos niveles, según el grado de abstracción o la orientación teórica del investigador. Independientemente del sistema, cabe definir tres niveles de representación:

- a) un nivel de representación fonética de la melodía, que incluiría:
 - la representación de la curva melódica original
 - la representación estilizada de la curva melódica
 - la representación simbólica de esa curva melódica
- b) un nivel de representación fonológica de la melodía
- c) un nivel de representación de la jerarquía de unidades prosódicas.

La complejidad del estudio de los fenómenos asociados a las variaciones melódicas en el habla requiere, como se ha intentado poner en relieve, una clara distinción metodológica entre el nivel fonético y el fonológico, así como una serie de procedimientos que permitan, mediante aproximaciones sucesivas con un grado cada vez más elevado de abstracción, representar la información fonéticamente relevante que se manifiesta en la evolución temporal de la frecuencia fundamental. Tal es, precisamente, el objetivo de las técnicas y sistemas expuestos en el presente trabajo.

REFERENCIAS

- Albelda, M. 2005. "Sistemas de transcripción de los corpus orales del español". In M. Carrió (Ed.), *Perspectivas interdisciplinarias de la lingüística aplicada* (pp. 381-387). València: Universitat Politècnica de València - AESLA, Asociación Española de Lingüística Aplicada.
- Alcina, J. & Blecua, J. M. 1975. *Gramática española*. Barcelona: Ariel.
- Astésano, C., Espesser, R., Hirst, D. & Llisterri, J. 1997. "Stylisation automatique de la fréquence fondamentale: une évaluation multilingue". In *Actes du 4e Congrès Français d'Acoustique* (pp. 441-443). Consultado el 24 de junio de 2007 en http://liceu.uab.es/~joaquim/publicacions/Astesano_et_al_97.pdf
- Baqué, L. & Estruch, M. 2003. "Modelo de Aix-en-Provence". In P. Prieto (Ed.), *Teorías de la entonación* (pp. 123-154). Barcelona: Ariel. Consultado el 24 de junio de 2007 en http://seneca.uab.es/lorraine_baque/Publicacions/ModeloAix-en-ProvenceV3.pdf
- Beckman, M. E. & Ayers, G. M. 1997. *Guidelines for ToBI Labelling*. Version 3. Consultado el 24 de junio de 2007 en http://ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf
- Beckman, M. E., Hirschberg, J. & Shattuck-Hufnagel, S. 2005. "The original ToBI system and the evolution of the ToBI framework". In S.-A. Jun (Ed.), *Prosodic*

- Typology. The Phonology of Intonation and Phrasing* (pp. 9-54). Oxford: Oxford University Press. Consultado el 24 de junio de 2007 en <http://www.ling.ohio-state.edu/~tobi/JunBook/BeckHirschShattuckToBI.pdf>
- Boersma, P. & Weenink, D. 2007. Praat: doing phonetics by computer (Version 4.6.09) [Programa informático]. Consultado el 24 de junio de 2007 en <http://www.praat.org/>
- Bruce, G. 1988. "2.3. Suprasegmental categories and 2.4. The symbolization of temporal events". *Journal of the International Phonetic Association*, 18(2), 75-76.
- Buenafuentes, C., Madrigal, N. & Garrido, J. M. 2000. "Análisis acústico de las variaciones micromelódicas en las curvas del F0 del español". *Español Actual*, 73, 65-77.
- Campione, E. & Véronis, J. 1998. "A Multilingual Prosodic Database". In *ICSLP 98, Proceedings of the 5th International Conference on Spoken Language Processing* (pp. 3163-3166). Consultado el 24 de junio de 2007 en <http://www.up.univ-mrs.fr/~veronis/pdf/1998icslp-database.pdf>
- Campione, E. & Véronis, J. 2000. "Une évaluation de l'algorithme de stylisation mélodique MOMEL". *TIPA, Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix*, 19, 27-44. Consultado el 24 de junio de 2007 en <http://aune.lpl.univ-aix.fr/lpl/tipa/19/tipa19-campione.pdf>
- Campione, E.; Hirst, D. & Véronis, J. 2000. "Automatic stylisation and symbolic coding of F₀. Implementations of the INTSINT model". In A. Botinis (Ed.), *Intonation: Analysis, Modelling and Technology* (pp. 185-208). Dordrecht: Kluwer. Consultado el 24 de junio de 2007 en <http://www.up.univ-mrs.fr/~veronis/pdf/2000Campione.pdf>
- Cantero, F. J. 1995. *Estructura de los modelos entonativos: interpretación fonológica del acento y la entonación en castellano*. Tesis doctoral no publicada, Universitat de Barcelona.
