

Juan María Garrido Almiñana

**MODELIZACIÓN
DE PATRONES MELÓDICOS
DEL ESPAÑOL
PARA LA SÍNTESIS
Y EL RECONOCIMIENTO DE HABLA**

Departament de Filologia Espanyola
Facultat de Lletres

UNIVERSITAT AUTÒNOMA DE BARCELONA

INDICE

| | pág. |
|--|------|
| 0. INTRODUCCION..... | 5 |
| 1. ALGUNAS CONSIDERACIONES GENERALES SOBRE LA ENTONACION | |
| 1.1. Definición..... | 7 |
| 1.1.1. Plano físico: parámetros que intervienen en la entonación..... | 8 |
| 1.1.2. Plano perceptivo: percepción de la entonación..... | 9 |
| 1.1.3. Plano funcional: funciones de la entonación..... | 10 |
| 1.2. El análisis de la entonación: del plano físico al plano funcional..... | 12 |
| 1.2.1. El análisis de las unidades entonativas: la "curva estilizada"..... | 12 |
| 1.2.1.1. El problema de la representación..... | 13 |
| 1.2.1.2. La eliminación de las fuentes de variación..... | 16 |
| 1.2.2. La búsqueda de la información funcional: los patrones melódicos..... | 21 |
| 1.2.2.1. Tipos de curvas melódicas..... | 22 |
| 1.2.2.2. La estructura de las curvas melódicas..... | 22 |
| 1.2.2.3. Elementos significativos dentro de las curvas melódicas..... | 23 |

| | |
|---|----|
| 1.2.2.4. Los patrones melódicos..... | 27 |
| 1.2.2.4.1. Tipos de patrones melódicos del español..... | 28 |
| 1.2.2.4.2. Forma típica de los patrones melódicos del español..... | 30 |
| 2. ANALISIS DE LA ENTONACION Y TECNOLOGIAS DEL HABLA..... | 34 |
| 2.1. Entonación y síntesis automática..... | 34 |
| 2.2. Entonación y reconocimiento automático..... | 36 |
| 3. UN SISTEMA DE ESTILIZACION DE CURVAS MELODICAS PARA EL ESPAÑOL..... | 40 |
| 4. ANALISIS Y CLASIFICACION DE LOS PATRONES MELODICOS DEL ESPAÑOL | |
| 4.1. Objetivo..... | 43 |
| 4.2. Acotación del objeto de estudio..... | 44 |
| 4.3. Procedimiento experimental..... | 44 |
| 4.3.1. Elaboración del <i>corpus</i> | 45 |
| 4.3.2. Grabación del <i>corpus</i> | 52 |
| 4.3.3. Análisis del <i>corpus</i> | 53 |
| 4.4. Resultados..... | 56 |
| 4.4.1. Frases enunciativas..... | 57 |
| 4.4.2. Frases interrogativas absolutas..... | 62 |
| 4.4.3. Frases interrogativas pronominales..... | 68 |
| 4.4.4. Frases interrogativas relativas..... | 73 |
| 4.4.5. Frases exclamativas..... | 77 |
| 4.4.6. Frases volitivas..... | 90 |
| 4.5. Clasificación de los esquemas entonativos obtenidos..... | 94 |
| 4.5.1. Patrones melódicos..... | 94 |
| 4.5.1.1. Patrón descendente..... | 96 |

| | |
|---|-----|
| 4.5.1.2. Patrón ascendente-interrogativo..... | 96 |
| 4.5.1.3. Patrón ascendente-enunciativo..... | 100 |
| 4.5.2. Formas superpuestas..... | 100 |
| 4.5.2.1. "Esquema ondulado"..... | 100 |
| 4.5.2.2. Aumento del rango frecuencial..... | 104 |
| 4.5.2.3. Esquema circunflejo..... | 104 |
| 4.5.2.4. Elevación de la altura tonal o la pendiente del primer pico..... | 104 |
| 4.5.2.5. Variación de la pendiente en el segmento final..... | 105 |
| 4.6. Incidencia de los resultados en la descripción de los patrones melódicos del español..... | 106 |
| 5. CONCLUSIONES GENERALES..... | 109 |
| 5.1. Aplicación de los resultados a la síntesis y el reconocimiento..... | 109 |
| 5.1.1. Aplicación a la síntesis..... | 111 |
| 5.1.2. Aplicación al reconocimiento..... | 111 |
| 5.2. Cuestiones pendientes..... | 112 |
| 5.2.1. Sistema de estilización..... | 112 |
| 5.2.2. Patrones melódicos..... | 113 |
| ANEXO..... | 115 |
| REFERENCIAS..... | 125 |

0. INTRODUCCION

Este trabajo presenta los primeros resultados de un estudio sobre la utilización de la información prosódica en sistemas de síntesis y reconocimiento de habla que se está desarrollando en el Laboratorio de Fonética de la Universidad Autónoma de Barcelona. Es una versión revisada del proyecto de investigación defendido, el mes de octubre de 1990, en el marco del programa de Tercer Ciclo "Lingüística e Informática" del Departamento de Filología Española de esta Universidad¹. Se trata de un proyecto multidisciplinar, tanto lingüístico como técnico, que está siendo llevado a cabo con la colaboración de Francesc Gudayol, ingeniero de Telecomunicaciones. No es un trabajo completo ni cerrado, puesto que, como se señala en las conclusiones, todavía quedan bastantes interrogantes que resolver. Lo que se presenta aquí quiere ser sobre todo la definición de una línea de trabajo y de un método de análisis que, sin duda, aún debe ser verificado y refinado experimentalmente².

Los capítulos que siguen pueden agruparse en dos bloques principales. El primero, más teórico, comprende los capítulos 1 y 2, y constituye ante todo la definición de la línea de trabajo que se inicia con este estudio. Estos capítulos incluyen un planteamiento general de los problemas que presenta el análisis de la entonación, a la vez que un repaso de los estudios realizados hasta el momento, tanto generales como para el español en particular. No pretende ser una presentación exhaustiva de todos los aspectos relacionados con la entonación, sino únicamente de aquéllos que tienen que ver directamente con este trabajo. En el segundo bloque (capítulos 3 y 4) se presentan los resultados del trabajo realizado hasta ahora: en primer lugar (capítulo 3), se propone un procedimiento de estilización para curvas melódicas del español utilizable en sistemas de síntesis y reconocimiento de habla;

¹ Mi más sincero agradecimiento a la doctora Dolors Poch, catedrática de Lingüística General en la Escuela Universitaria de Traductores e Intérpretes de la Universidad Autónoma de Barcelona, y al doctor Climent Nadeu, profesor de la Escuela Técnica Superior de Ingenieros de Telecomunicaciones de la Universidad Politécnica de Cataluña, ambos miembros del tribunal, por sus comentarios y sugerencias acerca de este trabajo. Mi agradecimiento también a Francesc Gudayol, y a Lourdes Aguilar por su colaboración en la revisión del texto. Y muy especialmente al doctor Joaquim Llisterrí, director del proyecto, por su gran ayuda y por la paciencia demostrada en las continuas revisiones del texto.

² Este trabajo está siendo llevado a cabo con el soporte de una ayuda a investigadores jóvenes de la CIRIT de la *Generalitat de Catalunya* (DOGC 11.02.90).

en el capítulo 4 se presenta una clasificación de patrones melódicos y otros recursos que se utilizarían en frases del español para la expresión de la modalidad oracional. En el capítulo 5, a modo de conclusión, se analizan las posibilidades de aplicación de los resultados en sistemas de reconocimiento y síntesis, y se plantean cuestiones pendientes que serán la base de futuros trabajos.

1. ALGUNAS CONSIDERACIONES GENERALES SOBRE LA ENTONACION

1.1. Definición

La entonación es ante todo un **fenómeno lingüístico**, relacionado con la sensación perceptiva que produce la variación **a lo largo de todo un enunciado** de tres parámetros físicos - frecuencia del fundamental (en adelante, Fo), amplitud y duración -, y que proporciona al receptor información de distintos tipos. A diferencia del acento, las relaciones y contrastes entre las variaciones de los tres parámetros se darían en **unidades mayores que la palabra, normalmente sintagmas o frases**.

De acuerdo con esto, la entonación puede entenderse como un fenómeno que pone en relación tres niveles diferentes (´t Hart y Collier, 1975, pp. 235-238):

- 1) En primer lugar, el **plano físico** (o acústico): en este sentido, la entonación es el resultado de la variación temporal de una serie de parámetros físicos.
- 2) A continuación, el **plano perceptivo**: el oído humano actúa como un filtro que en cierta medida transforma la señal sonora que llega a él, desechando algunas de las variaciones de esos parámetros físicos.
- 3) Finalmente, el **plano semántico-funcional**: el oyente extrae de estas variaciones, convenientemente filtradas, diversas informaciones de tipo lingüístico, o incluso extralingüístico.

Un análisis de la entonación ha de intentar, como objetivo último, poner en relación el plano físico - las variaciones de los parámetros acústicos - con el plano funcional - la información que se transmite con la entonación. Este proceso ha de tener en cuenta necesariamente el plano perceptivo, para poder desechar aquellas variaciones de estos parámetros físicos que no están relacionadas con la transmisión de la información entonativa.

1.1.1. Plano físico: parámetros que intervienen en la entonación

Tradicionalmente, se ha señalado que las variaciones de la Fo, juntamente con las de la amplitud y duración de los elementos segmentales, son los tres parámetros acústicos responsables de la entonación. Estos tres parámetros son los que intervienen también en otros fenómenos suprasegmentales, como el ritmo o el acento. Como se señala más adelante, esto hará que en ocasiones sea difícil atribuir la variación de un parámetro determinado a un fenómeno u otro.

Bolinger (1955) presenta casos en inglés en los que la sensación de entonación no es resultado de un cambio en la Fo, sino de la amplitud con que se emite una determinada porción del grupo fónico. Y los trabajos de Denes (1959; 1962) demuestran experimentalmente esta interrelación. Igualmente, por poner un ejemplo del español, la R.A.E., en su *Esbozo* de 1973 afirma que ciertos tipos de frases - las órdenes y los ruegos, concretamente - se expresan mediante cambios en la amplitud de algunos segmentos, además de los cambios en la curva melódica: "Las frases de solicitud y sus réplicas se desarrollan, en general, en formas melódicas y con recursos fonéticos (tensión articulatoria, intensidad del acento, etc.), además del tono, más suaves y atenuadas que las frases en que se da una orden" (R.A.E, 1973, pág. 118).

Con todo, se considera que en la mayoría de los casos las variaciones de la Fo son el responsable físico principal del fenómeno de la entonación. De ahí que los principales estudios - Rossi *et al.* (1981), Thorsen (1979; 1980), 't Hart *et al.* (1990), por mencionar los más conocidos - se hayan centrado únicamente en el análisis de la forma y variaciones de las curvas de Fo. Incluso en bastantes de las definiciones que se han dado de la entonación se considera la Fo como el único parámetro que interviene en la producción de la misma. Así ocurre, por ejemplo, en la definición que dan Alcina y Blecua - "línea melódica con que se produce un mensaje" (Alcina y Blecua, 1975, pág. 452) -, o en la que da León - "*variations melodiques de l'enoncé, perçues par l'auditeur*" (León, 1970, pág. 58).

Sin embargo, de acuerdo con la definición de entonación establecida al comienzo del apartado 1.1. - variaciones no sólo de la Fo, sino también de la amplitud y duración de los elementos segmentales que se dan en el ámbito del **grupo fónico** -, consideraremos en este trabajo melodía y entonación como dos fenómenos diferentes, aunque relacionados entre sí. Así, definiremos melodía (*pitch* en inglés) como el fenómeno que se relaciona con la curva de Fo o **curva melódica** de un grupo fónico.

Dado el mayor peso de este parámetro en el fenómeno de la entonación, los estudios que se presentan en este trabajo se han centrado únicamente en el análisis de curvas de Fo. Por ello, los trabajos que se citan en este capítulo son principalmente los que han abordado el estudio de dichas curvas.

1.1.2. Plano perceptivo: percepción de la entonación

No todas las variaciones en la frecuencia de la F_0 son percibidas por el oído humano. Se han realizado estudios en este sentido para determinar cuál es la variación mínima que el oído puede percibir: son los llamados en inglés *just noticeable differences* (JND), o **umbrales diferenciales**, de percepción de la F_0 .

Tres son los estudios principales realizados en esta línea: el de Flanagan y Saslow (1958), el de Klatt (1973), y el de 't Hart (1974). Los tres difieren un tanto en las condiciones en que se realizaron, por lo que sus resultados son también algo diferentes.

a) Flanagan y Saslow emplean vocales sintetizadas con una F_0 constante para la realización de tests de percepción. Estos tests indicaron que el umbral se situaba entre los 0,3 y los 0,5 Hz.

b) Klatt también utiliza vocales sintetizadas, pero no sólo con la F_0 constante, sino además con pendiente ascendente y descendente. La justificación que Klatt da para ello es que en el habla real la F_0 no es siempre igual, sino que va variando continuamente en el tiempo. Sus resultados con vocales de tono constante confirman los de Flanagan y Saslow (0,3 Hz). En las vocales con tono descendente, sin embargo, este umbral es más alto, situándose en los 2 Hz.

c) 't Hart continúa la línea de investigación de Klatt, profundizando en el estudio de la percepción de tonos ascendentes y descendentes. Sus resultados indican que en contorno ascendente el umbral se sitúa en 1,5 semitonos, mientras que en contornos descendente, el umbral aumenta hasta los 3 semitonos³.

³ Tampoco está aún muy claro cuál es el mecanismo mediante el cual el oído capta la sensación de tono o *pitch*. Se han postulado diferentes teorías, que pueden agruparse en dos grandes bloques:

a) El primero incluiría aquellas teorías que consideran que la sensación de melodía es el resultado de comparar la distancia entre las excitaciones producidas por cada armónico en puntos diferentes de la membrana basilar. Estas teorías reciben el nombre genérico de **modelos de comparación de patrones** (*pattern recognition models*). Dentro de este grupo estarían las propuestas de Thurlow (1963), Whitfield (1967; 1970) y Goldstein (1973), entre otros.

b) El segundo grupo reúne a las teorías que se pueden denominar, de forma general, **modelos temporales**. La idea central de estos modelos es que la percepción del *pitch* es el resultado de comparar las diferencias en el tiempo de las excitaciones que se producen, a intervalos periódicos, en la membrana basilar. El principal defensor de este modelo es Schouten (1940; 1970).

Otro aspecto importante, que sin duda ha sido menos tratado, es la determinación de las **variaciones relevantes desde el punto de vista entonativo**, es decir, las que afectan a la transmisión de la información contenida en la entonación. Quilis señala (Quilis, 1981, pág. 359) citando a su vez a León (1971b), que en la percepción de la Fo “deben contar tanto los factores físicos como los factores lingüísticos”. Así, por ejemplo, la posición en que se dé una variación dentro de la curva influirá también en su percepción: si se da en un punto altamente significativo, bastará un cambio mínimo para que se perciba una diferencia, y viceversa.

Hay que destacar en esta línea los trabajos de Thorsen (1980), para el danés, y de ‘t Hart y su grupo, para el holandés. Sin embargo, el análisis de las variaciones de Fo relevantes desde este punto de vista es aún una cuestión pendiente en el caso del español.

1.1.3. Plano funcional: funciones de la entonación

En autores como Quilis (1981, pp. 377-395) o Martin (1973), entre otros, podemos encontrar clasificaciones de las diferentes funciones de la entonación. Con alguna variación en las categorías, las dos coinciden en que la entonación transmite los siguientes tipos de información:

1) Información de tipo **lingüístico**, referente al mensaje, que es generalmente información **sintáctica**. Esta información sintáctica puede ser principalmente de dos clases:

a) La entonación sirve, en primer lugar, para integrar y delimitar las diferentes frases o sintagmas de una frase, haciendo que formen una unidad y separándolas del resto de frases del discurso. Es lo que algunos autores - Quilis, por ejemplo (1981, pág. 384 y ss.), siguiendo a Rossi (1977) - denominan **funciones integradora y delimitadora**. El estudio de esta función de la entonación ha sido el principal objetivo de los trabajos de Martin (1982), Rossi (1981; 1985) o Di Cristo (1978), para el francés, o de Bolinger, en un reciente estudio (1989), para el inglés. Sobre la relación entre sintaxis y entonación en español, merece destacarse el trabajo general de Fant (1984).

Hasta el momento, ninguna de las dos teorías ha sido capaz de explicar todos los fenómenos observados experimentalmente. Es posible incluso que las dos explicaciones sean complementarias. Lo que parece claro es que la Fo no es la responsable directa de la sensación de melodía, sino los diferentes armónicos (Schouten, 1940), y en especial el tercero, cuarto y quinto, según sugieren los trabajos de Terhardt (1972a; 1972b). Una prueba de ello es que la entonación se percibe también a través del teléfono, en el que la Fo está ausente.

b) Igualmente, la entonación permite distinguir entre enunciados no finales, situados al principio o en medio de una oración, y los situadas al final de las mismas (para una distinción más detallada, *vid.* apartado 1.2.2.1). Asimismo puede aportar información sobre el tipo de oración de que se trata (la **modalidad oracional**), oponiéndola de este modo al resto de oraciones del discurso. Esta función, que Quilis (1981, pág. 381 y ss.) denomina **distintiva**, es la que ha recibido más atención en los estudios entonativos del español - en Navarro (1948), por ejemplo.

