

TÉCNICAS Y PROCEDIMIENTOS PARA LA REPRESENTACIÓN DE LAS CURVAS MELÓDICAS

Mònica Estruch
Juan María Garrido
Joaquim Llisterri
Montserrat Riera

Departament de Filologia Espanyola
Universitat Autònoma de Barcelona

1. Introducción

La representación de las curvas melódicas es uno de los principales problemas que se plantean en el estudio de la entonación. La primera dificultad para el análisis prosódico del habla aparece ya 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) 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 enfoques. 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.

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 emplean normalmente 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 frecuencial. 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 cálculo¹.

En la figura 1 se presenta un ejemplo de curva melódica (correspondiente a la oración "Una vez en el interior, los agentes encontraron a tres de los niños", seleccionada en este trabajo para la ejemplificación de los diferentes sistemas), que se ha obtenido 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 d'Aix-en-Provence*, e integradas en el programa 'Phonédit', comercializado por la empresa SQLab. En la figura se pueden apreciar las diferencias de estimación existentes entre uno y otro sistema.

¹ Una presentación general de los diferentes sistemas de estimación de la F_0 puede encontrarse en Hess (1983).

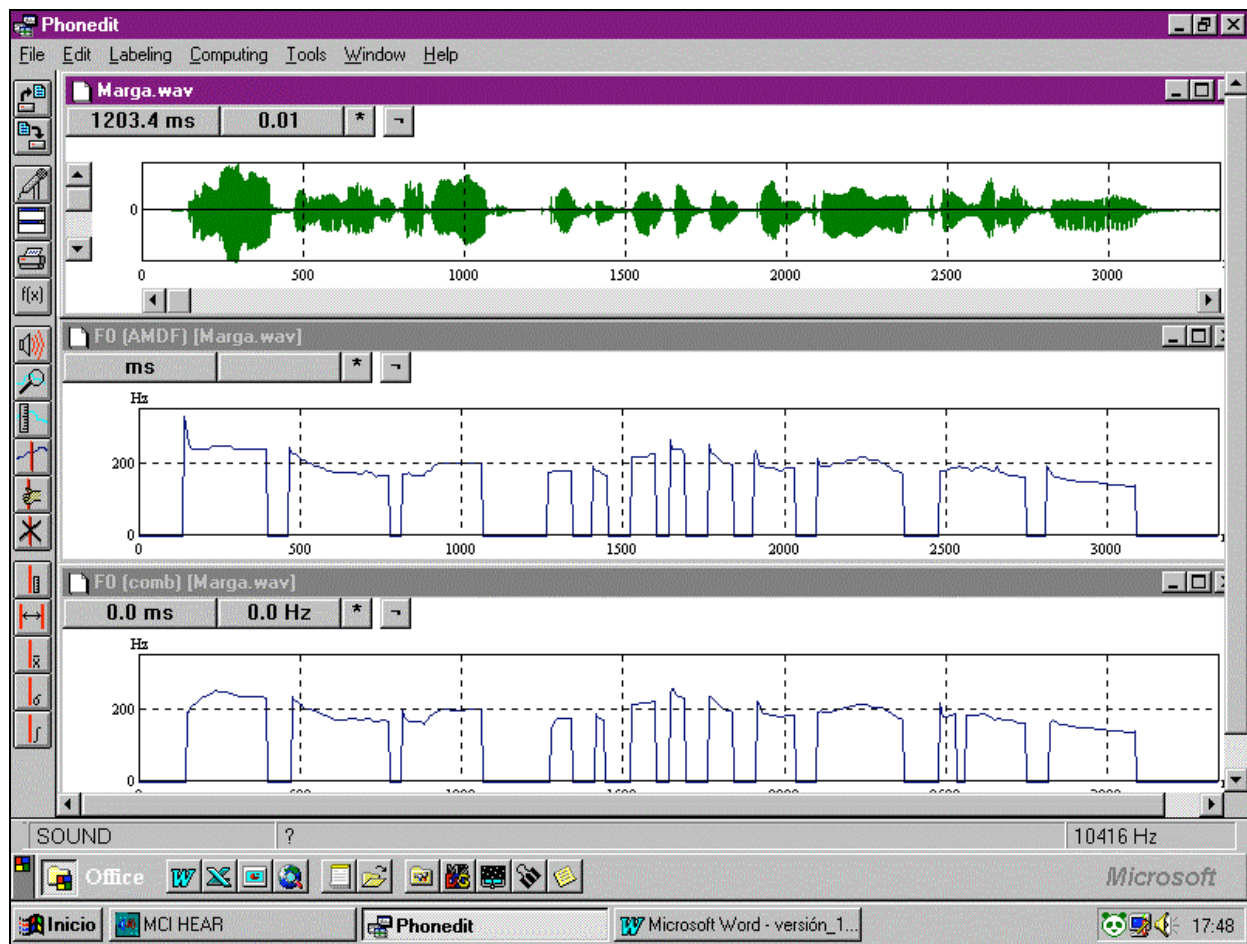


Figura 1. Oscilograma (ventana superior), curva melódica obtenida mediante un sistema de estimación temporal (ventana central) y curva melódica obtenida mediante un sistema de estimación frecuencial (ventana inferior) correspondiente al enunciado "Una vez en el interior, los agentes encontraron a tres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

Las curvas melódicas que se obtienen tanto con los sistemas temporales como con los 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 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 Hz. y 300 Hz., mientras que en el caso de un locutor femenino, el rango oscila entre 130Hz. 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 (obtenida con el sistema de delimitación del rango) han desaparecido los errores de estimación presentes en la primera estimación de la curva (obtenida sin delimitar el rango). En este caso, se ha utilizado una técnica frecuencial, basada en el sistema descrito en Secret y Doddington (1983), que está integrada en el programa 'xwaves', de la empresa Entropic.

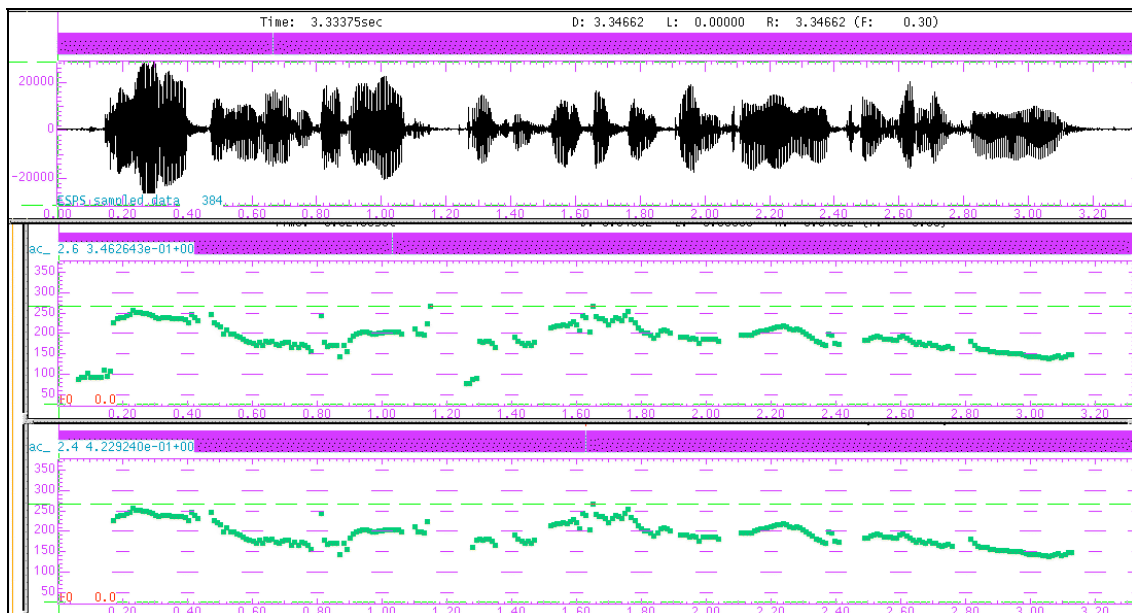


Figura 2. Oscilograma (ventana superior), curva melódica obtenida mediante un sistema de estimación frecuencial sin delimitación del rango (ventana central) y con delimitación de rango (ventana inferior) correspondiente al enunciado "Una vez en el interior, los agentes encontraron a tres de los niños", 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 proceso 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 d'Aix-en-Provence*), cada valor de F_0 se compara con los valores anteriores y posteriores y, 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 (como sucede en el programa 'SoundScope', de GW Instruments).