- Cantero, F. J. 2002. *Teoría y análisis de la entonación*. Barcelona: Publicacions de la Universitat de Barcelona.
- Cosi, P. 2002. "Metodologie e sistemi per l'annotazione linguistica". *Quaderni dell'Istituto di Fonetica e Dialettologia*, 4. Consultado el 24 de junio de 2007 en <http://www.pd.istc.cnr.it/Papers/quaderni2002.zip>
- D'Alessandro, C. & Mertens, P. 1995. "Automatic pitch contour stylization using a model of tonal perception". *Computer Speech & Language*, 9, 257-288.
- De la Mota, C. 1995. *La representación gramatical de la información nueva en el discurso*. Tesis doctoral no publicada, Universitat Autònoma de Barcelona.
- Di Cristo, A. 1982. *Prolegomènes à l'étude de l'intonation. Micromélie*. Paris: Éditions du CNRS.
- Di Cristo, A.; Hirst, D.; Boudouresques, N. & Louis, M. 2002. "Écrire l'intonation: le système INTSINT, fondements théoriques et illustrations". *Revue Parole*, 22-23-24, 175-212.
- Dybkjaer, L., Berman, S., Kipp, M., Wegener Olsen, M., Pirrelli, V., Reithinger, N. et al. 2001. *Survey of existing tools, standards and user needs for annotation of natural interaction and multimodal data*. (Deliverable D11. 1. January 2001). ISLE Natural Interactivity and Multimodality Working Group. Consultado el 24 de junio de 2007 en <http://isle.nis.sdu.dk/reports/wp11/D11.1-14.2.2001.pdf>
- Edwards, J. A. & Lampert, M. D. (Eds.). 1993. *Talking Data: Transcription and Coding in Discourse Research*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Estruch, M. 2000. "Évaluation de l'algorithme de stylisation mélodique MOMEL et du système de codage symbolique INTSINT avec un corpus de passages en catalan". *TIPA, Travaux Interdisciplinaires du laboratoire Parole et Langage d'Aix-en-Provence*, 19, 45-61. Consultado el 24 de junio de 2007 en <http://www.lpl.univ-aix.fr/lpl/tipa/19/tipa19-estruch.pdf>
- Estruch, M. & Garrido, J. M. 1995. *Análisis y clasificación de los contornos melódicos finales en un corpus de frases aisladas del español*. Comunicación presentada en el XXV Simposio de la Sociedad Española de Lingüística, Zaragoza. Resumen publicado en *Revista Española de Lingüística* 26(1), 138-139.
- Estruch, M.; Garrido, J. M.; Gudayol, F.; Jiménez, J. M.; Renom, J. & Riera, M. 1999. "Validación perceptiva de un sistema de estilización automática de contornos melódicos". In *Actes del I Congrés de Fonètica Experimental* (pp. 217-223). Tarragona: Universitat Rovira y Virgili.
- Estruch, M.; Garrido, J. M.; Llisterri, J. & Riera, M. 1996. "Una aproximación fonética al estudio de la entonación". *Philologia Hispalensis*, 11, 281-293. Consultado el 24 de junio de 2007 en http://liceu.uab.cat/~joaquim/publicacions/Sevilla_96.pdf
- Estruch, M.; Mimó, M.; Renom, J. & Riera, M. 1995. *Validació perceptiva dels patrons melòdics del català*. Manuscrito no publicado, Universitat Autònoma de Barcelona.
- Fant, L. 1984. *Estructura informativa en español. Estudio sintáctico y entonativo*. Upsala: Acta Universitatis Upsaliensis.