2) Información no estrictamente lingüística:

a) De tipo **sociolingüístico**, acerca de la procedencia geográfica y social del hablante. Si éste es un extranjero, también puede deducirse de su entonación la lengua materna del mismo. Se trata de un campo en el que no se ha entrado todavía en profundidad: Quilis (1981), cita algunos estudios para el francés, como los de Szmidt (1968), o Martinet (1956), y también para el español, como el de Canellada (1941) sobre el habla extremeña, además de algunos estudios sobre variantes locales hispanoamericanas. Sobre las variaciones en los esquemas entonativos de un hablante provocadas por la interferencia de otra lengua, puede encontrarse un repaso al estado de la cuestión en De Bot (1986, pp. 115-17).

b) De tipo **expresivo**, acerca del estado de ánimo del hablante. Es el campo de estudio de la **fonoestilística**, en el que destacan las aportaciones de León (1971a, entre otros).

Respecto a esta clasificación, sin embargo, cabe hacer una matización: no parece muy claro a cuál de los dos grupos pertenece el plano **expresivo**: puede considerarse un tipo de información lingüística más (de hecho, la modalidad expresiva se ha considerado tradicionalmente un tipo más del modo oracional), o bien una información suplementaria acerca del individuo. Incluso podría considerarse que es un nivel intermedio que se solapa con los otros dos. La mayoría de las gramáticas del español, sin embargo, consideran la existencia de frases con modalidad expresiva -*vid.*, entre otros, R.A.E. (1973, pág. 357-8) -, y los estudios sobre la entonación del español hablan en muchos casos de patrones melódicos expresivos - Navarro (1948, pp. 152 y ss.), por ejemplo.

También hay que señalar que los diferentes tipos de funciones pueden superponerse, coexistir en una misma oración. Por tanto, un problema suplementario que se plantea es distinguir los correlatos correspondientes a cada función.

1.2. El análisis de la entonación: del plano físico al plano funcional

Para estudiar la relación entre las variaciones de la Fo en un determinado enunciado y la información que un oyente extrae de dichas variaciones pueden seguirse, en principio, dos caminos:

a) El del **análisis acústico**: las curvas melódicas se analizan acústicamente y se extraen sus características generales; se intenta hacer una descripción de los distintos esquemas haciendo referencia a las variaciones de los parámetros que más se repitan en los distintos casos estudiados. Ejemplos de este método serían los estudios de Thorsen (1978) o de O'Shaughnessy (1979).

b) El del **análisis perceptivo**: la descripción se realiza haciendo hincapié en aquellos indicios que son relevantes desde el punto de vista perceptivo. Implica normalmente la realización de tests de percepción. Es una línea por el momento menos seguida, en parte debido a dificultades de tipo técnico (hasta hace poco se carecía del instrumental adecuado: sistemas de generación y manipulación artificial de melodía, por ejemplo). Hay grupos, sin embargo, que llevan trabajando casi 25 años en esta línea, como es el caso de 't Hart y sus colaboradores, en sus estudios de la entonación del holandés - Cohen y 't Hart (1967); 't Hart (1979). También estudios como el de Studdert-Kennedy y Hadding (1973) para el inglés han ido en esta línea.

De hecho, ambos procedimientos pueden entenderse como dos fases en el proceso de relacionar los planos físico y funcional de la entonación: en la primera se analizarían los parámetros acústicos, y se extraerían las características más relevantes, los esquemas más repetidos, a la vez que se eliminarían los aspectos que se considerasen irrelevantes. Después, el análisis perceptivo confirmará o negará la validez de los esquemas establecidos, y permitirá saber si en dichos esquemas se ha preservado la información entonativa. Esta es, como se indica más adelante, la línea que se pretende seguir en el presente trabajo.

1.2.1. El análisis de las unidades entonativas: la "curva estilizada"

En Rossi *et al.* (1981) se establecían nueve factores diferentes que pueden determinar la forma que toma una curva de Fo. En un estudio posterior, Jassem y Demenko (1986) ampliaron este número hasta catorce, que son los que se relacionan a continuación:

- 1) Condiciones dependientes de la actitud del hablante, situacionales y del discurso.
- 2) Situación del acento temático.
- 3) Realización del acento léxico o gramatical.
- 4) Longitud de la curva, condicionada segmental o léxicamente.
- 5) Variación alotónica libre.
- 6) Efectos de estados emocionales.
- 7) Rasgos fisiológicos personales del hablante.
- 8) Rasgos patológicos de la voz.
- 9) Cambios momentáneos en el tono medio (confidencias, frases parentéticas...).
- 10) *Tempo* personal de elocución.
- 11) Cambios momentáneos del *tempo*.
- 12) Efectos de los rasgos segmentales.
- 13) Irregularidades no patológicas en la fonación.
- 14) Estilo.

Ahora bien, de todas estas fuentes de variación, hay algunas que transmiten informaciones relevantes para el receptor, y otras que no. La cuestión que se plantea a la hora de abordar el análisis de la entonación es determinar cuáles son significativas y cómo eliminar de la representación de la entonación aquellas que no lo sean.

El análisis de la entonación implica, por tanto, un proceso de **estilización** de la curva de la Fo. Este concepto no es nuevo, y de hecho está implícito en cualquier aproximación al análisis de curvas de Fo. Sólo en algún caso, sin embargo, se ha definido de forma sistemática y explícita el procedimiento de estilización utilizado. 't Hart y su grupo, por ejemplo, han desarrollado el concepto de "*Close copy stylizations*", que definirían como representaciones de la curva melódica formadas con "*the minimum number of straight lines that are necessary to imitate the original utterance without any audible difference*" (Collier, 1989, pág. 39). En este caso, los criterios de estilización son perceptivos, pero también pueden definirse sistemas de estilización basados en criterios acústicos, como el presentado en Rossi *et al.* (1981).

La estilización de curvas de Fo conlleva una serie de problemas que hay que solventar, y que se tratan a continuación.

1.2.1.1. El problema de la representación

La primera abstracción que se realiza en cualquier aproximación al estudio de las curvas melódicas va implícita en la forma de representación escogida para las mismas. La elección de un método en particular implica, como se verá a continuación, que se está dando preponderancia a unos aspectos determinados de la curva.

Podemos encontrar en la bibliografía dos aproximaciones fundamentales al análisis de los grupos entonativos: el sistema de niveles y el sistema de contornos.

a) **Análisis por niveles:**

Es la representación propia de la tradición americana, aunque también ha habido fonetistas europeos que la han seguido. Se considera la entonación como una serie de elementos discretos, los **tonos** o **niveles**, que se van yuxtaponiendo a lo largo de la frase. En este sistema de representación se da preferencia a la altura tonal de cada segmento, más que a las diferencias de tono entre uno y otro.

Un ejemplo de representación mediante niveles sería el modelo que presentan León y Martín (1969, pág. 52) para el francés. En este artículo, establecen en cuatro el número de tonos que pueden encontrarse en francés, y los criterios que hay que seguir para determinarlos en cada hablante. Más tarde, León (1970), propuso la existencia de un quinto tono, por encima de los otros cuatro, que sería típico de las oraciones con contenido expresivo. Dichos tonos serían los siguientes:

- a) **Tono 1:** el más bajo de todos, sería el propio de los incisos y paréntesis, y en general de los segmentos terminales.
- b) **Tono 2:** sería el propio de los inicios de las frases enunciativas.
- c) **Tono 3:** tono propio de los grupos de continuidad.
- d) **Tono 4:** se encontraría en los inicios de las exclamaciones, las interrogaciones y las órdenes.
- e) **Tono 5:** tono propio de las frases con contenido emocional o expresivo⁴.

También se ha utilizado bastante el sistema de niveles para la representación de las curvas melódicas del inglés (véase en el apartado siguiente, por ejemplo, la descripción del sistema de Pierrehumbert).

Para el español, podemos encontrar descripciones de los esquemas melódicos realizadas con este sistema en Stockwell *et al.* (1956), en el *Esbozo* de la R.A.E. (1973, pp. 105-6) y en Quilis (1981).

Tanto Stockwell como Quilis distinguen sólo tres niveles, **bajo** (1), **medio** (2) y **alto** (3), sin dar ninguna explicación sobre dónde aparecería cada uno.

⁴ Una alternativa al sistema de León y Martín para el francés sería el propuesto en Rossi *et al.*, (1981, pág. 76), que considera la existencia de hasta 6 niveles tonales diferentes: "0 (*Médium*); +1 (*Infra-Aigu*); +2 (*Aigu*); +3 (*Sur-Aigu*); -1 (*Grave*); -2 (*Infra-Grave*)".

En el *Esbozo* de la R.A.E. (1973), por su parte, se distinguen cuatro tonos diferentes:

- a) Tono 1: sería el **tono grave** "unos cuatro semitonos por debajo del 2".
- b) Tono 2, o **tono semigrave**, cuatro semitonos por debajo del 3.
- c) Tono 3, o **tono normal**, que no especifica dónde se situaría.
- d) Tono 4, o **tono agudo**, cuatro semitonos por encima del 3.

El principal problema que plantea el uso de este sistema es que resulta muy difícil llevarlo a la práctica: al analizar una curva, es bastante complicado determinar en qué nivel se halla cada segmento de la misma, y dónde se da el paso de un nivel a otro. Además, al estar limitado a unos pocos tonos el número de unidades, la cuantificación hace que se puedan perder variaciones que quizá son relevantes.

Los mismos León y Martín (1969) señalan la relatividad de este sistema, basándose en los resultados de sus investigaciones realizadas en el Laboratorio de Fonética de Toronto.

La representación de la curva melódica con notación musical es, de hecho, otra forma de representación en niveles, y presenta los mismos problemas que ésta. Navarro, en su estudio de la entonación castellana (1948) utiliza los tonos, semitonos y octavas para expresar los cambios en el Fo durante una frase. Con todo, parece un sistema de notación más acertado que el de unos pocos niveles, puesto que se trata de una escala mucho más discretizada.

b) **Análisis por contornos:**

Este tipo de representación de las curvas entonativas es propio de la tradición europea, aunque su máximo defensor es el norteamericano Bolinger. La idea básica que sustenta este tipo de análisis es que la percepción de la entonación se basa fundamentalmente en los contrastes entre segmentos dentro del mismo grupo, más que en la altura tonal que alcanza cada segmento. Bolinger (1951) se muestra partidario de este sistema, a la vista de los resultados de algunos experimentos realizados por él mismo y que describe en dicho artículo. Parte de la idea de que si la representación de niveles tiene una base real en que sustentarse, frases con el mismo contorno pero con el tono medio modificado habrían de percibirse de manera diferente. Por contra, el experimento de percepción que él realiza con frases de este tipo demuestra que los oyentes tienden a percibirlos como iguales.

Sin embargo, la afirmación de Bolinger no siempre se cumple. Como se señala más adelante, una elevación general de la F_0 en una frase, sin que la forma de la curva se altere, puede tener relevancia perceptiva, sobre todo en la transmisión de estados emocionales.

En esta línea cabría también incluir los sistemas de representación utilizados por Thorsen (1979, entre otros trabajos), y por 't Hart y su grupo - Cohen y 't Hart (1967), por ejemplo.

De acuerdo con lo expuesto, puede afirmarse que ninguno de los dos sistemas es capaz de representar correctamente por sí solo todas las variaciones de una curva entonativa. Por ello Danes (1960) propuso un sistema de representación por **configuraciones de niveles**, que combinase los dos anteriores.

Podría decirse que, en principio, los sistemas de niveles son más adecuados para aquellas aproximaciones que toman la sílaba como unidad de análisis, en tanto que los contornos serían más apropiados para los estudios que analizan la curva de la frase en su conjunto.

1.2.1.2. La eliminación de las fuentes de variación

Tal como se ha señalado antes, un análisis de la entonación supone dejar de lado una serie de rasgos presentes en la curva melódica, o en las variaciones de los otros dos parámetros, que son producto de "fuentes de variación" irrelevantes desde el punto de vista lingüístico. En este apartado se hace mención de algunas de estas fuentes de variación, y de las propuestas que se han hecho para su tratamiento.

a) La aparición de **segmentos sordos** a lo largo de la frase provoca interrupciones en la curva melódica. El contorno de la misma de alguna manera queda cortado, y sin embargo, los oyentes no parecen tener una pérdida de información entonativa por ello. Parece perfectamente válido considerar que el receptor realiza un proceso de reconstrucción del contorno en las interrupciones producidas en las curvas de F_0 por los segmentos sordos de una frase. Sin embargo, no hay todavía datos experimentales que apoyen esta idea.

b) Como ya se ha indicado, el **acento** introduce variaciones no sólo en la curva melódica, sino en la amplitud y la duración de los segmentos de un enunciado. Nos remitimos de nuevo al estudio de Bolinger (1955) para conocer las interferencias entre acento y entonación en inglés. De hecho, algunas de las aproximaciones al estudio de las curvas melódicas del inglés consideran éstas como el resultado de la suma de componentes acentuales y entonativos. Una

muestra la encontramos en el modelo descrito en Pierrehumbert (1987). En dicho modelo, que utiliza representaciones por niveles, la curva melódica es el resultado de la interacción de tres fenómenos diferentes:

a) Los llamados **tonos de límite** (*boundary tones*), que aparecen sólo al principio o al final de un grupo. Establece dos tonos de límite diferentes, uno alto (representado como H%) y otro bajo (L%).

b) El **acento de frase** o *phrase accent*, que se encontraría "*near the end of the word with the nuclear stress*" (Pierrehumbert, 1987, pág. 15). Los niveles posibles para este acento de frase serían también dos, uno alto (H⁻) y otro bajo (L⁻).

c) Los **acentos de tono** o *tone accents*, que pueden adoptar hasta siete formas diferentes: H*, L*, L*+H⁻, L⁻+H*, H*+L⁻, H⁻+L*, H*+H⁻. Se situarían en las diferentes sílabas acentuadas del grupo, con excepción de la que contiene el acento de frase.

Una vez asignado el nivel correspondiente de Fo al tono de límite y a cada punto del enunciado marcado con algún tipo de acento, la curva melódica se construiría mediante interpolación entre los distintos puntos.

En español, los diferentes trabajos indican que las principales inflexiones de la curva melódica suelen coincidir con sílabas tónicas, especialmente en la primera y la última del grupo fónico. No está tan claro, sin embargo, que el acento en español se traduzca siempre en un aumento de la Fo en la sílaba tónica. Según los datos de Gili Gaya (1924, pág. 167), las variaciones que introduce el acento en español pueden llegar a ser importantes (90 Hz por término medio en el sujeto de su estudio) en palabras aisladas. Sin embargo, su estudio de un discurso continuo revela que "el acento de intensidad no siempre coincide con la mayor elevación de tono cuando nos hallamos en la rama ascendente o descendente de la curva", en tanto que, en la parte central, "el acento de intensidad determina una elevación de tono mayor o menor" (Gili Gaya, 1924, pág. 170). El estudio presentado por Enríquez *et al.* (1989), realizado con estímulos sintetizados, demuestra la importancia de la Fo en la percepción del acento en palabras aisladas. En habla continua, sin embargo, y de acuerdo con lo expuesto por Gili Gaya, el acento puede manifestarse mediante cualquiera de los tres parámetros citados. La influencia del acento en español, de acuerdo con esto, parece limitada a unos puntos determinados de la curva melódica, aunque faltan estudios que confirmen y concreten esta idea. Se trata sin duda de una cuestión abierta todavía.

c) También pueden influir en la curva de F_0 los elementos segmentales. Cada sonido llevaría asociado un "fundamental intrínseco", que incidiría en la forma general de la curva. Hombert (1978) presenta un resumen de los estudios realizados sobre el tema: para las vocales, cita los trabajos de Black (1949), de Lehiste y Peterson (1961), el de Peterson y Barney (1952), y el de House y Fairbanks (1953). También merece destacarse el estudio más reciente de Di Cristo (1982). Todos llegan a conclusiones semejantes: "*High vowels (low F_1) have a higher fundamental frequency than low vowels*" (Hombert, 1978, pág. 97). Las diferencias pueden llegar a ser de hasta 20 Hz, entre vocales diferentes, según los datos aportados por Lehiste y Peterson (1961): su estudio da unos valores medios que se sitúan por encima de los 180 Hz en el caso de [i,u], que serían los más altos, y algo superiores a los 160 Hz en [a], que sería el valor más bajo⁵. El estudio de Mateo (1988) presenta unos resultados semejantes con vocales del español.