(b) En segundo lugar, los valores de F_0 considerados como errores de detección se aproximan a los valores adyacentes (programa 'Pitch', desarrollado por miembros del Departamento de Acústica de *Enginyeria La Salle* de la *Universitat Ramon Llull* de Barcelona; Jiménez, 1994; Martínez, 1995), se neutralizan ('Mes') o se eliminan de la curva melódica ('SoundScope').

En la figura 3 se puede observar un ejemplo de aplicación de esta técnica: en la ventana superior 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 'SoundScope', con los correspondientes errores de detección; en la ventana central aparece la misma curva tras el proceso de alisado, en la que han desaparecido la mayoría de los errores que se apreciaban en la ventana superior.

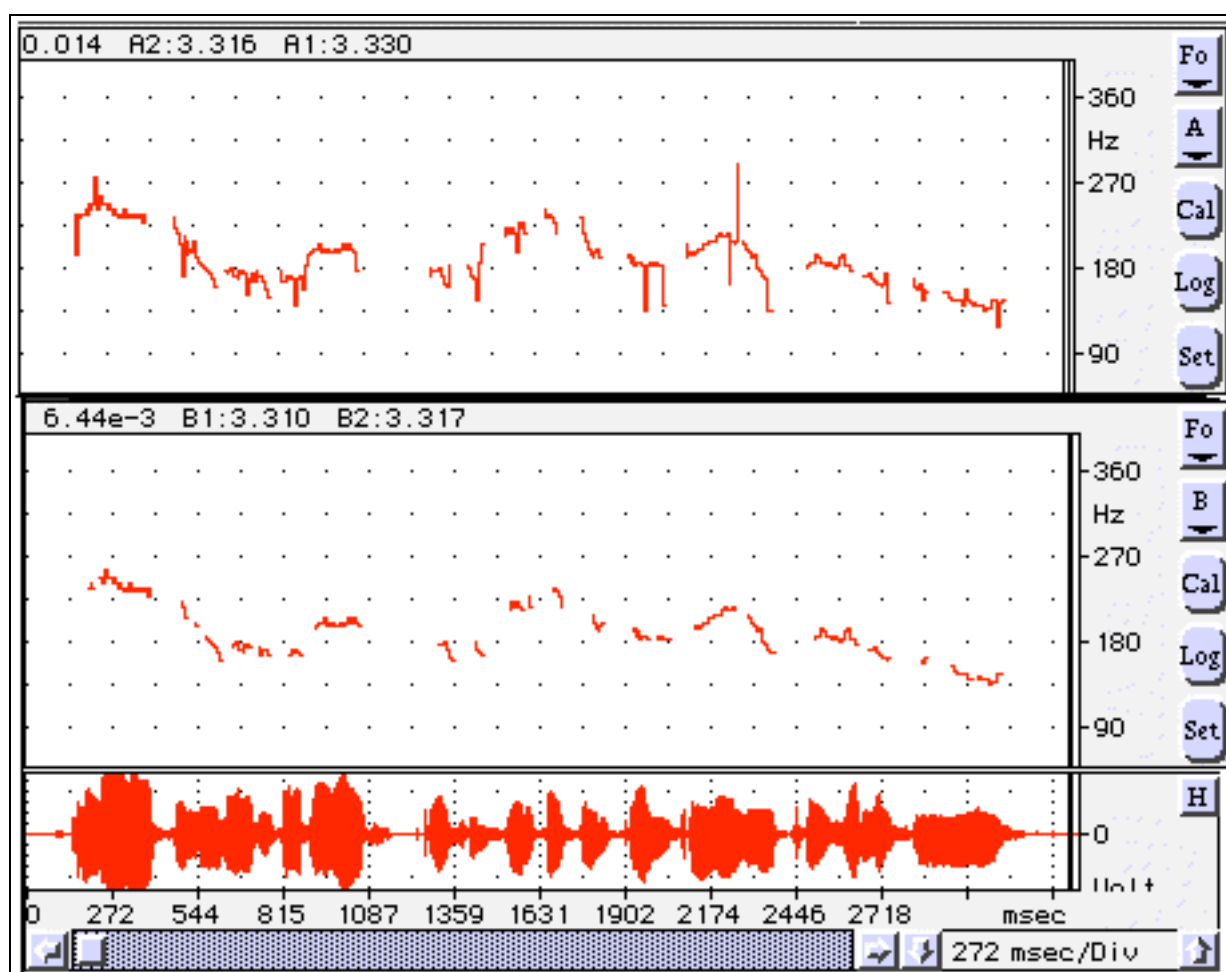


Figura 3. Oscilograma (ventana inferior), curva melódica obtenida mediante un sistema de estimación temporal sin alisado (ventana superior) y con alisado (ventana central) correspondiente al enunciado "Una vez en el interior, los agentes encontraron a tres de los niños", pronunciado por un locutor femenino.

3. Estilización

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; Llisterri, 1984, para el catalán).
- 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).
- 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, para el español).

En la figura 4 se aprecian algunos ejemplos 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.

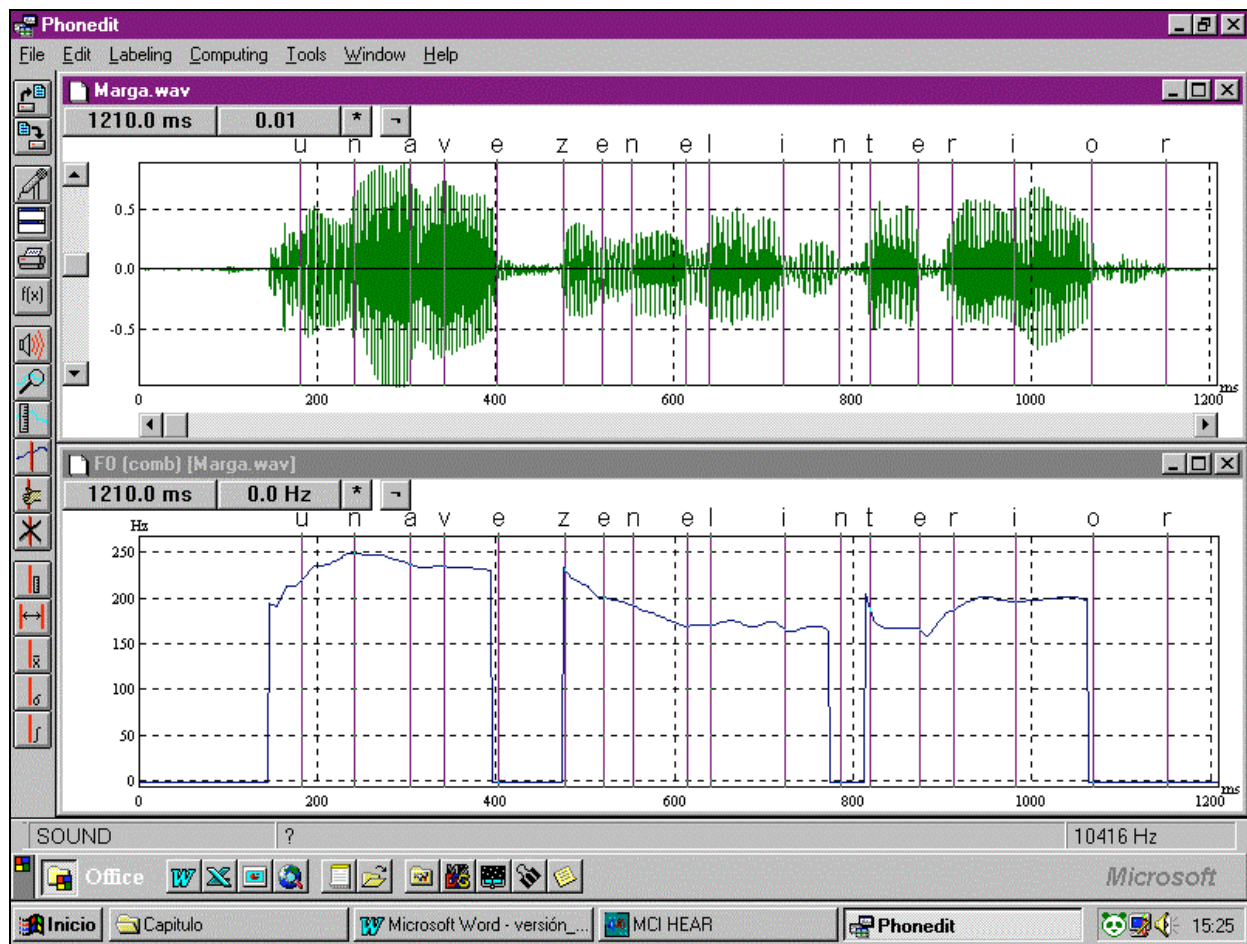


Figura 4. Oscilograma y curva melódica correspondientes al enunciado "Una vez en el interior", pronunciado por un locutor femenino. Las líneas verticales representan los límites entre segmentos.