- Font, D. 2007. *L'entonació del català*. Barcelona: Publicacions de l'Abadia de Montserrat.
- García Lecumberri, M. L. 2003. "Análisis por configuraciones: la escuela británica". In P. Prieto (Ed.), *Teorías de la entonación* (pp. 35-62). Barcelona: Ariel.
- Garrido, J. M. 1991a. *Modelización de patrones melódicos del español para la síntesis y el reconocimiento de habla*. Bellaterra: Universitat Autònoma de Barcelona.
- Garrido, J. M. 1991b. "Modelización de patrones melódicos del español para sistemas de conversión texto-habla". *Procesamiento del Lenguaje Natural*, 11, 209-219. Consultado el 24 de junio de 2007 en <http://www.sepln.org/revistaSEPLN/revista/11/11-Pag197.pdf>
- Garrido, J. M. 1996. *Modelling Spanish Intonation for Text-to-Speech Applications*. Tesis doctoral no publicada, Universitat Autònoma de Barcelona. Consultado el 24 de junio de 2007 en <http://liceu.uab.es/juanma/tesis.html>
- Garrido, J. M. 2001. "La estructura de las curvas melódicas del español: propuesta de modelización". *Lingüística Española Actual*, 23(2), 173-209.
- Garrido, J. M. 2003. "La escuela holandesa: el modelo IPO". In P. Prieto (Ed.), *Teorías de la entonación* (pp. 97-122). Barcelona: Ariel.
- Garrido, J. M.; Llisterri, J.; De la Mota, C. & Ríos, A. 1993. "Prosodic differences in reading style: Isolated vs. Contextualized Sentences". In *Eurospeech 93, 3rd European Conference on Speech Communication and Technology* (pp. 573-576). Consultado el 24 de junio de 2007 en http://liceu.uab.cat/~joaquim/publicacions/Berlin_93/berlin_93.html
- Garrido, J. M.; Llisterri, J.; De la Mota, C. & Ríos, A. 1995. "Estudio comparado de las características prosódicas de la oración simple en español en dos modalidades de lectura". In A. Elejabeitia, & A. Iribar (Eds.), *Phonetica. Trabajos de fonética experimental* (pp. 173-194). Bilbao: Universidad de Deusto. Consultado el 24 de junio de 2007 en <http://liceu.uab.cat/~joaquim/publicacions/Deusto95.pdf>

- Garrido, J. M.; Llisterri, J.; Marín, R.; De la Mota, C. & Ríos, A. 1995. "Prosodic markers at syntactic boundaries in Spanish". In *ICPhS 95, Proceedings of the XIIIth International Congress of Phonetic Sciences* (pp. 370-373). Consultado el 24 de junio de 2007 en http://liceu.uab.cat/~joaquim/publicacions/Stockholm_95/stockholm_95.html
- Gerhard, D. 2003. *Pitch Extraction and Fundamental Frequency: History and Current Techniques* (Technical Report TR-CS 2003-06). Regina, Saskatchewan: Department of Computer Science, University of Regina. Consultado el 24 de junio de 2007 en <http://www.cs.uregina.ca/Research/Techreports/2003-06.pdf>
- Gili Gaya, S. 1924. "Influencia del acento y de las consonantes en las curvas de entonación". *Revista de Filología Española*, 11, 154-177.
- Hess, W. 1983. *Pitch Determination of Speech Signals: Algorithms and Devices*. Berlin: Springer.
- Hirst, D. J. 2005. "Form and function in the representation of speech prosody". *Speech Communication*, 46, 334-347.