Una buena revisión de los estudios sobre el fundamental intrínseco en las consonantes sonoras puede encontrarse en Di Cristo (1982). A juicio de Di Cristo, los trabajos más completos sobre el tema son los llevados a cabo por Boë (Boë, 1973; Larreur y Boë, 1973). Los resultados de sus análisis indican que, en general, todas las consonantes sonoras presentan una F_0 más baja que la vocal adyacente, y que la forma de la curva y la pendiente de ésta durante la consonante puede variar de una a otra. Di Cristo (1982, pp. 87-110) presenta también los resultados de un estudio semejante, realizado por él mismo, que viene a confirmar las conclusiones de Boë. Los valores medios que da Di Cristo, para el descenso de la F_0 en las consonantes con respecto a la vocal adyacente, oscilan entre los 2 y los 6 cuartos de tono. Por su parte, Gili Gaya (1924) señala para el español que las consonantes no parecen tener una F_0 inherente, sino que es simplemente el resultado de la interpolación entre las F_0 de las vocales adyacentes, sin especificar nada más. Hoy en día, aún se carece de datos precisos sobre el fundamental inherente en las consonantes del español.

Igualmente, se ha demostrado que las consonantes que anteceden o siguen a una vocal pueden hacer variar el valor de su F_0 , y, consecuentemente, la forma

⁵ Se han formulado diferentes hipótesis para explicar este fenómeno, que resumimos a continuación:

- a) Flanagan y Landgraf (1968) lo atribuyen a motivos acústicos: según ellos, se produciría un acoplamiento entre la fuente glotal y el tracto vocal.
- b) Gandour y Weinberg (1980), achacan el fenómeno a una mayor presión subglótica cuando se pronuncia una vocal alta.
- c) Lehiste (1970) opina que se produce un aumento de la tensión en las cuerdas vocales por la elevación de la lengua al pronunciar las vocales altas.

de la curva melódica. El artículo de Hombert (1978) incluye también un resumen del estado de la cuestión, en el que se mencionan los trabajos ya citados de House y Fairbanks (1953) y de Lehiste y Peterson (1961), junto con resultados de experimentos propios. También Di Cristo (1982) presenta un análisis experimental de este fenómeno. Los resultados de estos estudios indican que, en general, "*Fo values of vowels after voiceless aspirated stops are higher than after voiced stops*" (Hombert, 1978, pág. 79).

Para el español, Gili Gaya (1924) analizó igualmente las diferencias en el Fo de las vocales según sea sorda o sonora la consonante adyacente. Los resultados se oponen a los presentados por Hombert: "por lo que se refiere a las consonantes sordas, el hecho de determinar interrupción en la curva suele influir en el tono de la vocal siguiente, puesto que las cuerdas vocales al empezar a vibrar de nuevo producen para vencer su inercia un sonido más grave que puede afectar aún a las vocales acentuadas" (Gili Gaya, 1924, pág 177). Como en el caso del fundamental inherente, son necesarios más estudios experimentales para confirmar o desmentir estas afirmaciones.

Finalmente, el punto de articulación de las consonantes adyacentes sería también otro factor que influiría en la Fo de una vocal. Mateo, en su estudio ya citado sobre la Fo de las vocales del español (1988) demuestra experimentalmente esta influencia.

Todas estas variaciones, denominadas genéricamente **micromelódicas**, no afectan en exceso a la percepción de la curva de Fo, y tal como afirma 't Hart (1979) "*apparently are not relevant to the perception of intonation at a global level*", aunque esta afirmación no ha sido comprobada experimentalmente de forma definitiva - *vid.* Di Cristo (1982, pp. 84-86 y 142-47) para un repaso de los diferentes estudios realizados sobre el tema. No parece necesario, por tanto, que las variaciones micromelódicas queden registradas en una representación estilizada de una curva de Fo.

En Rossi *et al.* (1981) se presentan dos procedimientos de estilización que buscan normalizar la altura tonal, la amplitud y la duración de los núcleos sílabicos, eliminando las variaciones micromelódicas. El segundo de ellos, desarrollado por Nishinuma y Rossi, comprende las siguientes etapas:

- a) Un proceso de adquisición de los datos acústicos, por el que se obtienen los valores de Fo, amplitud y duración de los distintos sonidos vocálicos a lo largo del enunciado.
- b) Una fase de tratamiento prosódico de los datos obtenidos en la etapa anterior, que comprendería una normalización de la duración, la amplitud y la Fo en las distintas vocales. Dicha normalización se realiza relativizando los valores obtenidos con respecto a los valores estándar de [a] para cada parámetro, y eliminando en segundo lugar la influencia de las consonantes adyacentes. También se eliminan las variaciones que se dan dentro de cada

vocal para cualquiera de los tres parámetros (“glissando”).

Mediante este procedimiento se obtiene una serie de alturas tonales para cada núcleo vocálico, que permiten generar una representación en niveles de cualquier enunciado. Sin embargo, para aplicar este procedimiento es necesario realizar previamente amplios estudios acústicos que permitan conocer los valores medios de Fo, amplitud y duración para cada vocal.

d) Las variaciones existentes entre los distintos locutores, especialmente si son de diferente sexo, son otro factor de variabilidad bastante importante. La frecuencia de la Fo en las mujeres y en los niños es considerablemente más alta que en los hombres (entre 200 y 500 Hz en los primeros; entre 90 y 175 Hz en los segundos). A esto hay que añadir que cada locutor tiene su "fundamental habitual" - *fundamental usuel* según la terminología de Szmidt (1979) -, por lo que carece de sentido hacer un estudio de los valores absolutos de Fo a lo largo de la curva, sino más bien de las **diferencias** entre dichos valores en puntos relevantes de la misma. O bien hay que realizar algún tipo de **normalización** que iguale de alguna manera los valores de Fo de todos los locutores.

Pierrehumbert (1987, pp. 25-28) presenta un sistema de normalización, que toma como punto de referencia una línea imaginaria formada por “*the lowest value he [el locutor] is disposed to produce at the given location in the utterance*” (Pierrehumbert, 1987, pág. 25). A esta línea imaginaria le da el nombre de *baseline* o **línea de base**, y afirma que tiene una pendiente ligeramente descendente a lo largo de la frase. El cálculo de la misma lo realiza experimentalmente, haciendo pronunciar la misma frase a distintos locutores y con diferentes tonos de voz. Una vez conocida esta línea de base, se aplica la siguiente fórmula para escalar los valores de una curva determinada con respecto a la misma:

$$\hat{P} = \frac{(P - B)}{B}$$

P = valor de Fo en un punto determinado de la curva.

B = valor de Fo en el punto de la línea de base correspondiente a P.

\hat{P} = valor de Fo escalado.

Este sistema implica la realización de un largo entrenamiento previo que permita conocer la línea de base del locutor.

Takefuta (1975) propone también un sistema para normalizar los valores de Fo, que toma como referencias el valor más alto de Fo y el más bajo registrados en un conjunto de frases analizadas para un locutor determinado. La fórmula que utiliza es la siguiente:

$$VN_n = \frac{V_n - \min}{\max - \min}$$

VN_n = valor normalizado de F_0 en un punto determinado.

V_n = valor de F_0 en un punto normalizado.

\min = valor más bajo de F_0 encontrado en todas las frases pronunciadas por el hablante.

\max = valor más alto de F_0 encontrado en todas las frases pronunciadas por el hablante.

Al igual que el método de Pierrehumbert, se trata de un sistema que precisa de información previa sobre las variaciones de F_0 de cada hablante conocido.

1.2.2. La búsqueda de la información funcional: los patrones melódicos

La información que transmite la entonación se reparte, como se ha señalado al principio, a la largo de unidades superiores a la palabra, normalmente sintagmas o frases, que son las unidades lingüísticas contenidas en los llamados **grupos fónicos**, o enunciados comprendidos entre dos pausas del discurso. Un análisis de la entonación implica, por tanto, estudiar las variaciones de los parámetros entonativos a lo largo de todo el grupo.

La información funcional se ha buscado tradicionalmente en la forma general de las curvas melódicas, o de algunas partes determinadas de éstas. Así, se ha asociado a cada tipo de información una forma típica de curva, o **patrón melódico**. Los patrones melódicos representan esquemáticamente la forma de las curvas de F_0 , y podrían identificarse con las representaciones estilizadas obtenidas tras la eliminación de las variaciones irrelevantes.

La curva melódica se constituye así en una unidad básica para el estudio de la entonación. En los subapartados siguientes se analizan los diferentes tipos de curvas melódicas y la estructura que éstas suelen presentar.

1.2.2.1. Tipos de curvas melódicas

Las curvas melódicas pueden dividirse en dos grandes clases, en función del tipo de grupo fónico al que pertenezcan. Los grupos fónicos serán **terminales** o **no terminales** según la posición en que aparezcan dentro de la oración:

- a) Los **grupos terminales** serían los que aparecerían al final de una oración. Contendrían básicamente información sobre el tipo de oración - lo que tradicionalmente se ha denominado la **modalidad oracional**.
- b) Los **grupos no terminales** serían los que aparecerían, en oraciones con más de un grupo fónico, en posición no final dentro de las mismas. La información que llevan relacionada se refiere básicamente al **tipo de sintagma o frase subordinada**, según los casos, contenido en su grupo fónico.

1.2.2.2. La estructura de las curvas melódicas

La bibliografía nos presenta diferentes opiniones sobre la estructura de las curvas melódicas:

- a) Cruttenden (1986, pág. 35), y en general los estudios realizados para el inglés - *vid.* también O'Connor y Arnold (1961, pág. 13 y ss.) -, consideran que los grupos fónicos están compuestos por dos partes, divididas por un **núcleo**. Dicho núcleo se define como la sílaba principal del grupo, la que lleva el acento principal, y que no tiene que ser necesariamente la última sílaba acentuada. Así, define el tono nuclear (*nuclear tone*) como el segmento que va desde el núcleo de un grupo entonativo hasta el final del mismo.

En función de esto, la curva melódica de un enunciado se dividiría en dos partes, la anterior y la posterior al núcleo.

- b) Sin embargo, las descripciones realizadas para el español, tal como señalan Alcina y Blecua en su gramática (1975, pág. 458), suelen distinguir tres partes dentro de un grupo fónico, que implicarían asimismo tres partes en su curva melódica:

- 1) La **rama inicial** de la curva, “formada por las sílabas átonas que llegan hasta el primer acento fuerte”.
- 2) El **cuerpo** de la unidad melódica, que “está formado por el conjunto de sílabas que comprenden la sílaba fuerte inicial hasta la sílaba inmediatamente anterior al último acento fuerte”.

3) La **rama final**, “que está integrada por la última sílaba fuerte y las siguientes débiles, en el caso de que las haya”.

La existencia de un acento de frase o tono nuclear en inglés ha llevado a algunos autores (Contreras, 1976) a afirmar la presencia de dicho acento en las frases del español. Otros, en cambio, como Fant (1984, pp. 48-57), consideran que la estructura entonativa difiere de unas lenguas a otras, y que dicho fenómeno no se da en las frases del español. Es posible que cada división sea adecuada para la lengua sobre la que se ha realizado el análisis. De acuerdo con esto, en español se distinguirían **dos núcleos entonativos**, uno al principio y otro al final del grupo, a diferencia del único núcleo que suele señalarse para el inglés.

Finalmente, al hablar de la estructura de las curvas melódicas, no se puede dejar de lado a la **sílaba**, como unidad considerada básica desde siempre en el análisis de la entonación. Como ya se ha indicado, hay efectos entonativos que se consiguen variando la amplitud y duración de sílabas determinadas de una frase. Y en las curvas melódicas, los núcleos entonativos suelen coincidir con sílabas acentuadas. Hay incluso quien ha considerado la sílaba como la unidad entonativa por excelencia, en lugar del grupo entonativo - Jassem (1952), entre otros.

1.2.2.3. Elementos significativos dentro de las curvas melódicas

La bibliografía ha tratado frecuentemente la cuestión de los elementos que contienen más información dentro de una curva melódica. Estos indicios pueden clasificarse según el tipo de información que transmitan:

A) Por lo que se refiere a la **información lingüística**, la mayoría de los estudios coinciden en considerar que el **final de la curva melódica** es el fragmento de la misma que contiene más información de este tipo. Un estudio bastante riguroso en este sentido es el realizado por Takefuta (1975), en el que aplica métodos estadísticos para determinar qué parte es la más informativa de la curva entonativa. Los resultados de su análisis corroboran la idea de que el final es la parte que contiene más información. También Studdert-Kennedy y Hadding (1973) llegan a una conclusión semejante, mediante el análisis perceptivo de curvas sintetizadas.

Esta porción final ha recibido diferentes nombres por parte de los que la han estudiado. En la tradición española, Navarro Tomás (1948, pág. 51) la denomina **tonema**, y Stockwell *et al.* (1956, pp. 411-12) y Quilis (1981, pág. 416) **juntura terminal**. Estos tonemas o junturas terminales transmiten diferentes informaciones de tipo sintáctico, principalmente si la frase está acabada o no, y su modalidad oracional. Su inicio se sitúa en el último cambio de dirección de la

curva de Fo, que coincide según estos estudios, con la última sílaba acentuada, y se prolonga hasta el final de la curva. En inglés, este segmento comenzaría en el tono nuclear (Cruttenden, 1986, pág. 35).

Existen diferentes tipos de tonemas, que se distinguen por la inclinación que tome la curva melódica al final del grupo entonativo. No hay acuerdo, sin embargo, en cuanto al número. Para el español, algunos autores distinguen 3 tonemas diferentes - Stockwell (1956, pp. 411-12), Quilis (1981, pág 416):

- a) **anticadencia**, o *raising*, cuando la pendiente es ascendente.
- b) **suspensión**, o *level*, cuando se mantiene plana.
- c) **cadencia** o *falling*, cuando la pendiente desciende.

Otros, como Navarro (1948), distinguen hasta 5 tonemas diferentes:

- a) **anticadencia**.
- b) **semianticadencia**.
- c) **suspensión**.
- d) **semicadencia**.
- e) **cadencia**.

Existe también alguna clasificación intermedia que incluye cuatro tipos, como la de Takefuta (1975, pág. 373) para el inglés:

- a) *fast rising*.
- b) *rising*.
- c) *level*.
- d) *falling*.

El problema de cualquiera de estas clasificaciones reside en determinar dónde está el límite entre uno y otro tonema. Para resolver esta cuestión, Takefuta (1975) calculó el valor de la pendiente de la recta de regresión para cada uno de los segmentos terminales de su *corpus* y estableció una clasificación de los mismos de acuerdo con el rango en que se movían. De esta manera, limitó cada uno de sus cuatro tipos con unos valores máximo y mínimo de pendiente. Los valores en que se expresan estos valores corresponden a pendientes de rectas de regresión:

- | | |
|-----------------------------|------------------------------|
| 1) <i>fast rising</i> | pendiente entre 1,0 y 5,0. |
| 2) <i>rising</i> | pendiente entre 0,4 y 1,0. |
| 3) <i>level</i> | pendiente entre -0,2 y 0,4. |
| 4) <i>falling</i> | pendiente entre -5,0 y -0,2. |

Se ha afirmado igualmente que el **inicio** del grupo entonativo es también portador de información lingüística, conjuntamente con el segmento final. Kullova (1988, pág. 50), por ejemplo, defiende esta idea para explicar la entonación del español. Llama a esta teoría "Concepción discontinua de la forma melódica relevante", según la cual se considera relevante "el contorno melódico de la inflexión inicial de la oración, comprendido como el intervalo tonal entre la sílaba inicial absoluta y la primera sílaba acentuada de la oración, la cual complementa la relevancia fonética de la inflexión final, comprendida como el contorno melódico de la porción de la oración delimitada por la última sílaba acentuada y la sílaba final absoluta de la misma". También Studdert-Kennedy y Hadding defienden en el trabajo ya citado (1973) la idea de que en las curvas del inglés al menos una parte de la información se halla al inicio de la curva, concretamente en el primer pico.

Finalmente, se han señalado en la curva otros índices importantes para la transmisión de la información lingüística:

a) Entre los investigadores anglosajones se ha prestado bastante atención al fenómeno de la declinación (*declination*), que podría definirse, tal como hace Ladd (1984, pág. 53), como la tendencia de la Fo "*to decline gradually during the course of utterances*". Ladd presenta en este mismo artículo una buena revisión de los estudios realizados hasta ahora sobre este fenómeno. Hay que destacar sobre todo los trabajos de 't Hart, Cohen y su grupo - Cohen y 't Hart (1967); Cohen *et al.* (1982); Collier (1983) -, las aportaciones de Fujisaki y sus colaboradores - Fujisaki *et al.* (1969; 1971; 1982) -, y los estudios de Maeda (1976), Pierrehumbert (1979; 1987) y Cooper y Sorensen (1981).

Los resultados de estos estudios se prestan a la controversia: mientras que unos consideran la declinación un elemento fundamental para la descripción entonativa de cualquier lengua - Pierrehumbert, por ejemplo -, otros opinan que se trata de un fenómeno secundario, explicable por la actuación de otros mecanismos que intervienen en la generación de Fo, como la aparición en general de segmentos con pendiente descendente al final de las frases, o el acento. Lo que parece claro es que se trata de un fenómeno que no parece transmitir ningún tipo de información al receptor, aunque se ha destacado su importancia en la naturalidad de la entonación, aspecto especialmente relevante en el terreno de la síntesis. Así, Collier afirma que "*synthetic intonation without declination does not sound natural*" (Collier, 1989, pág. 42). Se ha destacado también su importancia como un "universal entonativo" (Vaissière, 1983), tras comprobar los resultados de otros análisis de la declinación en lenguas diferentes del inglés - los de Fujisaki, por ejemplo, para el japonés. Sin embargo, aún se carece de datos sistemáticos para el español.

b) Determinados estudios señalan como elemento relevante para distinguir entre algunos tipos de oración el segmento anterior a la última sílaba acentuada. Kullova (1982, pág. 75), por ejemplo, señala que para la distinción en español entre oraciones enunciativas e interrogativas absolutas es importante "el intervalo tonal entre la sílaba acentuada del núcleo de centro enunciativo de la oración y la sílaba inmediatamente anterior".