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.

Otra posible clasificación de los sistemas de estilización se basa en la diferencia entre métodos manuales y automáticos. Partiremos de esta distinción para la descripción de los diferentes procedimientos de estilización en las secciones siguientes.

3.3.1. Métodos manuales de estilización

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. Es el caso de los tres métodos que se presentan a continuación.

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. En los estudios sobre el danés realizados por Thorsen (1983), por ejemplo, la F_0 de cada una de las vocales y de las consonantes silábicas se mide en un punto determinado de éstas (a 2/3 de la distancia temporal desde el inicio del núcleo silábico).

Otro método de estilización manual puede consistir en la obtención de la F_0 en el centro de las vocales que constituyen núcleo silábico, como en Garrido *et al.* (1993, 1995a, b). En estos estudios, en los que se llevó a cabo un análisis de 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.

El tercer sistema que presentamos (utilizado en Estruch y Garrido, 1995) tiene 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 obtiene partiendo del valor de la F_0 en el inicio, centro y final de la vocal, tal como ilustra la figura 5.

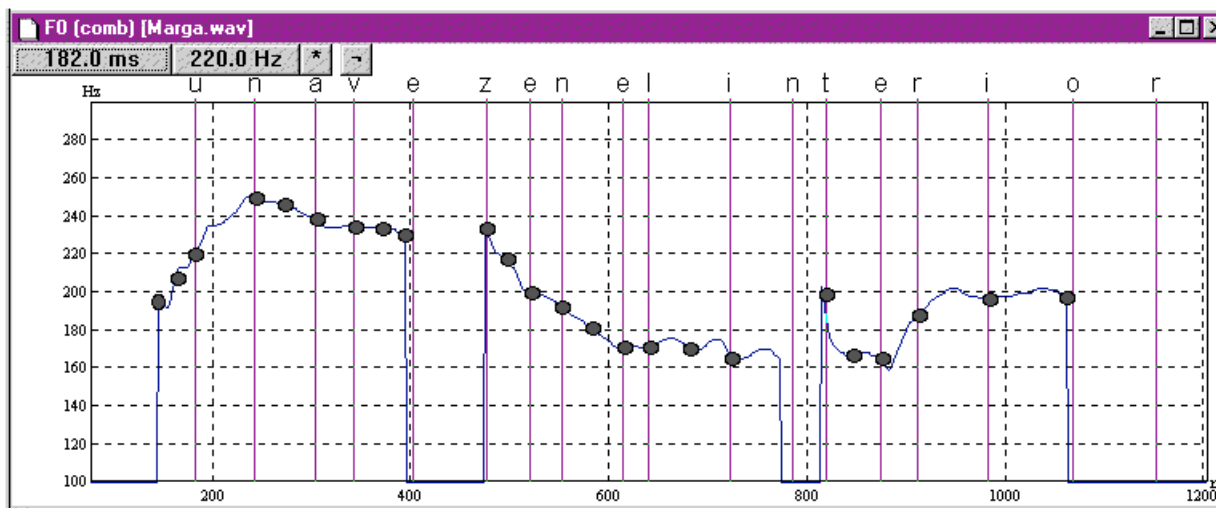


Figura 5. Curva melódica y estilizada correspondientes al enunciado "Una vez en el interior", pronunciado por un locutor femenino. Los círculos indican los puntos de la curva que se corresponden con el inicio, centro o final de los segmentos vocálicos del enunciado.

Sin embargo, no todos los métodos manuales de estilización se basan en criterios lingüísticos. También los hay que emplean criterios de tipo acústico para la determinación de los puntos que compondrán la curva estilizada.

Uno de estos sistemas es el desarrollado por los investigadores del IPO en Eindhoven, descrito en t'Hart, Collier y Cohen (1990). 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*.

Otro procedimiento basado en criterios acústicos es el descrito en Garrido (1991a, b), donde se define un sistema de estilización que tiene en cuenta las diferencias frecuenciales entre los distintos puntos de la curva melódica. 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 comprende 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.

Estos diferentes métodos resultan adecuados para el análisis de corpus no demasiado extensos, como en el caso de los estudios mencionados. Para la estilización de curvas en un corpus más amplio, es preferible, sin embargo, el uso de sistemas automáticos como los que se presentan a continuación.

3.3.2. Métodos automáticos de estilización

Los métodos automáticos de estilización se basan generalmente en criterios acústicos semejantes a los utilizados en los correspondientes métodos manuales, ya que los procedimientos basados en criterios lingüísticos son mucho más difíciles de automatizar; existe, sin embargo, algún sistema de este tipo, como el descrito en d'Alessandro y Mertens (1995). Los procedimientos automáticos de estilización convierten la curva melódica en una serie de puntos de inflexión unidos mediante líneas. Las diferencias entre los diferentes métodos 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.

Uno de los sistemas de estilización automática existentes es MOMEL (*MOdelling MELody*) desarrollado en el *Laboratoire Parole et Langage* de la *Université d'Aix-en-Provence* (Hirst y Espesser, 1993). El algoritmo MOMEL (integrado en los programas 'MES' y 'Phonédit') 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 marco del proyecto MULTEXT (*Multilingual Text Tools and Corpora*) (Hirst *et al.*, 1994), para realizar la anotación prosódica del corpus EUROM1 en varias lenguas europeas, entre otras el español (Campione y Véronis, 1998; Llisterra (Ed.), 1996). Actualmente se está llevando a cabo la anotación de un corpus semejante en catalán basada en el mismo procedimiento (Estruch, en preparación). En el marco de este proyecto se realizó una validación perceptiva del sistema, 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) es 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 6.