- Hirst, D. J. & Di Cristo, A. 1998. "A survey of Intonation Systems". In D. Hirst & A. Di Cristo (Eds.), *Intonation Systems. A Survey of Twenty Languages* (pp. 1-44). Cambridge: Cambridge University Press. Consultado el 24 de junio de 2007 en <http://www.lpl.univ-aix.fr/~hirst/articles/1998%20Hirst&DiCristo.pdf>
- Hirst, D. J. & Espesser, R. 1993. "Automatic modelling of fundamental frequency using a quadratic spline function". *TIPA, Travaux de l'Institut de Phonétique d'Aix*, 15, 75-85. Consultado el 24 de junio de 2007 en <http://www.lpl.univ-aix.fr/~hirst/articles/1993%20Hirst&Espesser.pdf>
- Hirst, D. J.; Di Cristo, A. & Espesser, R. 2000. "Levels of representation and levels of analysis for the description of intonation systems". In M. Horne (Ed.), *Prosody: Theory and Experiment. Studies presented to Gösta Bruce* (pp. 51-88). Dordrecht: Kluwer. Consultado el 24 de junio de 2007 en <http://www.lpl.univ-aix.fr/~hirst/articles/2000%20Hirst&al.pdf>
- Hirst, D. J., Ide, N. & Véronis, J. 1994. "Coding fundamental frequency patterns for multi-lingual synthesis with INTSINT in the MULTTEXT project". In *Conference Proceedings of the Second ESCA/IEEE Workshop on Speech Synthesis* (pp. 77-80).
- Hualde, J. I. 2003. "El modelo métrico y autosegmental". In P. Prieto (Ed.), *Teorías de la entonación* (pp. 155-184). Barcelona: Ariel.
- International Phonetic Association. (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- Jiménez, J. M. 1994. *Implementació d'un mètode d'estilitzat de corbes melòdiques*. Manuscrito no publicado, Enginyeria La Salle, Universitat Ramon Llull, Barcelona.
- Klein, M.; Bernsen, N. O.; Davies, S.; Dybkjaer, L.; Garrido, J. M.; Kasch, H. et al. 1998. *Supported Coding Schemes* (Deliverable D1.1). LE Telematics Project LE4-8370 MATE. Consultado el 24 de junio de 2007 en <http://mate.nis.sdu.dk/about/D1.1/>
- Leech, G.; Myers, G. & Thomas, J. (Eds.). (1995). *Spoken English on Computer: Transcription, Markup and Applications*. Harlow: Longman.
- Lehiste, I. & Peterson, G. E. 1961. "Some basic considerations in the analysis of intonation". *Journal of the Acoustical Society of America*, 33(4), 419-425.

- Lieberman, P. & Blumstein, S. E. 1988. *Speech Physiology, Speech Perception and Acoustic Phonetics*. Cambridge: Cambridge University Press.
- Llisterri, J. 1994. *Prosody Encoding Survey* (Deliverable 1.5.3). LRE Project 62-050 MULTEXT. Consultado el 24 de junio de 2007 en http://liceu.uab.cat/~joaquim/publicacions/Prosody_encoding_94.pdf
- Llisterri, J. 1999. "Transcripción, etiquetado y codificación de corpus orales". *Revista Española de Lingüística Aplicada* (Volumen Monográfico "Panorama de la Investigación en Lingüística Informática"), 53-82. Consultado el 24 de junio de 2007 en http://liceu.uab.es/~joaquim/publicacions/RESLA_99.pdf
- Llisterri, J. (Ed.). 1996. *Prosody Tools Efficiency and Failures* (Deliverable 4.5.2). LRE Project 62-050 MULTEXT. Consultado el 24 de junio de 2007 en http://liceu.uab.cat/~joaquim/publicacions/Prosody_tools_96.pdf
- Llisterri, J.; Machuca, M. J.; De la Mota, C.; Riera, M. & Ríos, A. 2005a. "La percepción del acento léxico en español". In *Filología y lingüística. Estudios ofrecidos a Antonio Quilis* (pp. 271-297). Madrid: Consejo Superior de Investigaciones Científicas - Universidad Nacional de Educación a Distancia - Universidad de Valladolid. Consultado el 24 de junio de 2007 en http://liceu.uab.es/~joaquim/publicacions/Llisterri_Machuca_Mota_Riera_Rios_05_Percepcion_Acento_Espanol.pdf
- Llisterri, J.; Machuca, M. J.; De la Mota, C.; Riera, M. & Ríos, A. 2005b. "Corpus orales para el desarrollo de las tecnologías del habla en español". *Oralia. Análisis del discurso oral*, 8, 289-325. Consultado el 24 de junio de 2007 en http://liceu.uab.cat/~joaquim/publicacions/Llisterri_Machuca_Mota_Riera_Rios_05_Corpus_Orales_Tecnologias_Habla_Espanol.pdf
- Martínez Celdrán, E. 2003. "Análisis por niveles: la escuela americana", in P. Prieto (Ed.), *Teorías de la entonación* (pp. 63-96). Barcelona: Ariel.