B) Como ya se ha indicado anteriormente, se han realizado pocos estudios sobre la **información de tipo sociolingüístico** que contienen las curvas melódicas. Quilis (1981, pág. 360) señala que según varios lingüistas - aunque no menciona cuáles - la información de tipo sociolingüístico se concentraría en la parte anterior al segmento final de la curva entonativa. Sin embargo, como indica también Quilis, hay quien opina que este segmento central de la curva no es relevante en absoluto.

C) La información de tipo **expresivo** puede manifestarse, en principio, con procedimientos bastante diversos, tal como señalan autores como Cruttenden (1986), para el inglés, o León (1970), para el francés. León distingue concretamente dos tipos de elementos que transmiten información expresiva:

1) Los que afectan a la totalidad de la frase de que se trate. Dentro de este grupo se encontrarían los siguientes:

a) La expresividad puede mostrarse, en primer lugar, ampliando el rango normal en que se mueven los valores de la F_0 , haciendo que ésta suba o baje más de lo habitual. Si estas variaciones se producen en el ámbito del tono nuclear, Cruttenden (1986, pp. 120-22) las denomina cambios en el *accent range* de una frase (**campo tonal**, en la traducción española). Si se dan en el marco de una curva entera, se etiquetan como cambios en la **tesitura** o *key* según Cruttenden (1986, pp. 129-30).

b) También puede marcarse la expresividad elevando el tono medio de la F_0 con que se pronuncia un enunciado. Este tono medio es lo que Cruttenden (1986, pp. 129-30) llama *register* (**registro** en la traducción española).

c) En ocasiones, según León, la forma de la curva melódica también puede proporcionar información de tipo expresivo.

d) La modificación de la **duración** de la frase es, para León, otro método de manifestar expresividad.

2) Los que sólo afectan a una parte de la curva melódica. Estos elementos serían, según León (1970):

a) En primer lugar, el ya mencionado **quinto tono** sobreagudo que considera típico de la expresividad en francés. Es un fenómeno que de hecho estaría relacionado con el cambio del rango en la frase.

b) El segundo método de este grupo es la utilización de una **marca tonal de implicación**, caracterizada por una forma peculiar de la curva melódica, que podría relacionarse con la que Navarro (1948, pág. 159) señala para el español, y que denomina **entonación circunfleja**: "La expresión de estados afectivos [...] utiliza [...] como principal recurso fónico la entonación circunfleja". Aunque Navarro Tomás no llega a definirla de una forma precisa, parece deducirse de sus explicaciones que se trata de un patrón melódico caracterizado por una subida de la Fo a la que sigue inmediatamente, incluso dentro de la misma sílaba, un rápido descenso. Su aparición, según él, es muy frecuente y sirve para expresar matices anímicos muy variados. Incluso señala la aparición de inflexiones circunflejas dobles, que sirven para expresar usualmente "el reproche afectuoso y la alegación exculpativa" (Navarro, 1948, pp. 162-3).

c) Finalmente, León afirma que un desplazamiento hacia arriba - más agudo - o hacia abajo - más grave - de la Fo en todo un segmento de una frase puede ser también una marca de expresividad.

1.2.2.4. Los patrones melódicos

Como se ha indicado anteriormente, se suele relacionar un tipo determinado de información entonativa con una forma concreta de la curva melódica, ya sea de una parte determinada o del contorno general de la misma. Cada una de estas formas típicas constituiría un **patrón melódico** diferente.

No es propósito de este trabajo entrar en la cuestión de si los patrones melódicos son universales o particulares a cada lengua, aunque parece claro que algunos esquemas, como los que sirven para distinguir entre enunciación e interrogación, parecen comunes a bastantes lenguas (Vaissière, 1983). Los apartados que siguen se centran únicamente en los patrones melódicos del español.

1.2.2.4.1. Tipos de patrones melódicos del español

Son varios los estudios que se han llevado a cabo sobre los patrones melódicos del español. Citaremos, como los más destacados, los realizados por Navarro (1948), Stockwell *et al.* (1956), Alarcos Llorach (1960), el del *Esbozo* de la Real Academia (R.A.E., 1973), los llevados a cabo por Quilis (Quilis y Fernández, 1979; Quilis, 1981), y por Kullova (1982; 1988), y el de Canellada y Kuhlmann (1987)⁶. La mayoría establece patrones melódicos a partir de la forma general de la curva, aunque alguno, como el de Stockwell *et al.* (1956), se centra únicamente en la parte final de la misma (las juntas terminales) como elemento diferenciador. De todos ellos, el estudio más completo es sin duda el de Navarro. En él se analizan únicamente los patrones melódicos del español en función de la modalidad oracional que exprese el enunciado. Navarro Tomás estudia cuatro grandes grupos de patrones:

1) Patrones **enunciativos**: son los que aparecen en las oraciones enunciativas, es decir, aquéllas que simplemente expresan "la conformidad o disconformidad lógica del sujeto con el predicado" (R.A.E., 1973, pág. 354), tal como ocurre en (1):

(1.a). Son las cinco en punto.

(1.b). Hoy es sábado.

2) Patrones **interrogativos**: en esta categoría, Navarro Tomás incluye gran cantidad de subgrupos (concretamente, distingue entre interrogativas absolutas, relativas, restrictivas, aseverativas, imperativas, pronominales, reiterativas, exclamativas, hipotéticas y alternativas), que podrían reducirse, siguiendo la propuesta de Canellada y Kuhlmann (1987, pág. 147 y ss.), a tres fundamentales:

2.1) Patrones de las **interrogativas absolutas**: este tipo de interrogaciones son las que han de tener "sí" o "no" obligatoriamente como respuesta; en palabras del *Esbozo*, son aquéllas en que "preguntamos la verdad o falsedad del juicio" que contiene la frase (R.A.E., 1973, pág. 359), como se observa en (2):

(2.a) ¿Recibió usted mi carta?

(2.b) ¿Vendrás hoy a cenar?

2.2) Patrones de las **interrogativas pronominales**: aparecen en aquellas preguntas que contienen una partícula interrogativa, y cuya respuesta ha de ser diferente de "sí" o "no". Ejemplos de este tipo serían

⁶ Una revisión de varios de estos trabajos puede encontrarse en Kvavik y Olsen (1974).

las oraciones de (3):

- (3.a) ¿**Qué** quieres?
- (3.b) ¿**Cuándo** ha venido?
- (3.c) ¿Por **qué** lo has hecho?

2.3) Patrones de las **interrogativas relativas**: serían, en términos generales, todas las interrogaciones que buscan confirmar algo que no se sabe con certeza, pero que se intuye de alguna manera. Canellada y Kuhlmann (1987, pág. 150) las definen como aquellas preguntas en las que "el que pregunta, sabe o presume saber, o, por lo menos, desea lo que le van a contestar". Los ejemplos que dan Canellada y Kuhlmann son los que aparecen en (4):

- (4.a) ¿Quiere usted pasar?
- (4.b) ¿Hacemos un alto?
- (4.c) ¿Nos quedamos un rato?
- (4.d) ¿Verdad que no?

3) Patrones **exclamativos**, que serían los que aparecen, de forma general, en todas aquellas oraciones con un contenido expresivo o emocional. Podrían señalarse tantos subgrupos de esta clase de frases como emociones o estados afectivos puedan ser expresados por un hablante. Navarro (1948, pág. 156) menciona bastantes tipos - los que incluyen las frases que expresan alegría, irritación, orgullo, tristeza, miedo, cariño, desdén, odio -, pero es prácticamente imposible dar una lista completa. Por ello, se incluyen aquí todos estos subgrupos dentro de un solo apartado, el de los patrones exclamativos en general. Ejemplos de frases exclamativas serían los de (5), tomados también de Canellada y Kuhlmann (1987):

- (5.a) ¡Qué maravilla!
- (5.b) ¡Qué le vamos a hacer!
- (5.c) ¡Pero si es muy tarde!
- (5.d) ¡Qué vergüenza!

4) Patrones **volitivos**, que aparecen en las oraciones exhortativas, es decir, las "que indican exhortación, mandato o prohibición" (R.A.E., 1973, pág. 361). Navarro distingue tres subgrupos:

- 4.1) patrones de **mandato**, propios de las oraciones imperativas.
- 4.2) patrones de **recomendación**, en oraciones que contienen mandatos atenuados, sugerencias o recomendaciones.
- 4.3) patrones de **deseo**, en oraciones que expresan deseos del emisor.

Las oraciones de (6) son ejemplos de oraciones exhortativas. La primera sería un caso de oración imperativa, la segunda, de mandato refrenado, y la tercera de deseo. Todas ellas están extraídas de Navarro (1948):

- (6.a) ¡Calla!
- (6.b) ¿Tendría la bondad de entregarle esta carta?
- (6.c) Me agradecería que tomase afición a los estudios astronómicos.

Navarro no hace una distinción demasiado clara entre patrones correspondientes a los ya mencionados **grupos terminales** y los **no terminales**, es decir, entre aquellos esquemas que aparecen en grupos entonativos que están al final de una oración, y aquéllos que aparecen en grupos entonativos no finales. De hecho, ninguno de los estudios citados la realiza explícitamente. Y sin embargo, muchos de los patrones que se presentan en la bibliografía para oraciones compuestas por más de un grupo entonativo son combinaciones de estos patrones básicos. Dichas combinaciones podrían denominarse **patrones complejos**. Estos patrones se dan en enunciados de distintos tipos, como los que se enumeran a continuación, ilustrados con ejemplos de Quilis y Fernández (1979):

1) las **enumeraciones**, formadas por una serie de grupos no terminales y un grupo terminal enunciativo:

- (7) Lunes, martes, miércoles y viernes.

2) las **interrogaciones alternativas**, combinación de un grupo no terminal y un grupo terminal descendente. Así ocurre en el ejemplo de (8):

- (8) ¿Fue Pedro o Enrique?

3) las **preguntas confirmativas**, también formadas por un esquema no terminal y un esquema terminal interrogativo, esta vez ascendente.

- (9.a) Es bonita, ¿no?
- (9.b) No estás cansado, ¿verdad?

1.2.2.4.2. Forma típica de los patrones melódicos del español

Cada categoría oracional tiene asociada, según los estudios realizados hasta el momento, uno (o dos, en algunos casos) esquemas melódicos determinados. A continuación se hace un rápido repaso de los mismos.

A) Por lo que se refiere a las oraciones enunciativas, en general todos los

estudios coinciden en señalar la presencia de un **segmento final descendente** como rasgo más característico. Así, por ejemplo, Navarro (1948, pág. 57) afirma que “la cadencia es el signo ordinario de la aseveración”, e igualmente, Canellada y Kuhlmann (1987, pág. 126) consideran que la entonación enunciativa “va esencialmente unida al descenso del tono final”. Una curva típicamente enunciativa presenta una subida inicial hasta la primera sílaba tónica, a partir de la cual se inicia un descenso que se prolonga hasta el final del grupo, y que puede acentuarse a partir de la última sílaba tónica.

El *Esbozo* de la R.A.E. (1973, pág. 106) señala también como rasgo propio de las enunciativas “la uniformidad del tono en el cuerpo central”, aunque no se especifica más qué se entiende por “uniformidad del tono” - si no varía, o si baja de una manera constante.

B) En los enunciados interrogativos, suelen distinguirse tres patrones diferentes, cada uno caracterizado por la presencia de un segmento final distinto:

1) El primer esquema, el típicamente interrogativo, acaba con un **final ascendente**. El grupo comienza con un ascenso de la Fo hasta la primera sílaba acentuada, sigue con un descenso pronunciado “hasta la última sílaba fuerte o hasta la que inmediatamente le precede” (Navarro, 1948, pag. 101), y finaliza con un tonema ascendente. Éste sería el esquema propio de las interrogaciones absolutas y, ocasionalmente, de las pronominales.

2) El segundo esquema termina, como en las enunciativas, con un **segmento descendente**. Tras una subida inicial, hasta la primera sílaba acentuada, comienza un descenso que alcanza hasta el final del grupo. Es el esquema típico de las preguntas pronominales, aunque éstas también pueden adoptar el esquema anterior, con final ascendente.

3) Finalmente, el tercer esquema se caracteriza por la presencia de un **final con forma circunfleja**, según Navarro (1948, pág.104). Se reconoce por la presencia de una subida inicial hasta la primera sílaba acentuada, un descenso y, ya en la última sílaba acentuada, la aparición del segmento circunflejo: “La voz vuelve a elevarse en esta sílaba [la última acentuada] para descender inmediatamente en ella misma o en la siguiente”.

C) Los enunciados exclamativos o afectivos, en los que domina la transmisión de emociones y estados de ánimo, también pueden presentar esquemas melódicos diferentes, de acuerdo con Navarro (1948, pág. 152 y ss.) o Canellada (1987, pp. 153-54):

1) Algunas frases exclamativas presentan un patrón melódico en el que la Fo va descendiendo progresivamente desde el principio hasta el final de la curva.

2) En otros casos, la Fo va subiendo de forma progresiva. Esta forma de exclamación es descrita por Navarro (1948, pp. 175-6) de la siguiente manera:

“La elevación de la voz se desarrolla suave y gradualmente a lo largo de la frase partiendo del tono semigrave de la sílaba inicial y continuando hasta alcanzar el nivel semiagudo en la que precede inmediatamente a la que lleva el último acento. En la última sílaba acentuada, la voz realiza un breve descenso y en seguida se eleva de nuevo dentro de esa misma sílaba o en la siguiente inacentuada quedando el tono finalmente suspendido hacia el nivel semiagudo”.

3) Finalmente, un tercer esquema exclamativo presentaría una forma ondulada, con subidas y bajadas a lo largo de la frase, coincidiendo respectivamente con las sílabas acentuadas e inacentuadas.

En cualquiera de los tres esquemas, la terminación típica sería el segmento circunflejo, que para Navarro (1948, pág. 159) es el medio principal de “expresión de estados afectivos”.

En otros estudios la descripción de los esquemas exclamativos es mucho menos detallada. Quilis (1981, pág. 442), por ejemplo, afirma que las curvas de este tipo de oraciones se caracterizan simplemente por un descenso muy acusado de la Fo desde la primera sílaba acentuada hasta el final, y por un nivel tonal algo más bajo a lo largo de toda la frase. Igualmente, se señala en el *Esbozo* de la R.A.E. (1973, pág. 115) que las exclamativas se caracterizan por un descenso final, sin aportar más detalles.

D) Finalmente, los enunciados volitivos presentarían dos esquemas diferentes, según Navarro (1948, pág. 130 y ss.):

1) Unos, con un **segmento final descendente**. Sería el esquema propio del mandato, la exhortación y la recomendación, principalmente. La expresión de uno u otro tipo de matiz se realizaría por medio de variaciones en la amplitud. Dicho esquema también se menciona en Quilis y Fernández (1979, pág. 181).

2) Otros, con un **segmento final ascendente**. En ocasiones, según Navarro (1948, pág. 136), "se refuerza la viveza del mandato terminando las frases con entonación ascendente".

E) Como ya se ha indicado, los patrones correspondientes a grupos no terminales no han sido estudiados de una manera sistemática hasta el momento; hay descripciones de algunos de estos esquemas en Navarro (1948), Quilis y Fernández (1979), Quilis (1981) y Canellada y Kuhlmann (1987), aunque también puede citarse algún trabajo más específico, como el de Garro y Parker (1983), sobre los rasgos suprasegmentales de las frases de relativo en español. En general, los esquemas de los grupos no terminales suelen presentar un final ascendente - como ocurre, por ejemplo, en los grupos no terminales de las enumeraciones -, aunque también se dan casos de final descendente. También podemos encontrar grupos caracterizados primordialmente por un descenso acusado de su nivel tonal medio con respecto al de los que le preceden o siguen. Es el esquema típico de las oraciones parentéticas o intercaladas.

2. ANALISIS DE LA ENTONACION Y TECNOLOGIAS DEL HABLA

El interés por el estudio de los fenómenos prosódicos se ha acrecentado en los últimos años con el desarrollo de las tecnologías del habla. De hecho, algunos de los análisis presentados hasta el momento - el de Pierrehumbert, por ejemplo - se han realizado en el marco de trabajos orientados hacia la utilización de la información prosódica en sistemas de síntesis o reconocimiento automático. Por lo que se refiere a la descripción de los patrones melódicos, la necesidad de integrar los resultados obtenidos en sistemas automáticos ha obligado a aumentar el rigor de los métodos de análisis utilizados y a presentar descripciones de los mismos mucho más precisas, que incluyan valores numéricos.

Este capítulo está dedicado a la revisión de los principales sistemas desarrollados, tanto de síntesis como de reconocimiento, que utilizan información prosódica, y especialmente información sobre las curvas melódicas.