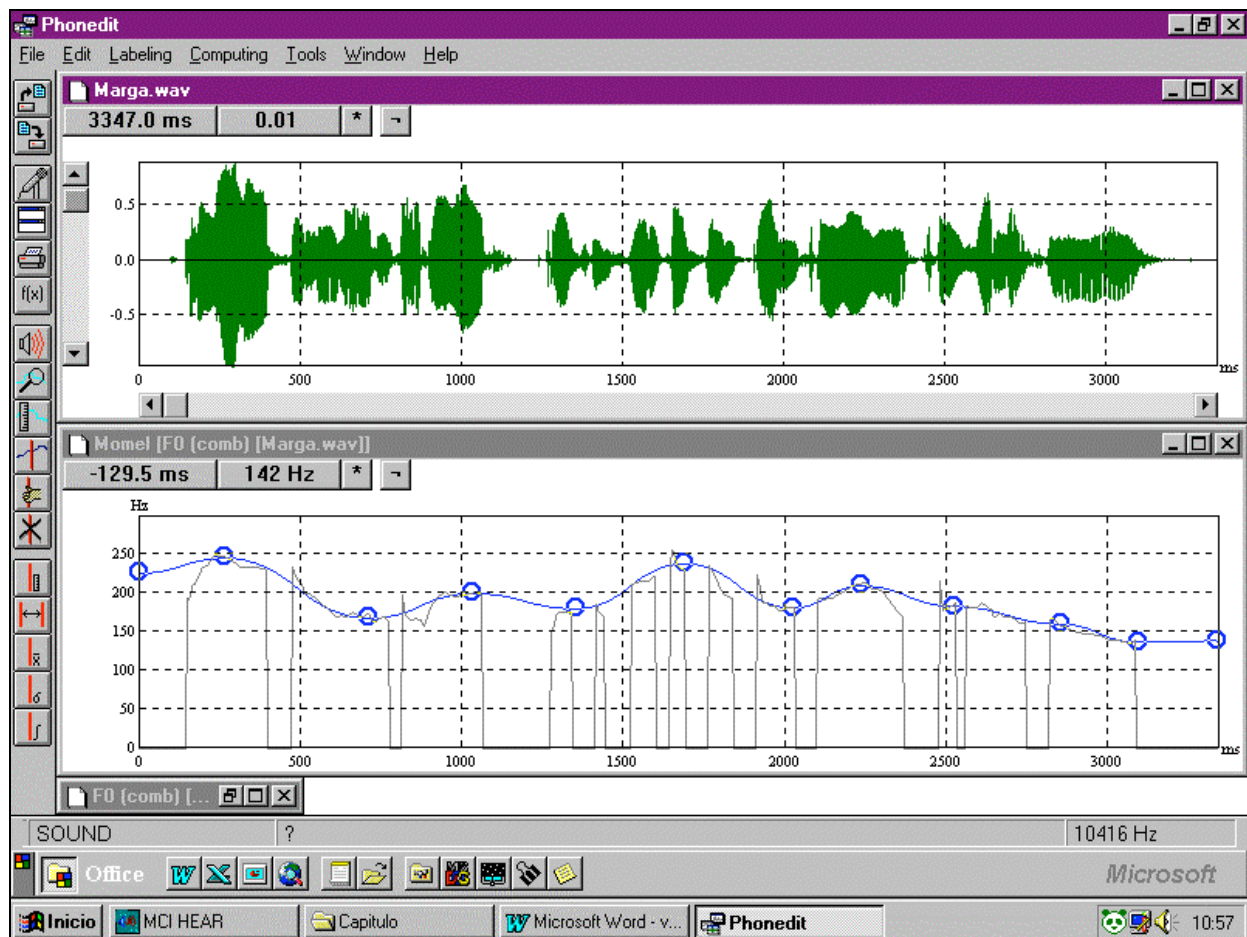


Figura 6. Oscilograma (ventana superior) y curva melódica (ventana inferior) correspondientes al enunciado "Una vez en el interior, los agentes encontraron a tres de los niños", pronunciado por un locutor femenino.

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 es el integrado en el programa 'Pitch', una aplicación para Windows que permite 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 determina automáticamente los puntos de inflexión en función de un 'umbral de semejanza' establecido previamente por el usuario, de manera que éste puede definir hasta qué punto la curva estilizada obtenida se asemejará a la curva original. Este sistema ha sido utilizado, entre otros, por Estruch *et al.* (1995) para el catalán y Garrido (1996)

para el español. La figura 7 presenta la curva melódica del enunciado de referencia estilizada mediante el programa 'Pitch'.

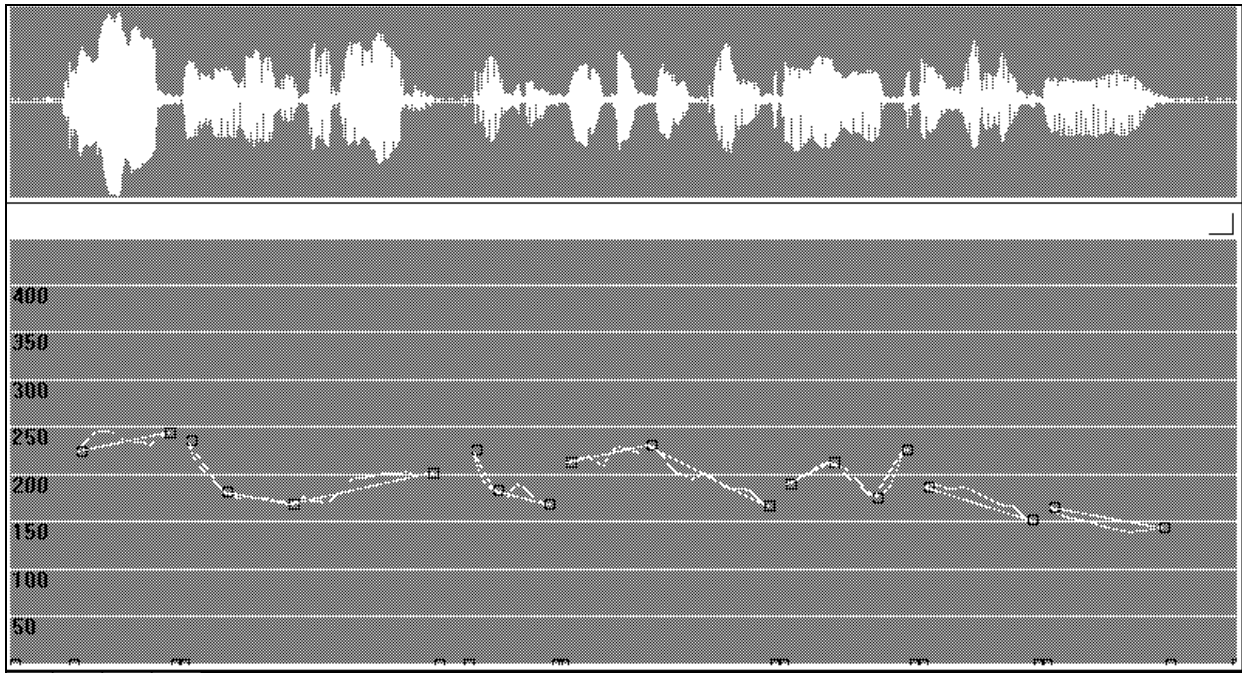


Figura 7. Curva melódica original y contorno estilizado, obtenidos mediante el programa 'Pitch', del enunciado "Una vez en el interior, los agentes encontraron a tres de los niños", pronunciado por un locutor femenino.

Una de las ventajas que presenta 'Pitch' es la posibilidad de que el usuario determine 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) y Estruch *et al.* (1995, en prensa). Mediante el umbral definido en estas pruebas es posible obtener curvas equivalentes perceptivamente a las originales sin necesidad de corrección manual.

Contrariamente al sistema MOMEL, descrito anteriormente, el algoritmo integrado en 'Pitch' interrumpe la curva estilizada cada vez que aparecen un segmento sordo o una pausa en el enunciado, tal y como puede observarse en la figura 7.

3.4. La normalización

El proceso de estilización permite eliminar 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), la altura (ascenso o descenso) se cuantifica por medio de porcentajes con respecto al total del rango del locutor. Un sistema similar ha sido adoptado por de la Mota (1995) para cuantificar la pendiente de las líneas que unen los diferentes puntos de inflexión.

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 8. Por otro lado, el uso de escalas logarítmicas refleja más fielmente el proceso de percepción de la melodía (t'Hart *et al.*, 1990).