- Martínez, D. 1995. *Sistema d'anàlisi, tractament i síntesi de la melodia en entorn Windows*. Manuscrito no publicado, Ingeniería La Salle, Universitat Ramon Llull, Barcelona.
- Mateo, A. 1988. "Experimento sobre el tono intrínseco de las vocales castellanas". *Estudios de Fonética Experimental*, 3, 157-179.
- Mertens, P. 2004. "The Prosogram: Semi-Automatic Transcription of Prosody based on a Tonal Perception Model". In *Proceedings of Speech Prosody 2004* (pp. 549-552). Consultado el 24 de junio de 2007 en <http://bach.arts.kuleuven.be/pmertens/papers/sp2004.pdf>
- Navarro Tomás, T. 1944/1974. *Manual de entonación española*. (4ª edición) Madrid: Guadarrama.
- Orlikoff, R. F. & Kahane, J. C. 1996. "Structure and function of the larynx". In N. J. Lass (Ed.), *Principles of Experimental Phonetics* (pp. 112-184). St Louis, Mosby.
- Payrató, L. 1995. "Transcripción del discurso coloquial". In L. Cortés (Ed.), *El español coloquial. Actas del I Simposio sobre Análisis del Discurso Oral* (pp. 43-70). Almería: Servicio de Publicaciones de la Universidad de Almería.
- Peterson, G. E. & Barney, H. L. 1952. "Control methods used in a study of vowels". *Journal of the Acoustical Society of America*, 24, 175-184.
- Pierrehumbert, J. B. 1980/1987. *The Phonology and Phonetics of English Intonation*. Tesis doctoral no publicada, MIT, Cambridge MA. Bloomington: Indiana University Linguistics Club. Consultado el 24 de junio de 2007 en <http://dspace.mit.edu/handle/1721.1/16065>

- Prieto, P. (Ed.). 2003. *Teorías de la entonación*. Barcelona: Ariel.
- Prieto, P.; Aguilar, L., Mascaró, I., Torres, F. J. & Vanrell, M. M. 2007. *CatToBI (Catalan Tones and Break Indices)*. Consultado el 24 de junio de 2007 en <http://seneca.uab.es/atlesentonacio/cat-tobi/Cat-ToBI.pdf>
- Quilis, A. 1981. *Fonética acústica de la lengua española*. Madrid: Gredos.
- Quilis, A. 1993. *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- Silverman, K.; Beckman, M.; Pitrelli, J.; Ostendorf, M.; Wightman, C.; Price, P. et al. 1992. "TOBI: A standard for labelling English prosody". In *ICSLP 92, Proceedings of the Second International Conference on Spoken Language Processing* (pp. 867-870). Consultado el 24 de junio de 2007 en http://ling.osu.edu/~tobi/ame_tobi/Silverman_etal1992.pdf
- Sosa, J. M. 2003. "La notación tonal del español en el modelo Sp-ToBI". In P. Prieto (Ed.), *Teorías de la entonación* (pp. 185-208). Barcelona: Ariel .
- T'Hart, J.; Collier, R. & Cohen, A. 1990. *A perceptual study of intonation. An experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.
- Thorsen, N. 1983. "Standard Danish sentence intonation - Phonetic data and their representation". *Folia Linguistica*, 17, 187-220.
- Wells, J. C. 1995. *SAMPROSA, SAM Prosodic Transcription*. Consultado el 14 de junio de 2007 en <http://www.phon.ucl.ac.uk/home/sampa/samprosa.htm>
- Wells, J. C. 2003. *SAMPA Computer Readable Phonetic Alphabet*. Consultado el 24 de junio de 2007 en <http://www.phon.ucl.ac.uk/home/sampa/home.htm>