2.1. Entonación y síntesis automática

El componente entonativo de los conversores texto-habla es uno de los principales responsables de que el habla generada sea natural. Algunos estudios - Sorin (1982), por ejemplo - han mostrado también su contribución a la inteligibilidad de los enunciados, al menos por lo que se refiere al habla natural. Por ello, se ha prestado bastante atención a este problema en los sistemas ya desarrollados. Los módulos prosódicos integran la información referente al acento y a la entonación, y serían los encargados de controlar la amplitud y duración de los diferentes sonidos, así como las variaciones de la F_0 a lo largo de los enunciados. Para este segundo cometido, los módulos prosódicos incluyen algoritmos de generación automática de curvas melódicas. Estos algoritmos pueden estar contruidos según dos enfoques diferentes:

- a) Los sistemas incluidos en el primer enfoque incluyen una gramática que permite calcular secuencialmente los valores de F_0 para los diferentes segmentos de la curva (cada sílaba, por ejemplo). Pierrehumbert (1987) describe un sistema de generación de la F_0 que utiliza una gramática de estados finitos, en la que cada estado representa una sílaba. Dicha gramática va asignando a cada sílaba un nivel tonal determinado, según sea sílaba inicial o final de frase, sílaba acentuada o inacentuada, o sílaba que contenga el tono nuclear (*phrase accent*) de la frase, siguiendo la filosofía descrita en apartados

anteriores. Este sistema se ha incorporado al conversor texto-habla desarrollado por Olive y Liberman (1979).

b) En el segundo enfoque, se utilizan modelos de curvas estilizadas que se superponen a la frase de que se trate, y que se van modificando según diferentes reglas para incorporar información suplementaria, por ejemplo la referente al acento. Un ejemplo de este sistema sería el algoritmo de generación de Fo de O'Shaughnessy desarrollado para el MITalk (O'Shaughnessy, 1987). Este algoritmo tiene determinados tres esquemas básicos de Fo, uno para frases enunciativas, otro para interrogativas absolutas y un tercero para interrogativas pronominales. Una vez escogido uno de estos tres esquemas, diferentes reglas se encargan de modificarlo para tener en cuenta, entre otros fenómenos, las variaciones de la Fo debidas al acento y a las características de cada elemento segmental (micromelodía). También utilizan patrones semejantes a los de O'Shaughnessy el sistema desarrollado para el español en la E.T.S.I.T. de Madrid (Olabe, 1983; Moreno *et al.*, 1989), y el descrito por Debello *et al.* (1987) para el italiano.

Tanto uno como otro enfoque implican un proceso de estilización, es decir, de eliminación de la información irrelevante. Las diferencias residen en lo que no se considera necesario mantener en cada caso. Así, tanto el sistema de Pierrehumbert como el de O'Shaughnessy tienen en cuenta el fenómeno de la declinación para la generación de las curvas. Sin embargo, mientras que Pierrehumbert utiliza una representación de niveles, O'Shaughnessy se decanta por el uso de contornos. El de O'Shaughnessy emplea también la información micromelódica para generar las curvas, aspecto que Pierrehumbert no tiene en cuenta. Akers y Lennig (1985) han comparado la calidad de las curvas de Fo generadas por los sistemas de O'Shaughnessy y Pierrehumbert mediante tests de evaluación. Los resultados son bastante semejantes en ambos casos, con una pequeña ventaja del primero, según los autores debida a las diferencias en la generación del acento existentes entre ambos algoritmos.

El sistema desarrollado por Olabe para el conversor texto-habla de la ETSIT de la Universidad Politécnica de Madrid sigue el segundo enfoque: utiliza una serie de patrones melódicos definidos mediante un análisis acústico previo, que se basan en la clasificación de patrones realizada por Quilis (1981). A estos patrones se le añadirían posteriormente las variaciones debidas al acento.

Teniendo en cuenta las particularidades melódicas del español (*vid.* apartado 1.2.1.2), parece más adecuado este segundo enfoque para la generación automática de curvas melódicas en esta lengua. El primero parece más adecuado para lenguas como el inglés, en las que el acento léxico juega un papel importante en la conformación de la curva melódica de una frase determinada. Un procedimiento de estilización por contornos como el utilizado en este trabajo permite presentar las curvas melódicas en un formato más apropiado para su empleo en un sistema que utilice este segundo enfoque.

2.2. Entonación y reconocimiento automático

Los intentos de reconocer de manera automática patrones entonativos o de incorporar información suprasegmental al reconocimiento de elementos segmentales han sido mucho menos habituales que en el caso de la síntesis. Los sistemas de reconocimiento han de enfrentarse a dificultades que no se plantean en un proceso de generación de melodía, como son los problemas inherentes a la extracción de F_0 : los procedimientos de estimación existentes hoy en día presentan errores que es necesario subsanar mediante procesos de **alisamiento** (*smoothing*).

Los sistemas desarrollados hasta ahora pueden clasificarse en función del tipo de información que intentan obtener de los parámetros prosódicos y del enfoque que subyace en el análisis de los mismos. Así, pueden distinguirse:

a) Aquéllos en los que subyace una aproximación por niveles, y toman la sílaba como unidad de análisis. Son sistemas que intentan extraer, a partir del análisis de los parámetros prosódicos de F_0 , amplitud y duración en cada sílaba, diversas informaciones sobre la posición de los acentos y sobre la localización de límites sintácticos. Muchos de ellos sirven sólo de ayuda para el reconocimiento de elementos segmentales. En este grupo se incluyen:

1) En primer lugar, el sistema automático de segmentación sintáctica desarrollado a partir del procedimiento de estilización de Rossi y su grupo (Rossi *et al.*, 1981, pp. 65-98), ya descrito en el apartado 1.2.1.2. La representación estilizada obtenida se utiliza para detectar los límites sintácticos y etiquetar los diferentes grupos fónicos, de acuerdo con una serie de "reglas transformacionales de enunciación": reglas generativas de enunciación, reglas de ajuste prosódico y reglas de correspondencia enunciación-sintaxis.

2) El sistema descrito por Mertens (1989) para el francés y el holandés, que utiliza los parámetros de F_0 y amplitud, además del número de cruces por cero de la señal temporal, para realizar una segmentación en sílabas. Se efectúa un proceso de "estilización" de los valores de F_0 y amplitud, que simula las transformaciones que realiza el oído humano sobre la señal acústica en el proceso de audición. También se procede a una corrección de los errores en la estimación de F_0 introducidos durante el proceso de parametrización. Posteriormente, utilizando los datos obtenidos tras la estilización, se extrae una serie de rasgos prosódicos para cada sílaba, como [□ acentuada] o [□ larga] entre otros. También se etiquetan como pausas los segmentos de silencio superiores a 300 ms. de duración. El autor no proporciona datos sobre la fiabilidad del sistema, aunque indica que los detectores de F_0 y de amplitud introducen errores que luego se propagan a través de todo el sistema (Mertens, 1989, pág. 49).

3) Vaissière (1989) describe un sistema semejante, que ha sido incorporado a KEAL, el reconocedor automático de palabras aisladas del CNET francés. El sistema extrae los siguientes parámetros prosódicos (PP):

- a) PP1: **duración** de las vocales y consonantes.
- b) PP2: movimientos de Fo dentro de cada vocal.
- c) PP3: Fo en la parte estable de cada vocal, que se etiqueta como **altura tonal** de dicha vocal.
- d) PP4, PP5 y PP6: corresponden a los **valores diferenciales** de PP1, PP2 y PP3 entre vocales consecutivas.
- e) PP7: **picos y valles de Fo**.
- f) PP8: **picos y valles de amplitud**.

Como en el sistema de Mertens, existe un proceso de corrección de los errores producidos durante la fase de parametrización, tanto en las duraciones de los segmentos como en la estimación de Fo, aunque, como reconoce la autora, "*the problem of erroneous parameters extraction is not yet solved*" (Vaissière, 1989, pág. 203).

A partir de los parámetros indicados, se obtiene una serie de informaciones lingüísticas:

- a) **Final de frase enunciativa**, detectado por la presencia de una bajada de Fo al final de un grupo.
- b) **Marca de grupo espiratorio no final**, caracterizado por una subida de Fo en una sílaba situada antes de pausa.
- c) **Palabra gramatical monosilábica al principio de una frase**, localizable por la presencia de un valor de Fo relativamente bajo en la primera sílaba de dicha frase.
- d) **Límite sintáctico (*major boundary*) dentro de un grupo fónico**, ligado a una duración de la sílaba por encima de lo normal, y a un esquema ascendente-descendente (*fall-rise*) de Fo.
- e) **Límites derecho e izquierdo de palabra**, marcados con subidas de Fo y alargamientos de sílabas.
- f) **Palabra gramatical en el interior de un grupo espiratorio**, caracterizada por un valor bajo de Fo, precedido de una bajada y seguido de una subida de Fo, así como un alargamiento de la sílaba previa.

Como se deduce del tipo de información que se obtiene mediante este procedimiento, la extracción automática de información suprasegmental sirve en KEAL sólo como apoyo al reconocimiento de las unidades segmentales, y no persigue el reconocimiento global de patrones entonativos.

El principal problema de estos sistemas reside en la detección de las sílabas: es necesario un procedimiento que segmente y etiquete previamente la señal correspondiente al enunciado y que detecte los núcleos silábicos para poder realizar el análisis prosódico, tarea que no siempre resulta fácil.

b) Aquéllos que utilizan representaciones de contornos, que intentan reconocer patrones melódicos de manera global, y que tratan de extraer información específicamente entonativa, como el modo de la oración, a partir de las características del patrón reconocido. Es una línea menos seguida que la anterior, aunque se describen en la bibliografía algunos trabajos en este sentido:

1) En Martín (1970) se propuso un esquema de reconocedor automático de patrones melódicos. Es un modelo teórico, realizado en una época en que los medios técnicos no permitían llevarlo a la práctica. Sin embargo, resulta interesante como una primera aproximación al problema.

2) Waibel (1988, pp. 96-102) describe un sistema de discriminación de patrones enunciativos e interrogativos para el inglés, que comprende las siguientes etapas:

a) Extracción de la curva de la Fo.

b) Aplicación de un proceso de alisamiento para eliminar los errores en la detección de la Fo. El procedimiento utilizado consiste en calcular medias sucesivas de 5 puntos de la curva y eliminar los puntos cuyo valor de la Fo excede del doble, o no llega a la mitad, del valor medio de la Fo en toda la curva.

c) Se extrae la línea de declinación de toda la frase mediante el cálculo de la línea de regresión (segundo orden) de todos los puntos de la curva, y se calcula la pendiente del segmento final. Estos dos parámetros son los que se utilizan después para el etiquetaje de la curva.

La fiabilidad del sistema se sitúa, de acuerdo con los tests de evaluación que presenta el autor, en torno al 77%. Una buena parte de los errores se deben, según él mismo, a fallos en el detector de la frecuencia fundamental.

3) Finalmente, aunque no se trata de un sistema desarrollado completamente, hay que mencionar aquí de nuevo los trabajos ya citados de 't Hart y su grupo: la extracción automática de *close copy stylizations*, y su comparación con modelos obtenidos previamente constituye una aproximación interesante al problema, sobre la que dicho grupo está trabajando en estos momentos. En 't Hart (1979) se describen los intentos de automatización del proceso de extracción de *close copy stylizations*. Dicho proceso comprende las siguientes fases:

- a) Tras la extracción de la F_0 , se procede a una corrección de errores por un sistema de **alisamiento**, consistente en calcular la media de diferentes tramas consecutivas.
- b) Se determinan después las líneas superior e inferior de la declinación en la frase analizada, y se colocan en la línea correspondiente los valores de F_0 situados entre 0 y 4 semitonos fuera de la región comprendida entre las dos líneas.
- c) Se procede al reconocimiento de la curva de F_0 por medio de una **gramática** que contemple las combinaciones posibles de segmentos que componen los diferentes patrones de F_0 .

Los autores de este sistema lo han aplicado a la generación automática de curvas de F_0 (Collier y Terken, 1987; Terken y Collier, 1989). Parece claro que también podría aplicarse al reconocimiento de patrones melódicos.

Los sistemas de este tipo, al no estar basados en un análisis silábico, no plantean los problemas señalados para el primer grupo. Sin embargo, desde el punto de vista del reconocimiento, proporcionan sólo información entonativa, en tanto que los primeros permiten obtener, además de la entonativa, también información acentual.

El procedimiento de estilización presentado en este trabajo se adecuaría más al segundo tipo de sistemas que al primero. Al no tomar la sílaba como unidad de análisis, ni tener en cuenta los parámetros de duración y amplitud de los distintos segmentos, no puede utilizarse en reconocedores automáticos que intenten extraer información relativa al acento, por ejemplo. Sí que podría utilizarse, en cambio, para la detección de la modalidad oracional o de otras informaciones de tipo sintáctico-semántico contenidas en la forma de la curva melódica.

3. UN SISTEMA DE ESTILIZACION DE CURVAS MELODICAS PARA EL ESPAÑOL

Como se deduce de los capítulos anteriores, el primer paso a la hora de afrontar un estudio de la entonación consiste en definir, de una manera más o menos explícita, el método de análisis. Esta ha sido la primera tarea abordada en el presente estudio.

La elección del método de análisis depende directamente de los objetivos que se persigan. Dichos objetivos serían, para el caso del estudio que aquí se presenta, los siguientes:

- a) El estudio se ha limitado, en principio, al análisis de las curvas melódicas. La duración y amplitud de los elementos segmentales quedaría fuera del ámbito del trabajo.
- b) Se ha pretendido extraer de las curvas melódicas únicamente la información entonativa. Las variaciones melódicas debidas al acento han quedado también excluidas de este estudio.
- c) La investigación se ha centrado de momento en las curvas melódicas del español, aunque podría ampliarse a otras lenguas con un comportamiento melódico similar, como el catalán.
- d) Se ha tenido en cuenta que los resultados obtenidos puedan utilizarse en sistemas de reconocimiento y síntesis.

Un procedimiento de análisis adecuado para estos objetivos debía consistir en un sistema de representación de las curvas melódicas del español que retuviese sólo las variaciones relevantes para la detección de la información entonativa contenida en las mismas, y que fuese un método de fácil automatización, aplicable tanto a sistemas de reconocimiento como de síntesis. La determinación del procedimiento que ha de utilizarse obliga a una serie de decisiones previas, en función de las cuestiones señaladas en los capítulos anteriores:

- a) Se ha elegido un método de representación **de contornos**, y no de niveles, dadas las dificultades ya reseñadas que implica el establecimiento de los distintos niveles

(dificultades que suponen, entre otras cosas, tener que utilizar información acerca de otras curvas del mismo oyente). Parece, por otro lado, un enfoque más adecuado para la representación de las curvas del español, y que se halla implícito en la mayoría de los estudios realizados hasta la fecha sobre el tema.

b) Se ha desechado la sílaba como unidad de análisis, dadas las dificultades que implica la detección automática de las mismas. El análisis se ha centrado en partes determinadas de las curvas melódicas, denominadas **segmentos** y **puntos de inflexión** (para una definición, véase más adelante, en este mismo apartado).

c) La representación resultante debía retener únicamente aquellas variaciones de la Fo relevantes para la transmisión de la información entonativa. Las variaciones micromelódicas y las variaciones inter-locutor habían de ser eliminadas.

A continuación se describe el método desarrollado para la realización del presente trabajo. Hay que señalar de entrada que se trata de un proceso de estilización que utiliza de momento sólo criterios acústicos, y no perceptivos. Tras el análisis acústico que se presenta aquí es necesario llevar a cabo otro perceptivo, que permita validar y refinar el procedimiento, sobre todo en lo que se refiere a la detección de las variaciones relevantes desde el punto de vista entonativo (punto c).

El proceso de estilización comprende actualmente los siguientes pasos:

a) En primer lugar, la extracción de una serie de valores de la curva de la Fo en unos puntos determinados de la curva:

- 1) Tiempo de inicio de la curva.
- 2) Valor de Fo en dicho inicio.
- 3) Localización temporal de los diferentes **puntos de inflexión**.
- 4) Valor de Fo en cada uno de estos puntos.
- 5) Tiempo al final de la curva.
- 6) Valor de Fo en ese punto.

Se considera **punto de inflexión** cualquier punto de la curva en el que se produce un cambio en el signo de la pendiente de la misma, siempre que la diferencia de frecuencia con respecto a la última inflexión sea de 10 Hz o más⁷.

⁷El valor asignado a este umbral está por encima de los umbrales diferenciales de percepción de Fo citados en el apartado 1.1.2. Pretende, por tanto, eliminar aquellas variaciones melódicas de una curva melódica imperceptibles para los oyentes. Sin embargo, debe considerarse como un punto de referencia inicial, que habrá de ser revisado en función de los resultados obtenidos en el estudio acústico que se describe a continuación, y en posteriores estudios perceptivos.

b) Realización de un proceso de **normalización frecuencial** con los valores obtenidos:

El propósito de esta normalización es eliminar las diferencias en el tono fundamental que se dan entre un hablante y otro. En el apartado 1.2.1.2 se ha hecho un repaso a los diferentes métodos propuestos en la bibliografía para normalizar el tono de los diferentes hablantes, y a los inconvenientes que presenta cada uno. Aquí se ha optado por un sistema diferente, más sencillo, teniendo en cuenta que las representaciones resultantes han de ser aplicables tanto a la síntesis como al reconocimiento, y que los métodos comentados sólo son válidos para la síntesis.