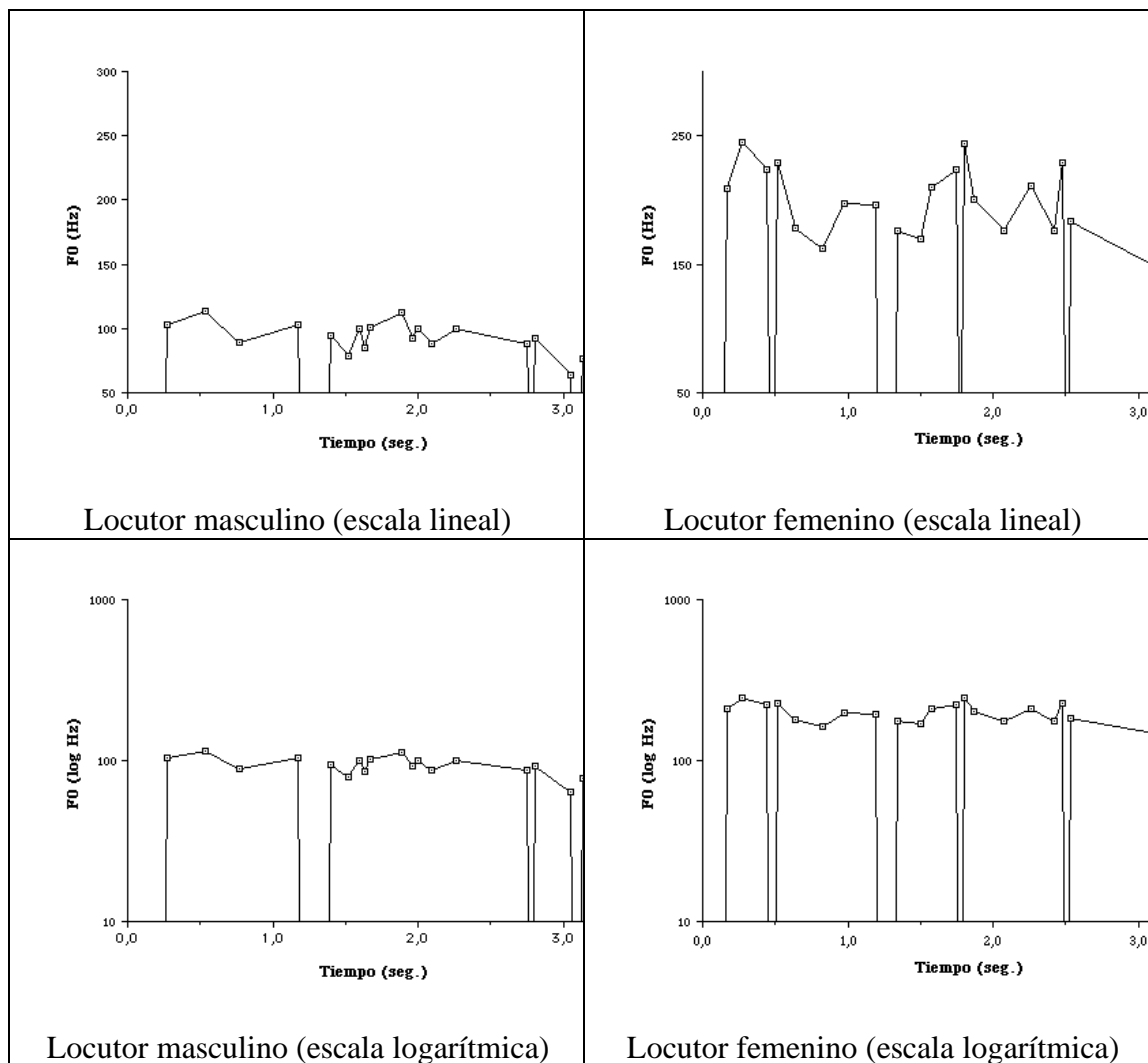


Figura 8. Curvas estilizadas correspondientes al enunciado "Una vez en el interior, los agentes encontraron a tres de los niños" pronunciadas por un locutor masculino (izquierda) y otro femenino (derecha), representadas utilizando una escala frecuencial lineal (arriba) y logarítmica (abajo).

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. Representación simbólica de las curvas melódicas: la anotación²

² '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 etiquetar directamente un texto ortográfico o una transcripción fonética ('transcripción'). En este trabajo utilizaremos al término 'anotación' para referirnos indistintamente a ambos tipos de etiquetado.

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 diferentes, 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 *et al.* (Eds.), 1995). Una presentación más detallada de los principales métodos de anotación se encuentra en Llisterrí (1994; 1996; 1997) y en Klein *et al.* (1998).

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').

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), en tanto que la segunda considera las curvas melódicas como el resultado de la concatenación de una serie de segmentos (análisis por contornos). Teniendo en cuenta esta clasificación, los diferentes sistemas de anotación fonética pueden clasificarse 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 9.

[‘una’βeθenelintə’rior ↑ | losa’xentəsəŋkon’tarona’treʃðlos’niɲos ↓]

Una vez en el interior (A), los agentes encontraron a tres de los niños (C)

Figura 9. 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) es quizá el mejor ejemplo del tipo de anotación al que nos estamos refiriendo. Este método se emplea para etiquetar los diferentes segmentos (líneas rectas que unen dos puntos de inflexión) de una curva previamente estilizada mediante el procedimiento de la *close-copy*. En la figura 10 puede observarse el enunciado de referencia transcrito mediante este sistema.

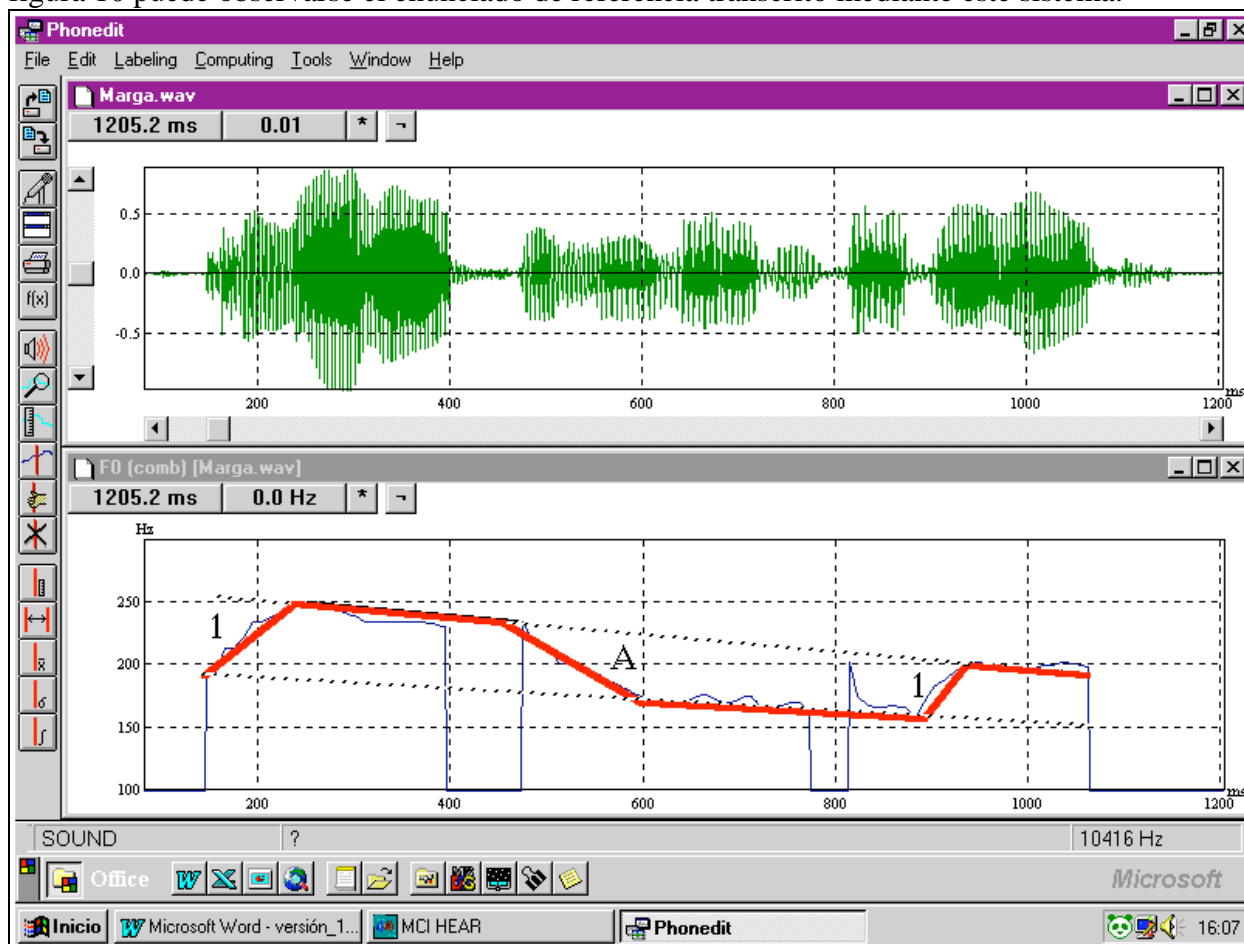


Figura 10. Anotación de la curva estandarizada correspondiente al enunciado "Una vez en el interior", según el método descrito en t'Hart *et al.* (1990).

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 sistema 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) 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 11. 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.

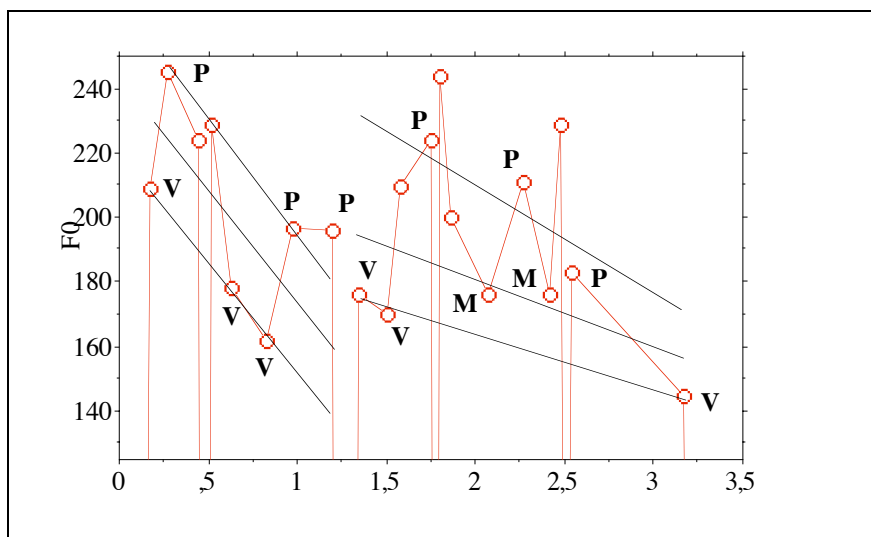


Figura 11. Curva estilizada y anotada según el sistema de Garrido (1996) correspondiente a la oración 'Una vez en el interior, los agentes encontraron a tres de los niños', pronunciada por un locutor femenino.

El sistema de anotación INTSINT (*IN*ternational *T*ranscription *S*ystem for *IN*Tonation), desarrollado en la *Université d'Aix-en-Provence* (Campione *et al.*, 1997; Hirst, 1991, 1994; Hirst y Di Cristo, 1998), 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'. 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 12 presenta un ejemplo de curva anotada mediante este sistema.

Este sistema se ha empleado para la anotación prosódica del corpus EUROM1 en español (Campione y Véronis, 1998; Llisterra (Ed.), 1996) y de la versión correspondiente en catalán (Estruch, en preparación).

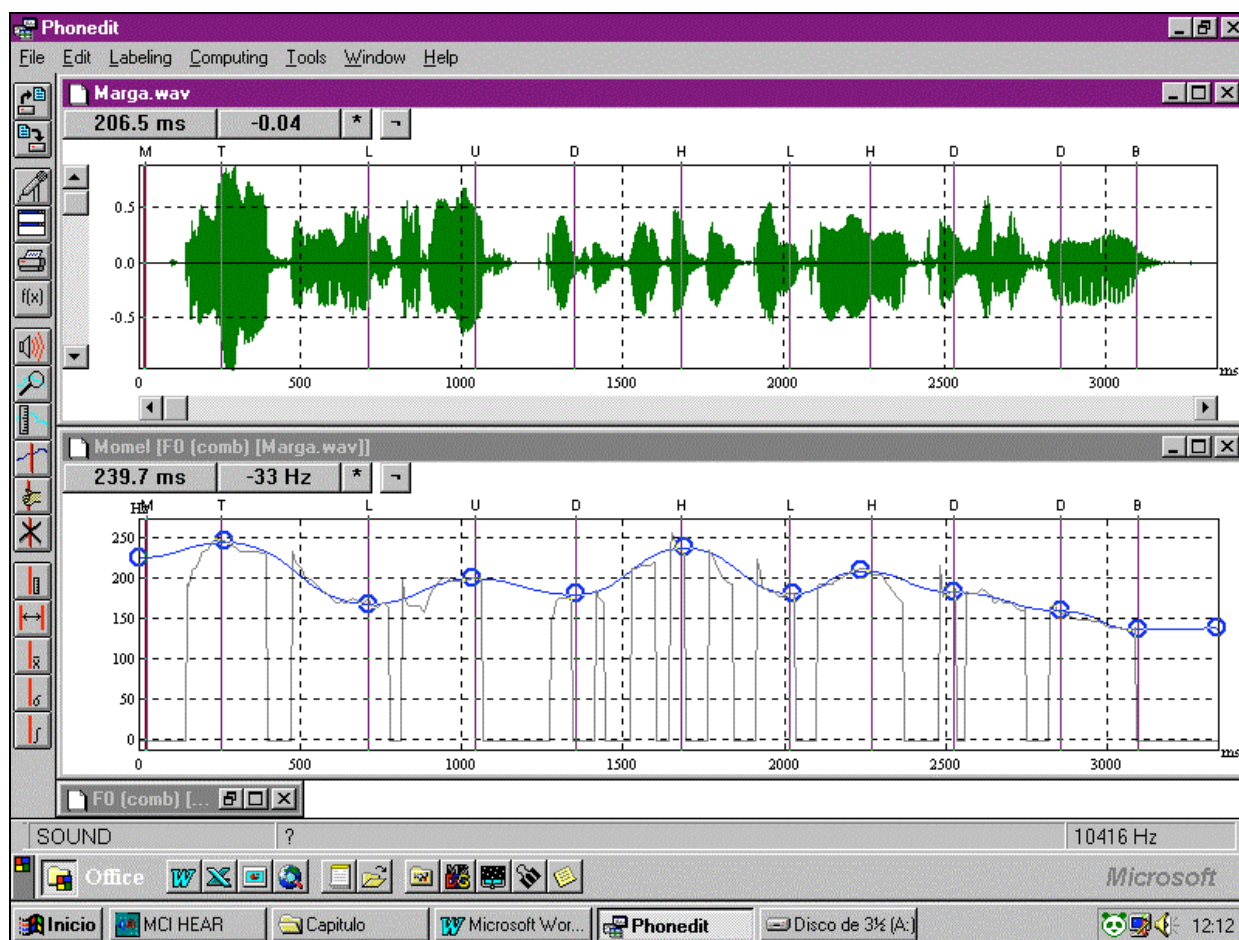


Figura 12. Oscilograma (ventana superior), curva melódica original, curva estilizada (sistema MOMEL) y anotación INTSINT (ventana inferior) correspondientes al enunciado "Una vez en el interior, los agentes encontraron a tres de los niños", pronunciado por un locutor femenino.

Al igual que el sistema de Garrido (1996), 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, el principal procedimiento 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, 1994). 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, 1991, 1995; Prieto *et al.*, 1995, 1996) y del catalán (Prieto, 1995).

El sistema recoge el inventario de unidades entonativas propuestas en el modelo de Pierrehumbert (Pierrehumbert, 1980; Beckman y Pierrehumbert, 1986), 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 13 presenta un ejemplo de curva anotada mediante ToBI.