El sistema utilizado en este trabajo consiste en asignar al punto inicial de la curva valor cero, y relativizar los demás valores con respecto a ese punto. Para ello, se resta el valor de la F_0 en el punto de inicio de la curva a todos los valores de la F_0 extraídos de la misma.

c) Finalmente, se unen mediante rectas los valores frecuenciales obtenidos. También podría utilizarse otro tipo de línea (sinusoidal, por ejemplo) para realizar la interpolación. El uso de una u otra será el objetivo de próximos trabajos, dada la importancia que este hecho tendría para la naturalidad de las curvas melódicas generadas mediante este procedimiento.

De esta forma se obtiene la correspondiente "representación estilizada" de cada curva, consistente en una serie de líneas o **segmentos** que representan cada una un fragmento de la curva comprendido entre dos puntos de inflexión.

Dos son las fuentes de variación principales que se pretenden eliminar con el método utilizado: las **variaciones micromelódicas** y el **fundamental habitual** de los distintos locutores. Las primeras, mediante el umbral de variación mínima, situado en principio en los 10 Hz; las segundas, mediante el procedimiento de normalización frecuencial. Tal como se ha señalado en el primer capítulo, se han descrito en la bibliografía otros procedimientos para eliminar estas variaciones. Sin embargo, se ha optado por estos sistemas más sencillos, que no precisan en principio un conocimiento previo de otras curvas melódicas de los mismos locutores. Como ya se ha indicado, la validez de los métodos utilizados para su eliminación está aún pendiente de tests perceptivos, así como la determinación de las variaciones relevantes desde el punto de vista entonativo.

4. ANALISIS Y CLASIFICACION DE LOS PATRONES MELODICOS DEL ESPAÑOL

4.1. Objetivo

El objetivo de este estudio es doble: por un lado, **validar experimentalmente el proceso de estilización desarrollado**, y por otro, **realizar una descripción de los patrones melódicos del español que permitiese su aplicación al desarrollo de sistemas automáticos de reconocimiento y síntesis de entonación**. Aunque una descripción de este tipo ya se había abordado en trabajos anteriores, se han considerado varias razones que justifican la realización de una más:

a) En primer lugar, la necesidad de completar las descripciones disponibles hasta el momento sobre los patrones melódicos del español. Hasta ahora, los estudios sobre la entonación habían visto limitados sus resultados debido, en parte, a la falta de precisión del instrumental utilizado. En un principio, los quimógrafos, y más adelante los sistemas analógicos de análisis de la señal (el espectrógrafo, por ejemplo), permitían un estudio limitado de la melodía. Por ello, a pesar de la exhaustividad de la mayoría de estos análisis, ninguno presenta valores numéricos sobre las curvas, al estilo de como se ha hecho para los elementos segmentales, sino únicamente una serie de esquemas de patrones melódicos. El estudio de Navarro (1948), sin duda el más completo que se ha publicado para el español, presenta este inconveniente, por basarse sus conclusiones en representaciones quimográficas fundamentalmente. En la actualidad, con la aparición de los sistemas digitales, y a pesar de que la detección de la frecuencia fundamental sigue siendo un problema sin resolver por completo, se pueden efectuar estudios mucho más precisos.

b) La necesidad de proporcionar a las nacientes "tecnologías del habla" descripciones de los fenómenos entonativos mucho más precisas y formalizadas que las que se han realizado hasta ahora, para poder avanzar en el campo de la síntesis y el reconocimiento automáticos del habla. Mientras que para el inglés (O'Shaughnessy, 1979; Cooper y Sorensen, 1981; Pierrehumbert, 1987) o el francés (Rossi *et al.*, 1981; Martin, 1982) podemos encontrar ya bastantes estudios de la entonación con este enfoque, para el español no existe prácticamente ninguno. Tan sólo los llevados a cabo en la Escuela Técnica Superior de Ingenieros de Telecomunicación de Madrid, en colaboración con el

Departamento de Lengua Española de la Universidad Nacional de Educación a Distancia, con vistas a la elaboración de un sistema de conversión texto-habla para el español, se orientan en esta línea: la tesis doctoral de Olabe (1983), y los trabajos posteriores de Pardo *et al.* (1987) y de Moreno *et al.* (1989). Sin embargo, son estudios enfocados a la síntesis y no al reconocimiento, y que tratan sólo unos cuantos tipos entonativos - los presentados por Quilis (1981).

4.2. Acotación del objeto de estudio

El estudio que se presenta a continuación no pretende abarcar todos los patrones melódicos posibles del español. Se procuró que el campo de estudio coincidiese con el de los trabajos realizados anteriormente, de manera que éstos pudieran servir de referencia a la hora de validar el procedimiento de estilización. Por otro lado, la información que puede contener una curva melódica es muy variada, tal como se ha indicado en el primer capítulo, y resultaría imposible analizar adecuadamente en un solo trabajo los patrones asociados a cada tipo de información. Así, el estudio presenta en este sentido una serie de limitaciones, que son las que se especifican a continuación:

- a) Se trata, como ya se ha señalado, de un estudio **acústico**, aunque servirá de base para posteriores estudios perceptivos.
- b) El único parámetro estudiado ha sido la **variación de la Fo** a lo largo de la curva, dejando de lado los parámetros de amplitud y duración.
- c) El estudio se ha limitado al análisis de los **patrones terminales** del español (*vid.* definición en el apartado 1.2.2.1).
- d) El estudio se ha centrado en la transmisión de información acerca de la **modalidad oracional** por medio de dichos patrones.

4.3. Procedimiento experimental

A continuación se describen con detalle las diferentes fases del estudio, que ha comprendido la elaboración del *corpus* (apartado 4.3.1), su grabación (apartado 4.3.2) y el análisis de las diferentes frases (4.3.3). Los resultados se presentan en el apartado 4.3.4. Finalmente, en el apartado 4.3.5 se establece una clasificación de los patrones melódicos terminales del español de acuerdo con dichos resultados, y en el 4.3.6 se comparan los resultados obtenidos con los de trabajos anteriores.

4.3.1. Elaboración del *corpus*

El *corpus* empleado para este estudio está compuesto por un total de 30 diálogos cortos, que incluyen las 54 frases analizadas, y que fueron leídos por diferentes locutores. Tanto las frases como los diálogos se presentan en el anexo final. Las frases que componen el *corpus* están subrayadas dentro de cada diálogo, y al lado de cada una se indica un número de identificación. Con dicho número se realizan, de aquí en adelante, las referencias a cada frase.

Frente a la opinión de algunos autores que defienden la utilización de *corpora* extraídos del habla espontánea - León y Martín (1970, pp.1-6), por ejemplo -, se ha preferido en este caso la utilización de un *corpus* leído, técnica empleada en estudios como el de Caelen (1981). Aunque los primeros ofrecen la ventaja de la **naturalidad**, los segundos permiten el máximo **control** de las frases del *corpus*, de su contenido y su estructura. El *corpus* empleado en este trabajo está construido de manera que se puedan establecer **pares mínimos** de frases, opuestas sólo en un rasgo, como se puede comprobar en los siguientes ejemplos:

- | | | |
|----|------------------------------|---------|
| a) | Ramón llegó en avión. | (1.1) |
| | Ramón llegó andando . | (1.2) |
| b) | Ramón llegó en avión. | (1.1) |
| | ¿Ramón llegó en avión? | (2.3.1) |

Así, hay en el *corpus* frases con el mismo número de sílabas pero con el acento final en diferente posición - ejemplo (a) -, o frases iguales con modalidad diferente - ejemplo (b) -, entre otras variables que se detallan más adelante.

Otro motivo que ha llevado a emplear un *corpus* leído se deriva del hecho de que algunos tipos concretos de frases son muy difíciles de "provocar" en el habla espontánea. Así ocurre, por ejemplo, con las interrogaciones relativas. El uso de diálogos perseguía precisamente dirigir al locutor en la elección del patrón melódico de cada frase, creando una situación que obligase a su empleo.

Evidentemente, con la utilización de frases leídas se ha perdido naturalidad en las realizaciones, especialmente en las formas exclamativas y volitivas, pero este hecho parece relativamente secundario en un estudio sobre la modalidad oracional, teniendo en cuenta que las frases leídas también contienen este tipo de información, al menos por lo que respecta a las curvas de Fo, que eran el objetivo del estudio.

Las variables que se han considerado a la hora de elaborar el *corpus* son:

a) La modalidad oracional:

Para la determinación de las modalidades oracionales incluidas en este estudio se ha tomado como referencia el trabajo de Navarro (1948). Así, al igual que hace Navarro, se han diferenciado cuatro tipos de frases en el *corpus* :

- 1) Frases enunciativas, desde la 1.1 a la 1.6.
- 2) Frases interrogativas: se han distinguido finalmente tres subgrupos, siguiendo la ya mencionada propuesta de Canellada y Kuhlmann (1987, pág. 147 y ss.):
 - 2.1 Interrogativas absolutas (frases 2.1.1 - 2.1.6).
 - 2.2 Interrogativas pronominales (frases 2.2.1 - 2.2.6).
 - 2.3 Interrogativas relativas (frases 2.3.1 - 2.3.12).
- 3) Frases exclamativas, también divididas en dos subgrupos:
 - 3.1 Exclamativas de admiración (frases 3.1. - 3.6).
 - 3.2 Exclamativas de desaprobación (frases 3.7 - 3.12).
- 4) Frases volitivas, que incluyen igualmente dos subtipos:
 - 4.1 Volitivas de orden (frases 4.1 - 4.6).
 - 4.2 Volitivas de ruego o recomendación (frases 4.7 - 4.12).

b) El número de sílabas de la frase:

Aunque el procedimiento de estilización empleado no toma la sílaba como unidad de análisis, se ha considerado interesante tener en cuenta como variable el número de sílabas de las frases. Incluyendo frases con distinto número de sílabas en el *corpus* se ha podido analizar, entre otras cosas, si la forma de las curvas melódicas depende del número total de sílabas, o del número de sílabas tónicas de la frase.

Atendiendo a este criterio, las frases del *corpus* pueden dividirse en dos grandes grupos:

- 1) Las que contienen más de una sílaba acentuada (dos o tres, según los casos), que serían todas las frases que van, dentro de cada grupo, del 1 al 4 (por ejemplo, de la 1.1 a la 1.4), y en los grupos que existen, de la 7 a la 10 (por ejemplo, de la 2.3.7 a la 2.3.12).
- 2) Las que contienen una sola sílaba acentuada, y que son las que están marcadas con los números 5, 6, 11 y 12 (estas dos últimas, no presentes en todos los grupos).

c) Posición de los acentos inicial y final en la frase:

Para poder estudiar la relación entre los puntos de inflexión de la curva y la posición de las sílabas tónicas en la frase - relación que se ha puesto

reiteradamente de manifiesto en estudios anteriores -, se ha procurado que en el *corpus* hubiera frases en las que el primer acento recaiga sobre la sílaba inicial de la frase, y frases en las que recaiga sobre la segunda sílaba. Igualmente, se han incluido frases cuyo acento final recae en la última sílaba de la frase, y frases con el acento final en la penúltima sílaba.

Se ha intentado además que estas cuatro posibilidades - acento inicial en la primera sílaba, acento inicial en la segunda, acento final en la última sílaba y acento final en la penúltima - aparezcan combinadas en el *corpus*. Esto ha dado como resultado cuatro esquemas acentuales diferentes para cada tipo de frase, especificados en las tablas 4.1 - 4.6.

En las frases con un solo acento, lógicamente sólo se ha tenido en cuenta la posición de la sílaba inicial, de forma que se incluyen sólo dos esquemas acentuales.

Finalmente, se ha optado por la utilización de frases compuestas únicamente por **sonidos sonoros**, con el fin de conseguir curvas sin rupturas. También se ha procurado que, dentro de cada grupo, las frases presenten sólo los cambios imprescindibles en cuanto a los elementos segmentales que las componen, para evitar en lo posible la influencia de las variaciones micromelódicas en las representaciones obtenidas, como puede observarse en los ejemplos siguientes (las negritas indican la parte de cada frase que coincide en sus elementos segmentales con la otra frase del par):

- a) **Ramón** llegó en avión. (1.1)
Ramón llegó andando. (1.2)

- b) **Llama a Ramón**. (4.1)
Llamad a Ramón. (4.3)

| Frase | Acento inicial | | Acento final | |
|-------|----------------|-----------|---------------|-------------|
| | 1ª sílaba | 2ª sílaba | Penul. sílaba | Ult. sílaba |
| 1.1 | | X | | X |
| 1.2 | | X | X | |
| 1.3 | X | | | X |
| 1.4 | X | | X | |
| 1.5 | | X | | |
| 1.6 | X | | | |

Tabla 4.1.- Posición de los acentos en las frases enunciativas

| Frase | Acento inicial | | Acento final | |
|-------|----------------|-----------|---------------|-------------|
| | 1ª sílaba | 2ª sílaba | Penul. sílaba | Ult. sílaba |
| 2.1.1 | X | | | X |
| 2.1.2 | X | | X | |
| 2.1.3 | | X | | X |
| 2.1.4 | | X | X | |
| 2.1.5 | X | | | |
| 2.1.6 | | X | | |

Tabla 4.2.- Posición de los acentos en las frases interrogativas absolutas

| Frase | Acento inicial | | Acento final | |
|-------|----------------|-----------|---------------|-------------|
| | 1ª sílaba | 2ª sílaba | Penul. sílaba | Ult. sílaba |
| 2.2.1 | X | | | X |
| 2.2.2 | X | | X | |
| 2.2.3 | | X | | X |
| 2.2.4 | | X | X | |
| 2.2.5 | X | | | |
| 2.2.6 | | X | | |

Tabla 4.3.- Posición de los acentos en las frases interrogativas pronominales

| Frase | Acento inicial | | Acento final | |
|--------|----------------|-----------|---------------|-------------|
| | 1ª sílaba | 2ª sílaba | Penul. sílaba | Ult. sílaba |
| 2.3.1 | | X | | X |
| 2.3.2 | | X | X | |
| 2.3.3 | X | | | X |
| 2.3.4 | X | | X | |
| 2.3.5 | | X | | |
| 2.3.6 | X | | | |
| 2.3.7 | X | | | X |
| 2.3.8 | X | | X | |
| 2.3.9 | | X | | X |
| 2.3.10 | | X | X | |
| 2.3.11 | X | | | |
| 2.3.12 | | X | | |

Tabla 4.4.- Posición de los acentos en las frases interrogativas relativas

| Frase | Acento inicial | | Acento final | |
|-------|----------------|-----------|---------------|-------------|
| | 1ª sílaba | 2ª sílaba | Penul. sílaba | Ult. sílaba |
| 3.1 | | X | | X |
| 3.2 | | X | X | |
| 3.3 | X | | | X |
| 3.4 | X | | X | |
| 3.5 | | X | | |
| 3.6 | X | | | |
| 3.7 | | X | | X |
| 3.8 | | X | X | |
| 3.9 | X | | | X |
| 3.10 | X | | X | |
| 3.11 | | X | | |
| 3.12 | X | | | |

Tabla 4.5.- Posición de los acentos en las frases exclamativas

| Frase | Acento inicial | | Acento final | |
|-------|----------------|-----------|---------------|-------------|
| | 1ª sílaba | 2ª sílaba | Penul. sílaba | Ult. sílaba |
| 4.1 | X | | | X |
| 4.2 | X | | X | |
| 4.3 | | X | | X |
| 4.4 | | X | X | |
| 4.5 | X | | | |
| 4.6 | | X | | |
| 4.7 | X | | | X |
| 4.8 | X | | X | |
| 4.9 | | X | | X |
| 4.10 | | X | X | |
| 4.11 | X | | | |
| 4.12 | | X | | |

Tabla 4.6.- Posición de los acentos en las frases volitivas

4.3.2. Grabación del *corpus*

Se procuró, tal como se ha hecho en otros trabajos - Caelen (1981, pág. 57), por ejemplo -, que los locutores escogidos para la grabación tuvieran unas características semejantes en cuanto a su edad, procedencia geográfica y nivel cultural. Para ello, se establecieron los siguientes criterios de selección:

- a) Que los locutores tuvieran edades comprendidas entre los 20 y los 35 años.
- b) Que fueran nacidos en Cataluña, pero cuya lengua de uso principal, al menos en el medio familiar, fuera el español. De esta manera se procuró evitar al máximo la influencia de la entonación catalana en las frases grabadas.
- c) Que tuvieran un nivel cultural medio-alto, es decir, que hubieran realizado, como mínimo, estudios a nivel de C.O.U.

De acuerdo con estos criterios, se seleccionaron 8 locutores, 4 hombres y 4 mujeres, con los que se procedió a realizar la grabación. Dichos locutores se identifican como L1, L2, L3, L4, L5, L6, L7 y L8.