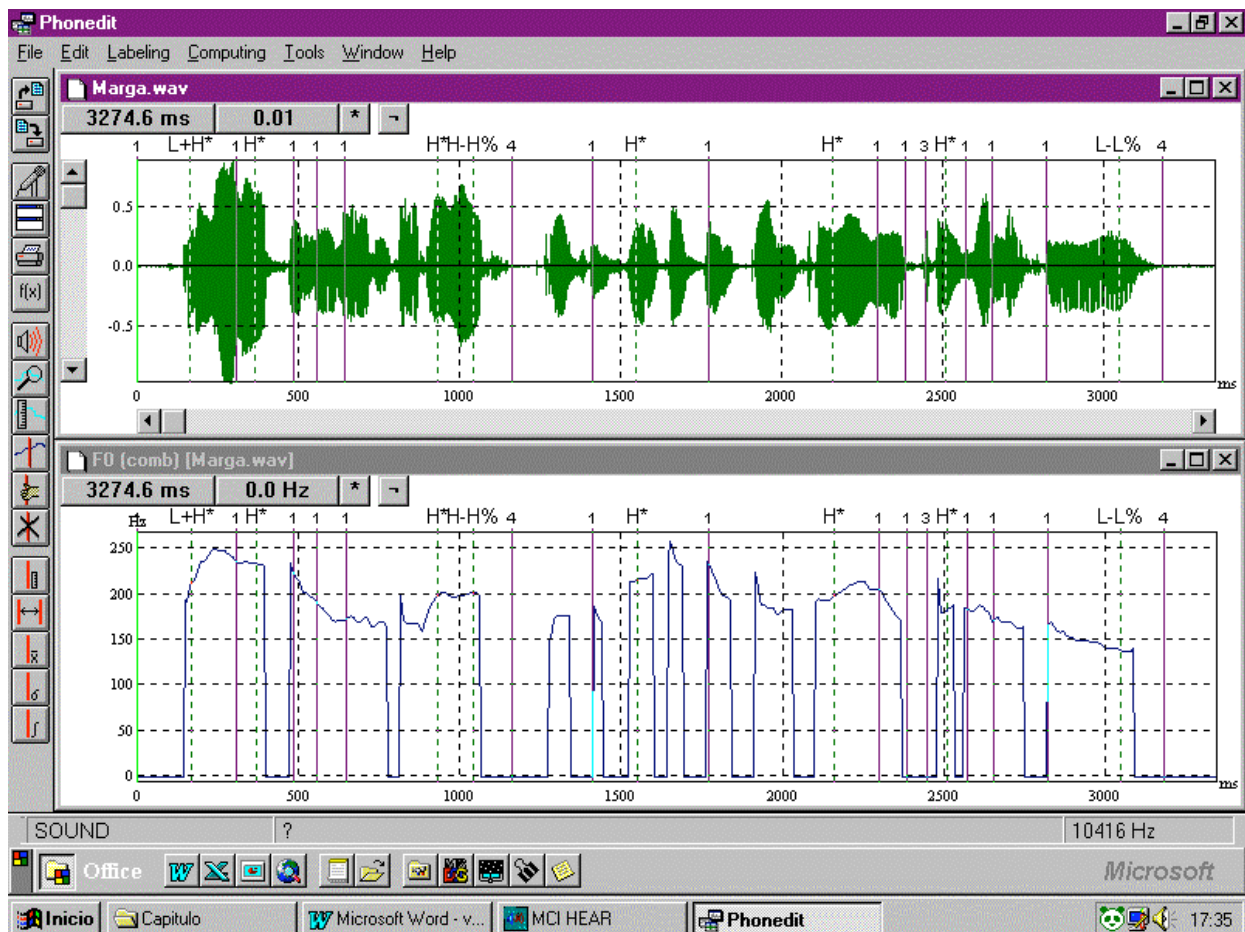


Figura 13. Oscilograma, curva melódica y anotación mediante el sistema ToBI del enunciado "Una vez en el interior, los agentes encontraron a tres de los niños", pronunciado por un locutor femenino.

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 más recientes en fonología prosódica.

5. Conclusión

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. De esta presentación se deduce que no parece existir hoy en día un procedimiento que pueda considerarse estándar (equivalente al Alfabeto Fonético Internacional o a SAMPA 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

Esta es la línea que orienta los trabajos sobre la entonación del castellano y el catalán descritos en Estruch *et al.* (1996) y la que se sigue igualmente en el marco del proyecto MATE (*Multilevel Annotation Tools Engineering*), que pretende definir un sistema para la anotación de diálogos en diversos niveles de análisis, entre ellos el prosódico, tarea en la que está implicado nuestro grupo actualmente.

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

ALCINA, J. - BLECUA, J.M. (1975).- *Gramática española*, Barcelona, Ariel.

BECKMAN, M.E. - PIERREHUMBERT, J.B. (1986).- "Intonational structure in English and Japanese", *Phonology Yearbook*, 3: 255-310.

BECKMAN, M.E. - AYERS, G.M. (1994).- *Guidelines for ToBI Labelling*. Version 2.0, February 1994. URL: <http://ling.ohio-state.edu/Phonetics/ToBI/ToBI.0.html>

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.

CAMPIONE, E. - FLACHAIRE, E. - HIRST, D. - VÉRONIS, J. (1997).- "Stylisation and symbolic coding of F₀: A quantitative model", en BOTINIS, A. - KOUROUPETROGLOU, G. - CARAYIANNIS, G. (eds.) (1997).- *Intonation: Theory, Models and Applications. Proceedings of an ESCA Workshop, September 18-20, Athens, Greece*, ESCA/University of Athens, pp. 71-74.

ESTRUCH, M.- GARRIDO, J.M.- LLISTERRI, J.- RIERA, M. (1999) "Técnicas y procedimientos para la representación de las curvas melódicas". Unpublished Ms. Grup de Fonètica, Departament de Filologia Espanyola, Universitat Autònoma de Barcelona,

CAMPIONE, E. - VÉRONIS, J. (1998).- "A Multilingual Prosodic Database", en *ICSLP'98, Proceedings of the 5th International Conference on Spoken Language Processing*, 30th November-4th December 1998, Sydney, Australia. Volume 7: 3163-3166.

CANTERO SERENA, F.J. (1995).- *Estructura de los modelos entonativos: interpretación fonológica del acento y la entonación en castellano*, Tesis doctoral, Universitat de Barcelona.

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, Departament de Filologia Espanyola, Universitat Autònoma de Barcelona.

DI CRISTO, A. (1982).- *Prolegomènes à l'étude de l'intonation. Micromélorie*. Paris, Éditions du CNRS.

EDWARDS, J.A. - LAMPERT, M.D. (Eds) (1993).- *Talking Data: Transcription and Coding in Discourse Research*. Hillsdale, N.J., Lawrence Erlbaum Associates.

ESTRUCH, M. (en preparació).- *Anàlisi melòdica i codificació simbòlica d'un corpus de paràgrafs en català*, Proyecto de tercer ciclo, Universitat Autònoma de Barcelona.

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, 11-14 de diciembre de 1995. Resumen publicado en *Revista Española de Lingüística* 26,1: 138-139.

ESTRUCH, M. - MIMÓ, M. - RENOM, J. - RIERA, M. (1995).- *Validació perceptiva dels patrons melòdics del català*, Manuscrito no publicado, Universitat Autònoma de Barcelona.

ESTRUCH, M. - GARRIDO, J.M. - LLISTERRI, J. - RIERA, M. (1996).- "Una aproximación fonética al estudio de la entonación", *Philologia Hispalensis* XI: 281-293.

ESTRUCH, M. - GARRIDO, J.M. - RIERA, M. (1997).- "Estudio perceptivo de los tonemas del español", Comunicación presentada en el XXVII Simposio de la Sociedad Española de Lingüística, Palma de Mallorca, 15-19 de diciembre de 1997. Resumen publicado en *Revista Española de Lingüística* 28, 1 (1998): 178-179.

ESTRUCH, M. - GARRIDO, J.M. - GUDAYOL, F. - JIMÉNEZ, J.M. - RENOM, J. - RIERA, M. (en prensa).- "Validación perceptiva de un sistema de estilización automática de contornos melódicos", *Actas del I Congrés de Fonètica Experimental*, Tarragona, 22-24 de febrero de 1999.

FANT, L. (1984).- *Estructura informativa en español. Estudio sintáctico y entonativo*. Upsala,

ESTRUCH, M. - GARRIDO, J.M. - LLISTERRI, J. - RIERA, M. (1999) "Técnicas y procedimientos para la representación de las curvas melódicas". Unpublished Ms. Grup de Fonètica, Departament de Filologia Espanyola, Universitat Autònoma de Barcelona,

Acta Universitatis Upsaliensis.

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*, Boletín nº 11: 209-219.

GARRIDO, J.M. - LLISTERRI, J. - de la MOTA, C. - RÍOS, A. (1993).- "Prosodic differences in reading style: Isolated vs. Contextualized Sentences", en *Eurospeech'93. 3rd European Conference on Speech Communication and Technology*. Berlin, Germany, 21-23 September 1993. Vol 1: 573-576.

GARRIDO, J.M. - LLISTERRI, J. - MARÍN, R. - de la MOTA, C. - RÍOS, A. (1995a).- "Prosodic markers at syntactic boundaries in Spanish", en ELENIUS, K.- BRANDERUD, P. (Eds.) *ICPhS 95, Proceedings of the XIIIth International Congress of Phonetic Sciences*. Stockholm, Sweden, 13-19 August, 1995. Vol. 2: 370-373.

GARRIDO, J.M. - LLISTERRI, J. - de la MOTA, C. - RÍOS, A. (1995b).- "Estudio comparado de las características prosódicas de la oración simple en español en dos modalidades de lectura", en ELEJABEITIA, A. - IRIBAR, A. (Eds.) *Phonetica. Trabajos de fonética experimental*. Bilbao: Universidad de Deusto (Serie Lingüística, 6) pp. 173-194.