De esta manera, cada frase analizada se etiquetará, de aquí en adelante, con el número que le corresponde por su tipo, como ya se ha explicado anteriormente - por ejemplo, 1.1 -, más el identificador correspondiente al locutor - por ejemplo, L1 - de la forma siguiente: 1.1/L1.

Las grabaciones se llevaron a cabo en la cámara insonorizada del Laboratorio de Fonética de la U.A.B., con un magnetófono REVOX A77 y un micrófono direccional SENNHEISER MD 441N, procurando que la distancia entre el locutor y el micrófono fuera constante - entre 15 y 20 cms.. El procedimiento seguido para la grabación fue el siguiente:

- a) Cada locutor leyó todo el *corpus* una vez, antes de realizar la grabación, para familiarizarse con el mismo.
- b) Después, se les instruyó sobre los detalles que habían de tener en cuenta durante la lectura (la separación entre diálogos, por ejemplo), y se procedió a la grabación, controlando los posibles errores del locutor.
- c) Finalizada la misma, en el que caso de que se hubieran detectado errores, se realizó una segunda grabación sólo de los diálogos con errores.

4.3.3. Análisis del *corpus*

Para la digitalización y análisis de las diferentes frases se ha utilizado un programa de análisis de señal estándar del mercado, MACSPEECH LAB, para ordenadores MACINTOSH II o superiores. El programa permite digitalizar, editar, reproducir y analizar señales de habla, e incluye un algoritmo de estimación de la Fo. De acuerdo con las indicaciones del manual de uso (GWI, 88, pág. 94), es un sistema basado en el análisis temporal de la onda, concretamente en la autocorrelación: "*The fundamental frequency is determined using a modified auto-correlation technique. The auto-correlation function displays local maxima at time corresponding to the inverse of Fo*". El analizador de Fo tiene dos modos de funcionamiento, uno rápido y otro lento. El primero consigue velocidades de análisis mayores analizando únicamente la parte central de la onda: "*The Fast method uses a center clipper to reduce the time waveform amplitudes to -1, 0 and 1*" (GWI, 1988, pág 94). Durante el análisis se emplearon ambos métodos indistintamente, buscando para cada frase el menor número posible de errores en la detección de la Fo.

El análisis se ha realizado únicamente sobre 4 de los 8 locutores grabados, concretamente L1, L2, L3 y L5, es decir, 2 hombres y 2 mujeres. Los otros 4 locutores se han reservado para posteriores trabajos.

El primer paso ha consistido en filtrar la señal a través de un filtro *anti-aliasing* y digitalizarla a una frecuencia de muestreo de 10 KHz. Una frecuencia de muestreo más alta habría permitido obtener una mayor precisión en la estimación de la Fo, pero se ha escogido ésta para no alargar en exceso el tiempo de cálculo. Algunas pruebas de análisis con frases muestreadas a 20 KHz han indicado que la diferencia no es tan grande como para escoger una frecuencia mayor.

Una vez digitalizada la señal, se ha procedido a la extracción de las curvas de Fo, utilizando el algoritmo del programa. Durante esta fase, algunas de las frases del *corpus* han sido desechadas, pues debido a la gran cantidad de errores en el cálculo de la Fo ha resultado imposible efectuar la medición de los parámetros estudiados.

Tras obtener las curvas melódicas se ha procedido a aplicar el proceso de estilización descrito en el capítulo anterior. Hay que señalar que, si bien en un principio se estableció el umbral en 10 Hz para considerar una variación de signo como punto de inflexión, en algunos casos ha habido que rebajarlo para no eliminar puntos que se consideraron importantes, especialmente en frases de corta duración. Por otro lado, el análisis posterior de las representaciones obtenidas ha revelado que algunas de las pequeñas variaciones (15-25 Hz) que se han mantenido en las mismas coinciden con oscilaciones de la Fo de tipo micromelódico, que podrían ser eliminadas. Todo esto parece indicar que el umbral de 10 Hz ha de ser revisado de cara a próximos trabajos.

Además del proceso de estilización, se ha aplicado a las frases estudiadas una **normalización temporal** que permitiera estudiar la posición de los diferentes puntos de inflexión independientemente de la duración de la curva. Igualmente, tal

proceso ha permitido eliminar las variaciones debidas al *tempo* particular de elocución de cada hablante. Los valores temporales se han relativizado con respecto al inicio (valor 0) y al final (valor 1). El método de normalización utilizado es el que se describe a continuación:

- a) **Normalización con respecto al punto de inicio.** Se resta a todos los valores temporales de cada curva el valor temporal correspondiente al inicio de la misma. De esta manera el valor del punto de inicio pasa siempre a ser 0, y los demás valores temporales de la curva se adaptan convenientemente.
- b) **Normalización con respecto al final,** consistente en dividir todos los valores temporales de la curva por el valor temporal del final resultante en la normalización anterior.

Un ejemplo de las representaciones estilizadas obtenidas, junto con la curva melódica de la que se ha extraído, puede hallarse en la figura 4.7.

Una vez obtenidas las representaciones, se ha procedido a la separación de las mismas, dentro de cada grupo, según el signo de su pendiente final, es decir, según tuvieran un final ascendente o descendente. De esta forma se ha obtenido un total de 12 grupos diferentes de curvas, sobre los cuales se han realizado nuevos análisis.

Dichos análisis se han centrado en las dos partes de la curva melódica que, como se ha señalado anteriormente, se consideran más ricas en información entonativa: la **parte inicial** - desde el inicio hasta la primera sílaba acentuada de la curva - y la **parte final** - el último segmento antes del final de la curva:

- 1) Análisis de la parte inicial de cada curva, concretamente del **primer segmento con pendiente ascendente** de la misma, que ha comprendido:
 - a) un estudio estadístico de los valores de F_0 en el primer pico de la curva.
 - b) un estudio estadístico de los valores de la pendiente en la subida hasta este primer pico.

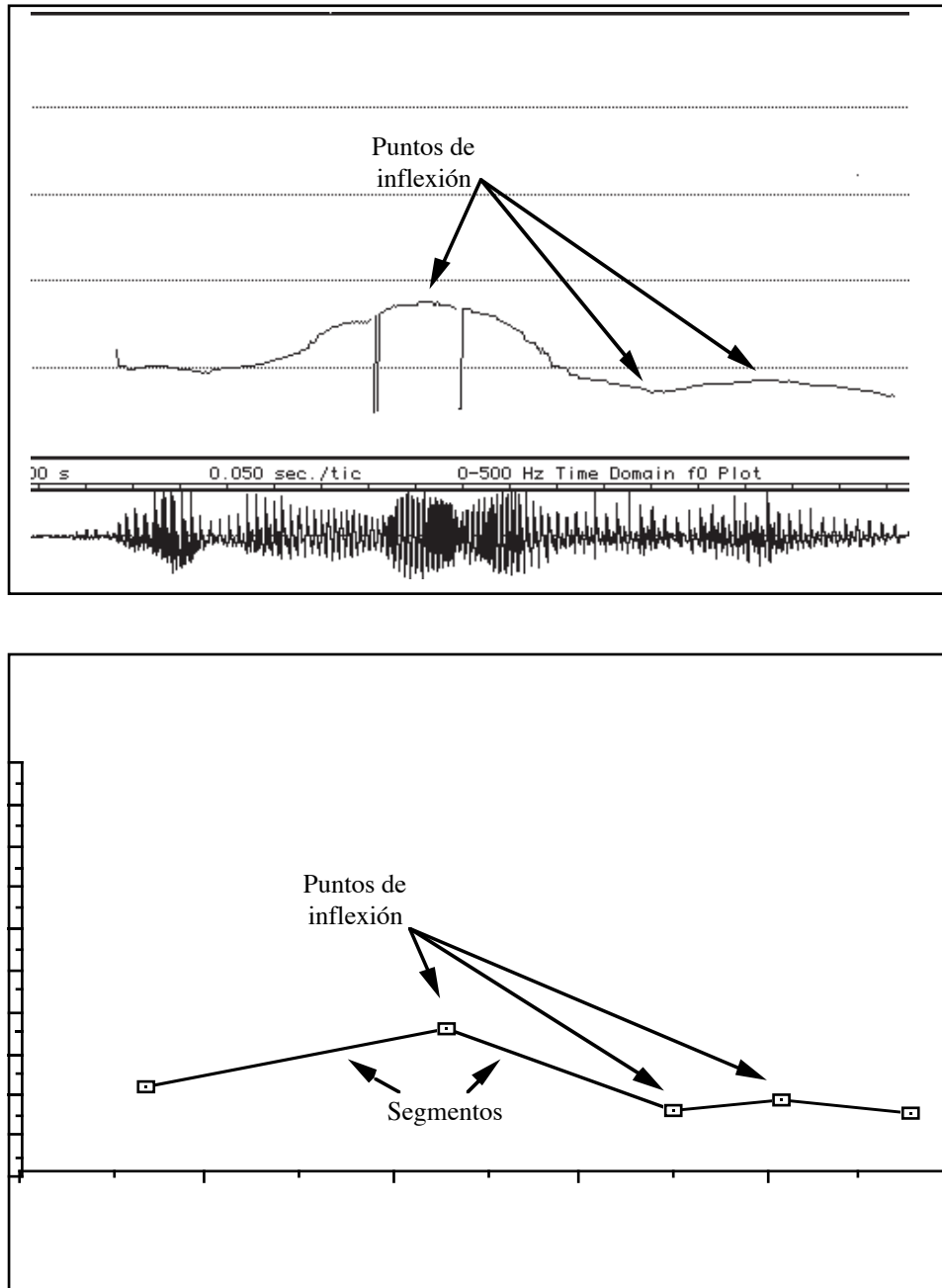


Figura 4.7.- curva original y representación estilizada de la frase 1.1/L1

2) Análisis del **segmento final** de la curva:

- a) estudio estadístico del valor de Fo al final de dicho segmento.
- b) estudio estadístico de los valores de la pendiente en el segmento.

También se ha realizado un estudio del **rango de la Fo** dentro de cada curva, parámetro relacionado tradicionalmente con la transmisión de información expresiva.

Para el cálculo de la pendiente, tanto en el segmento inicial como en el final, se ha utilizado la fórmula siguiente:

$$p = F_{o_f} - F_{o_i} / T_f - T_i$$

donde p es la pendiente, F_{o_f} es el valor de la Fo al final del segmento estudiado, F_{o_i} es el valor de la Fo al inicio del segmento, T_f es el instante de tiempo en que acaba el segmento, y T_i el instante de tiempo en que comienza. El resultado se expresa en Hz/ms.

Hay que señalar que el cálculo de las pendientes se ha realizado con los valores temporales reales, previos a la aplicación de la normalización temporal. Para el estudio de los valores de la Fo en el primer pico y al final de la curva se han tomado los resultados obtenidos tras la normalización frecuencial descrita anteriormente.

En el análisis del rango frecuencial, se ha procedido a buscar el valor de Fo más alto y bajo dentro de cada curva, obteniéndose el rango simplemente mediante la resta entre los dos valores.

En los tres casos, el análisis estadístico ha consistido en el cálculo de la media, en la determinación de los valores máximo y mínimo para cada grupo, y en un análisis de la distribución de los valores mediante histogramas. También se han realizado tests estadísticos para evaluar la relación entre grupos de datos.

4.4. Resultados

Los resultados del proceso de estilización han permitido efectuar un primer análisis visual de los esquemas de cada modo oracional. Se presentan ahora, grupo por grupo, dichos resultados. También se describen los resultados de los diferentes análisis estadísticos.

4.4.1. Frases enunciativas (1.1 - 1.6)

La práctica totalidad de las frases de este grupo presenta un **segmento final descendente**. Sólo en 2 casos, 1.1/L.2 y 1.2/L.5, de los 23 analizados, se observa un pequeño ascenso al final. En función de este rasgo se han establecido dos subgrupos diferentes.

El análisis visual de los esquemas en ambos subgrupos revela la presencia, tanto en uno como en otro, de como mínimo una subida importante de la Fo en algún punto de la curva. En las frases con una sola sílaba acentuada, se aprecia en la mayoría de los casos sólo una subida, situada en las proximidades de la sílaba acentuada, como se puede observar en la figura 4.8. En las frases con más de una sílaba acentuada, la subida más importante está situada igualmente en la primera sílaba tónica o cerca de ella; aparecen además una serie de picos secundarios, localizados aproximadamente, según el locutor, en la última o penúltima sílaba acentuada, tal como se ejemplifica en la figura 4.9.

La altura de la Fo con respecto al inicio en su primera subida presenta bastantes variaciones: en las frases con tonema final descendente, el rango de valores medidos va desde los - 4 hasta los 152 Hz normalizados; en las 2 frases con final ascendente, los valores oscilan entre 30 y 107 Hz por encima del valor de referencia al inicio de la frase, lo que revela también bastante variabilidad. Analizando la distribución de estos valores en el histograma, se distinguen dos grandes bloques, uno entre - 4 y 30 Hz aproximadamente, y otro con valores superiores a 40 Hz. El valor medio de la Fo en este primer pico es, en las frases con final descendente, de 59 Hz.

El análisis de la pendiente de esta primera subida presenta igualmente una fuerte dispersión de los valores obtenidos (entre 0,11 y 0,60 Hz/ms.), según se desprende del histograma. El valor medio obtenido para las pendientes - teniendo en cuenta sólo las frases con final descendente - es de 0,39 Hz/ms.

Es de destacar igualmente el hecho de que también en español se cumple el fenómeno de la declinación, según se desprende de la observación de las frases estudiadas (véanse ejemplos de la figura 4.10). En todas las frases enunciativas analizadas con más de un pico, salvo en 1.5/L.5, el primer pico de cada frase es el que presenta un valor de la Fo más alto, mientras que los demás, a medida que se aproxima el final de la frase, presentan un valor cada vez más bajo.

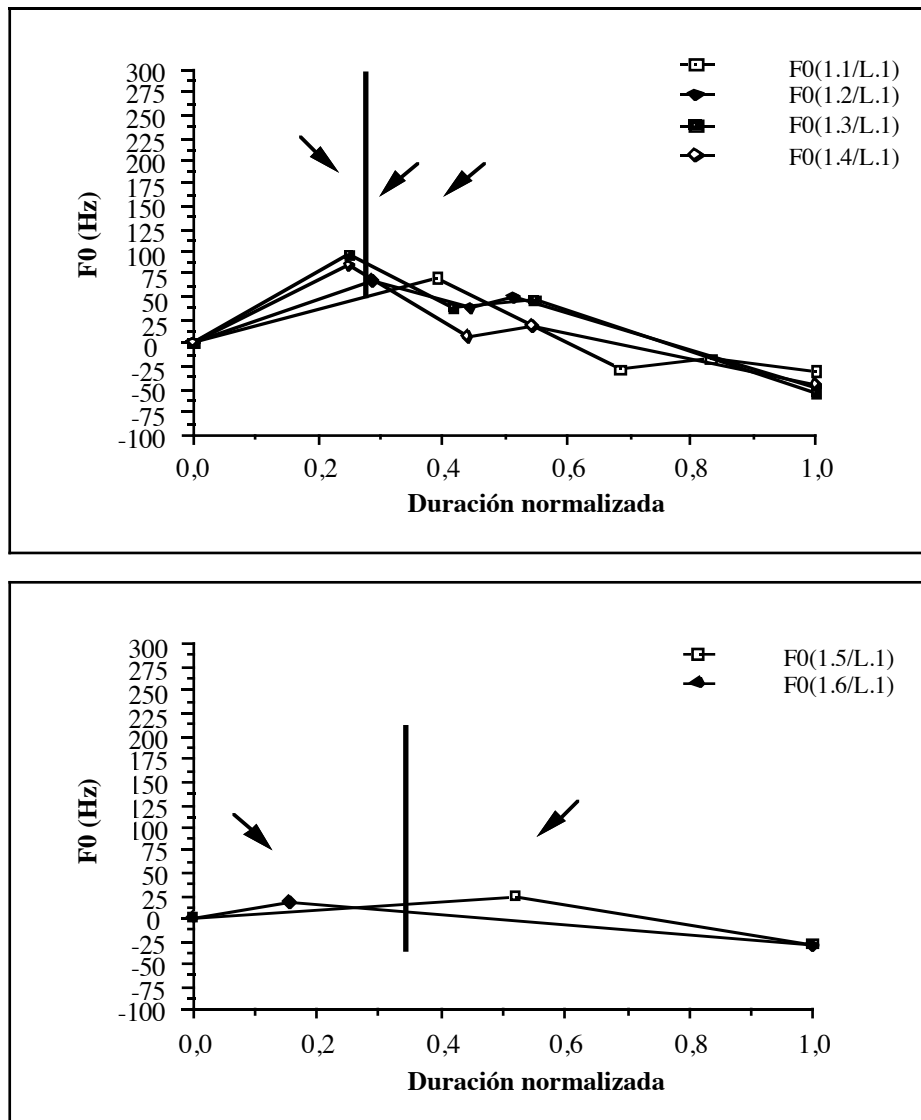


Figura 4.8.- La presente figura ilustra las diferencias existentes en la posición del primer pico según cuál sea la primera sílaba acentuada en las frases enunciativas. En los dos gráficos puede observarse cómo, en general, aquellas frases en las que la primera sílaba es tónica presentan el primer pico más cerca del inicio de la frase que aquellas cuya primera sílaba tónica se encuentra en la segunda sílaba.