GARRIDO, J.M. (1996).- *Modelling Spanish Intonation for Text-to-Speech Applications*, Tesis doctoral, Departament de Filologia Espanyola, Universitat Autònoma de Barcelona. URL: <http://liceu.uab.es/~juanma/tesis.html>

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*, Berlín, Springer-Verlag.

HIRST, D.J. (1991).- "Intonation models: Towards a third generation", en *Actes du XIIème Congrès International des Sciences Phonétiques. 19-24 août 1991, Aix-en-Provence, France*. Aix-en-Provence: Université de Provence, Service des Publications. Vol. 1: 305-310.

HIRST, D.J. - ESPESSER, R. (1993).- "Automatic modelling of fundamental frequency using a quadratic spline function", *Travaux de l'Institut de Phonétique d'Aix*, 15: 75-85.

HIRST, D.J. (1994).- "The symbolic coding of fundamental frequency curves: from acoustics to phonology", en FUJISAKI, H. (Ed.) *Proceedings of International Symposium on Prosody, Satellite Workshop of ICSLP 94*, Yokohama, September 1994.

HIRST, D.J. - IDE, N. - VÉRONIS, J. (1994) "Coding fundamental frequency patterns for multi-

ESTRUCH, M.- GARRIDO, J.M.- LLISTERRI, J.- RIERA, M. (1999) "Técnicas y procedimientos para la representación de las curvas melódicas". Unpublished Ms. Grup de Fonètica, Departament de Filologia Espanyola, Universitat Autònoma de Barcelona,

lingual synthesis with INTSINT in the MULTEXT project", en *Conference Proceedings of the Second ESCA/IEEE Workshop on Speech Synthesis*. September 12-15, 1994. Mohonk Mountain House, New Paltz, New York, USA. pp. 77-80 .

HIRST, D.J. - DI CRISTO, A. (Eds.) (1998).- *Intonation Systems. A Survey of Twenty Languages*. Cambridge, Cambridge University Press.

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, Barcelona, Enginyeria La Salle, Universitat Ramon Llull.

KLEIN, M. - BERNSEN, N.O. - DAVIES, S. - DYBKJAER, L. - GARRIDO, J.M. - KASCH, H. - MENGEL, A. - PIRRELLI, V. - POESIO, M. - QUAZZA, S. - SORIA, C. (1998).- *Supported Coding Schemes*, Multilevel Annotation Tools Engineering, LE Telematics Project LE4-8370, Deliverable 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. (1984).- *Anàlisi i síntesi de vocals catalanes. Determinació experimental de la freqüència dels tres primers formants en un corpus de vocals en context realitzades per quatre parlants de català central*. Tesis de licenciatura, Universitat Autònoma de Barcelona.

LLISTERRI, J. (1994).- *Prosody Encoding Survey*. WP 1 Specifications and Standards. T1.5. Markup Specifications. Deliverable 1.5.3. Final version, 15 September, 1994. LRE Project 62-050 MULTEXT.

LLISTERRI, J. - MARÍN, R. - de la MOTA, C. - RÍOS, A. (1995).- "Factors affecting F₀ peak displacement in Spanish", en PARDO, J.M.- ENRÍQUEZ, E. - ORTEGA, J. - FERREIROS, J. - MACÍAS, J. - VALVERDE, F.J. (Eds.) *Eurospeech'95. 4th European Conference on Speech Communication and Technology*. Madrid, Spain, 18-21 September, 1995. Vol. 3: 2061-2064.

LLISTERRI, J. (Ed.) (1996).- *Prosody Tools Efficiency and Failures*. WP 4 Corpus. T4.6 Speech Markup and Validation. Deliverable 4.5.2. Final version. 15 October 1996. LRE Project 62-050 MULTEXT.

ESTRUCH, M.- GARRIDO, J.M.- LLISTERRI, J.- RIERA, M. (1999) "Técnicas y procedimientos para la representación de las curvas melódicas". Unpublished Ms. Grup de Fonètica, Departament de Filologia Espanyola, Universitat Autònoma de Barcelona,

LLISTERRI, J. (1997).- *Transcripción, etiquetado y codificación de corpus orales*. Fundación Duques de Soria, Seminario de Industrias de la Lengua, 15 de julio de 1997. URL: <http://liceu.uab.es/~joaquim/publicacions/FDS97.html>

MARTÍNEZ, D. (1995).- *Sistema d'anàlisi, tractament i síntesi de la melodia en entorn Windows*, Manuscrito no publicado, Barcelona, Enginyeria La Salle, Universitat Ramon Llull.

MATEO, A. (1988).- "Experimento sobre el tono intrínseco de las vocales castellanas", *Estudios de Fonética Experimental III*: 157-179.

NAVARRO TOMÁS, T. (1944).- *Manual de entonación española*. New York: Hispanic Institute. Cuarta edición: Madrid, Guadarrama (Punto Omega, 175), 1974.

ORLIKOFF, R.F. - KAHANE, J.C. (1996).- "Structure and function of the larynx", in LASS, N.J. (Ed.) *Principles of Experimental Phonetics*. St Louis, Mosby, pp. 112-184.

PAYRATÓ, L. (1995).- "Transcripción del discurso coloquial", in CORTÉS RODRÍGUEZ, L. (Ed.) *El español coloquial. Actas del I Simposio sobre Análisis del Discurso Oral*. Almería, 23-25 de noviembre de 1994. Almería, Universidad de Almería, Servicio de Publicaciones, pp. 43-70.

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).- *The Phonology and Phonetics of English Intonation*. Bloomington, Indiana University Linguistics Club.

PRIETO, P. (1995).- "Aproximació als contorns entonatius del català central", *Caplletra, Revista Internacional de Filologia* (València), 19: 161-186.

PRIETO, P. - van SANTEN, J. - HIRSCHBERG, J. (1995).- "Tonal Alignment Patterns in Spanish", *Journal of Phonetics* 23: 429-451.

PRIETO, P. - SHIH, S. - NIBERT, H. (1996).- "Pitch downtrend in Spanish", *Journal of Phonetics* 24,4: 445-473.

QUILIS, A. (1981).- *Fonética acústica de la lengua española*. Madrid, Gredos.

QUILIS, A. (1993).- *Tratado de Fonología y Fonética españolas*. Madrid, Gredos.

SECRET, B.G. - DODDINGTON, G.R. (1983).- "An integrated pitch tracking algorithm for speech systems", *Proceedings ICASSP83*, pp. 1352-1355.

SILVERMAN, K. - BECKMAN, M. - PITRELLI, J. - OSTENDORF, M. - WIGHTMAN, C. - PRICE, P. - PIERREHUMBERT, J. - HIRSCHBERG, J. (1992).- "TOBI: A standard for labelling English prosody", in OHALA, J.J. *et al.* (Eds.) *Proceedings of the Second International*

ESTRUCH, M.- GARRIDO, J.M.- LLISTERRI, J.- RIERA, M. (1999) "Técnicas y procedimientos para la representación de las curvas melódicas". Unpublished Ms. Grup de Fonètica, Departament de Filologia Espanyola, Universitat Autònoma de Barcelona,

Conference on Spoken Language Processing, ICSLP-92. Banff, October 1992, pp. 867-870.

SOSA, J.M. (1991).- *Fonética y Fonología de la entonación del español hispanoamericano*, Ph. D., UMI, University of Massachusetts.

SOSA, J.M. (1995).- "Nuclear and pre-nuclear tonal inventories and the phonology of Spanish declarative intonation", en ELENIUS, K. - BRANDERUD, P. (Eds.) *ICPhS 95, Proceedings of the XIIIth International Congress of Phonetic Sciences*. Stockholm, Sweden, 13-19 August, 1995. Vol. 4: 646-649.

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)*. URL: <http://www.phon.ucl.ac.uk/home/sampa/samprosa.htm>

ESTRUCH, M.- GARRIDO, J.M.- LLISTERRI, J.- RIERA, M. (1999) "Técnicas y procedimientos para la representación de las curvas melódicas". Unpublished Ms. Grup de Fonètica, Departament de Filologia Espanyola, Universitat Autònoma de Barcelona,