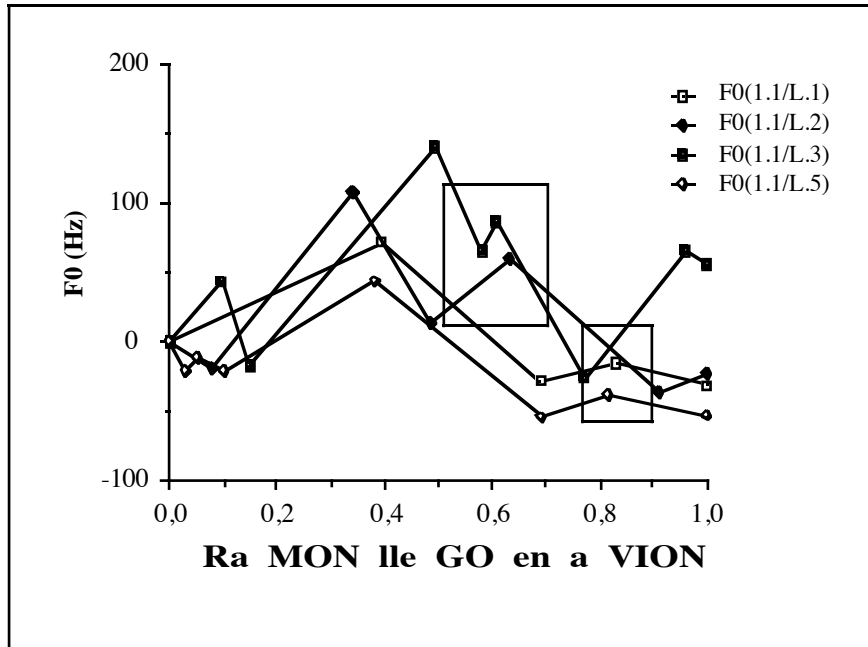


Figura 4.9.- En la presente figura se muestra un ejemplo de cómo en una misma frase, el pico secundario puede recaer en sílabas acentuadas diferentes. En la frase 1.1, "Ramón llegó en avión", dos locutores marcaron el segundo pico en las proximidades de la segunda sílaba acentuada de la frase (1.1/L.2 y 1.1/L.3), en tanto que los otros dos lo hicieron en la tercera (1.1/L.1 y 1.1/L.5). La posición de dicho pico en los dos casos está marcada con un recuadro.

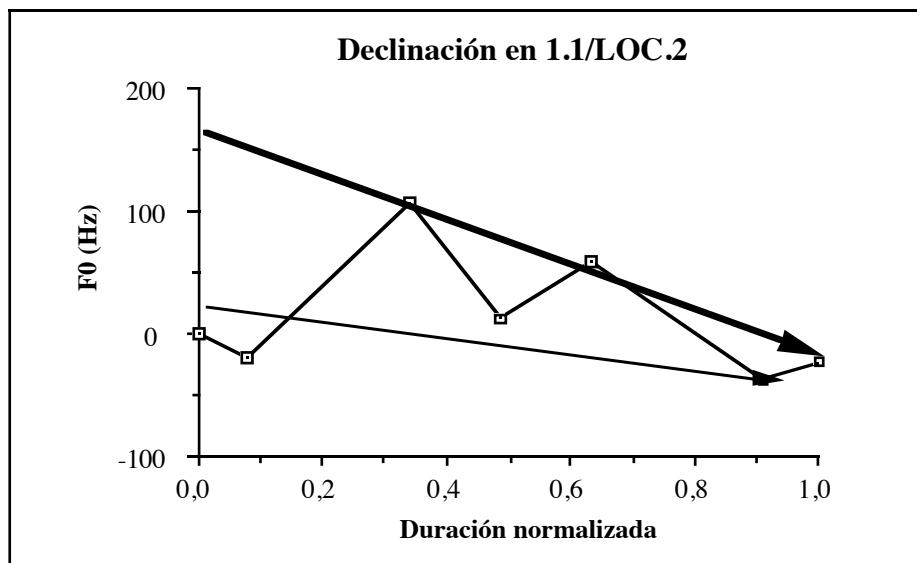
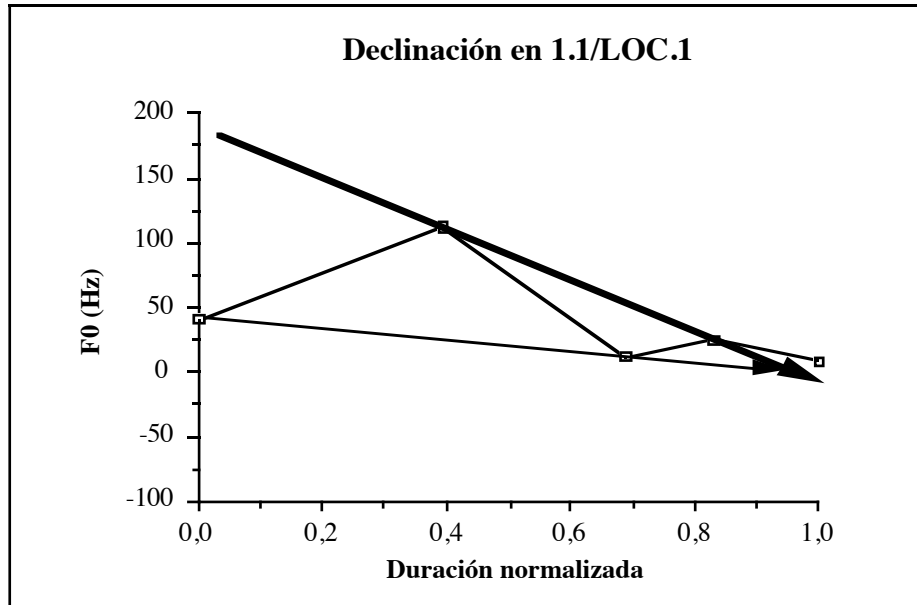
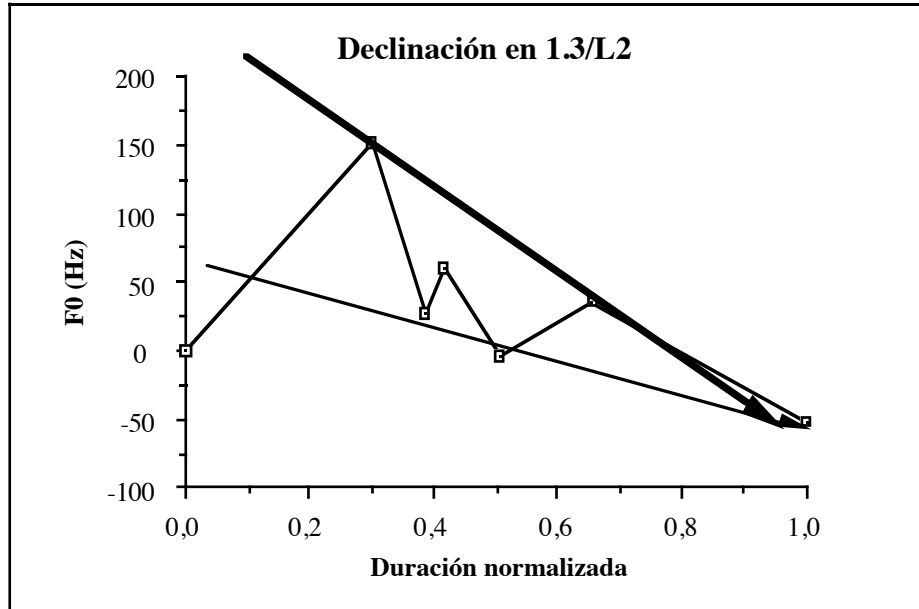
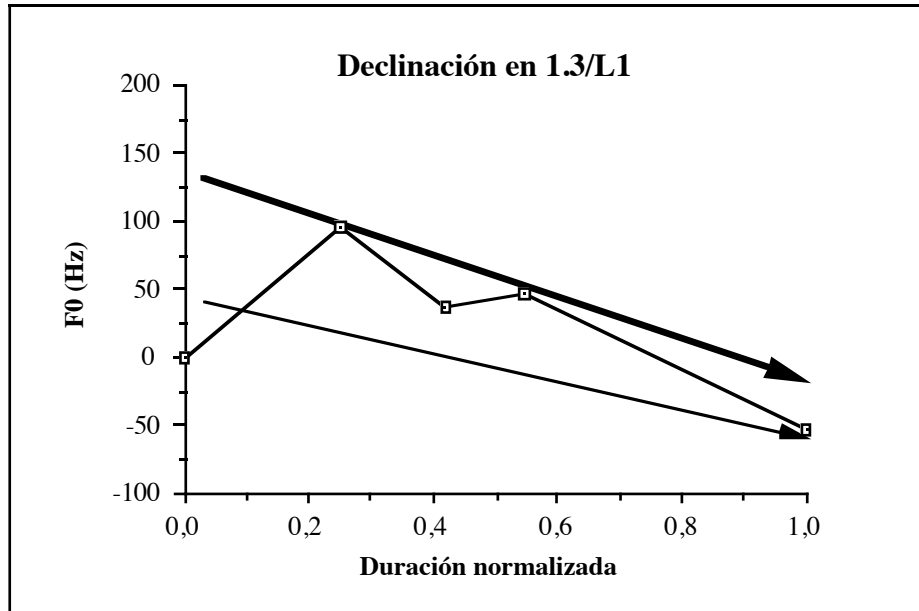


Figura 4.10.- En la presente figura se muestran diferentes ejemplos del fenómeno de la declinación en las frases enunciativas del español. La línea más gruesa une los diferentes picos de la curva. La más fina representaría la línea de base.



El principio del segmento final presenta algunas variaciones de una frase a otra, dependiendo bastante de la posición del pico secundario: mientras que en unas frases comienza en la última sílaba acentuada, tal como señala la bibliografía, en algunos casos se inicia en la penúltima. Así ocurre en 1.1/L2, 1.1/L3, 1.2/L1, 1.2/L5, 1.3/L1, 1.3/L2, 1.3/L5, 1.4/L1 y 1.4/L5.

La pendiente de este segmento final oscila, cuando ésta es negativa, entre los - 0,08 y los - 0,66 Hz/ms., aunque la dispersión es menor que en el caso de la primera subida. El valor medio para dicha pendiente es de - 0,27 Hz/ms.

En cuanto a los valores de Fo al final de este segmento, éstos oscilan entre los - 84 y los 55 Hz. El histograma muestra que los valores están bastante agrupados alrededor de los - 50 Hz.

Los dos casos de pendiente positiva encontrados no parecen tener mucho en común. Hay que reseñar, sin embargo, que aunque su final es ascendente, el valor de la Fo al final de la curva se sitúa por debajo del 0, correspondiente al valor de la Fo al principio de la curva (- 66 y - 24 Hz exactamente). Este sería un hecho que acercaría a ambos grupos, y que quizá podría indicar que para etiquetar una frase como enunciativa bastaría con que su valor de Fo final esté por debajo del inicial. Sin embargo, son muy pocos los datos como para poder afirmarlo con rotundidad.

Finalmente, el rango frecuencial que abarca por término medio este tipo de frases es de 112 Hz, teniendo sólo en cuenta las frases con final descendente, aunque se aprecian bastantes diferencias en función de la longitud, como puede comprobarse en la figura 4.11 (mayor rango cuanto más larga sea la frase), aspecto que queda reflejado también en los dos picos que aparecen en el correspondiente histograma.

4.4.2. Frases interrogativas absolutas (2.1.1 - 2.1.6)

Sin duda, el rasgo más característico de las interrogativas absolutas reside, tal como ha señalado repetidamente la bibliografía, en la pronunciada subida que se da en su segmento final. En el presente estudio, sólo 3 frases de las 24 analizadas presentan una pequeña bajada al final, tras una subida pronunciada: 2.1.3/L3, 2.1.4/L3, y 2.1.2/L5. El resto presentan un segmento final claramente ascendente.

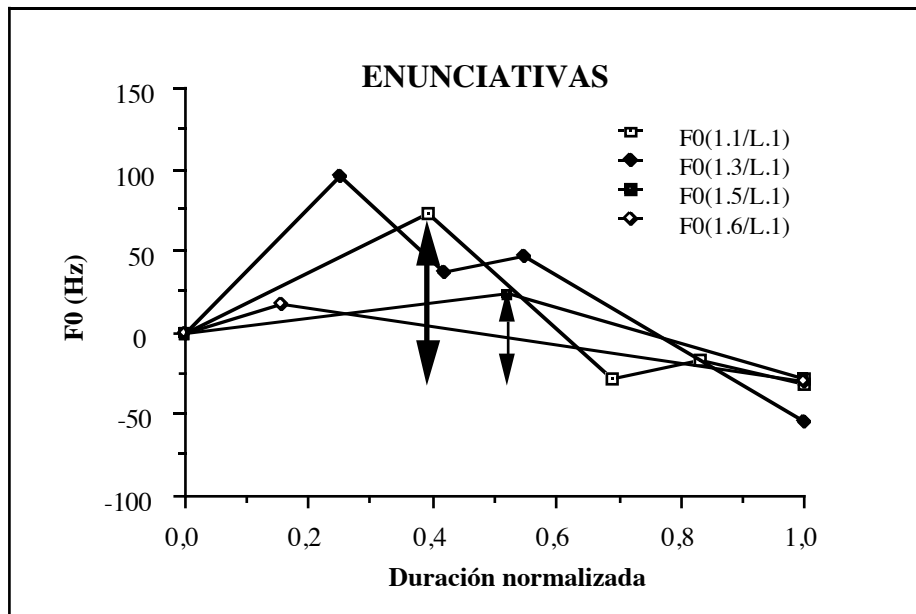


Figura 4.11.- La presente figura ilustra las diferencias existentes en el rango frecuencial, para las frases enunciativas, entre las frases largas y las frases cortas. La línea más gruesa indica el rango en las frases largas (que en el cuadro son 1.1/L1 y 1.3/L1), y la más delgada el rango en las frases cortas (1.5/L1 y 1.6/L1). Se puede apreciar cómo las frases cortas presentan un rango menor que las largas.

Las representaciones correspondientes a estas frases muestran, al igual que en el caso de las enunciativas, la presencia de una subida importante en las proximidades de la primera sílaba acentuada, según se desprende de la figura 4.12. Sin embargo, hay que señalar el especial comportamiento de las oraciones con una sola sílaba acentuada, en las que esta subida inicial casi no se aprecia (2.1.6/L2) o simplemente no existe (2.1.6/L3).

El valor medio de la pendiente en esta subida es 0,65 Hz/ms., algo superior al de las enunciativas. La distribución de los valores, según el histograma, se acerca bastante a una distribución normal. Por lo que se refiere a la altura tonal de esta primera subida, el valor más alto en estas oraciones es de + 171 Hz con respecto a la Fo inicial; el más bajo, + 17 Hz. El valor medio, teniendo en cuenta sólo las frases con final ascendente, es de 96 Hz, un valor claramente superior al de las enunciativas, y el histograma, al igual que en éstas, revela la presencia de dos grandes grupos, entre 17 y 64 Hz, y entre 94 y 171 Hz, respectivamente.

A diferencia de las frases enunciativas, no se aprecia normalmente un segundo pico antes del final de la frase. Sólo en algún caso aislado (2.1.2/L1) se ha encontrado este pico secundario.

Como ya se ha señalado, en la gran mayoría de los casos el segmento final de estas frases es ascendente. Los valores de la pendiente en el último segmento oscilan entre los 0,42 y los 1,62 Hz/ms., cuando es ascendente, y entre los - 0,52 y los - 0,34 Hz/ms. en los tres casos que es descendente. Los valores tienden a agruparse, en las frases con pendiente positiva, alrededor de los 0,8 Hz/ms. (valor medio 0,87).

Los valores de la Fo al final de esta subida oscilan entre 294 Hz (2.1.5/L5) y 52 Hz (2.1.2/L3) por encima del tono inicial. Hay que destacar también el hecho de que son las frases más cortas las que alcanzan los valores de Fo más altos en los 4 locutores.

También hay que destacar el hecho de que las 3 curvas con final descendente presentan valores por encima del 0 inicial (valores entre los + 73 y los + 91 Hz normalizados), lo cual reforzaría la hipótesis planteada para las enunciativas: inversamente a éstas, en las interrogativas absolutas bastaría con que el valor final de la Fo presente una cierta elevación sobre el valor inicial para que la frase se etiquete como tal interrogativa, aunque el último segmento sea descendente.

Merece señalarse igualmente que el punto de inicio de esta subida parece bastante independiente de la posición del acento final en la frase, según se observa en la figura 4.13: en las frases poliaccentuales, se sitúa entre el 63,3% (2.1.2/L2) y el 80,9% (2.1.3/L2) de la longitud total de la frase; en las monoacentuales, entre el 38,1% (2.1.5/L3) y el 54,7% (2.1.6/L2). También es bastante fija su altura tonal: se sitúa por lo general unos pocos Hertzios por debajo del valor de Fo al principio de la frase. En principio este punto parece coincidir con el inicio de la última sílaba, sea ésta tónica o átona.

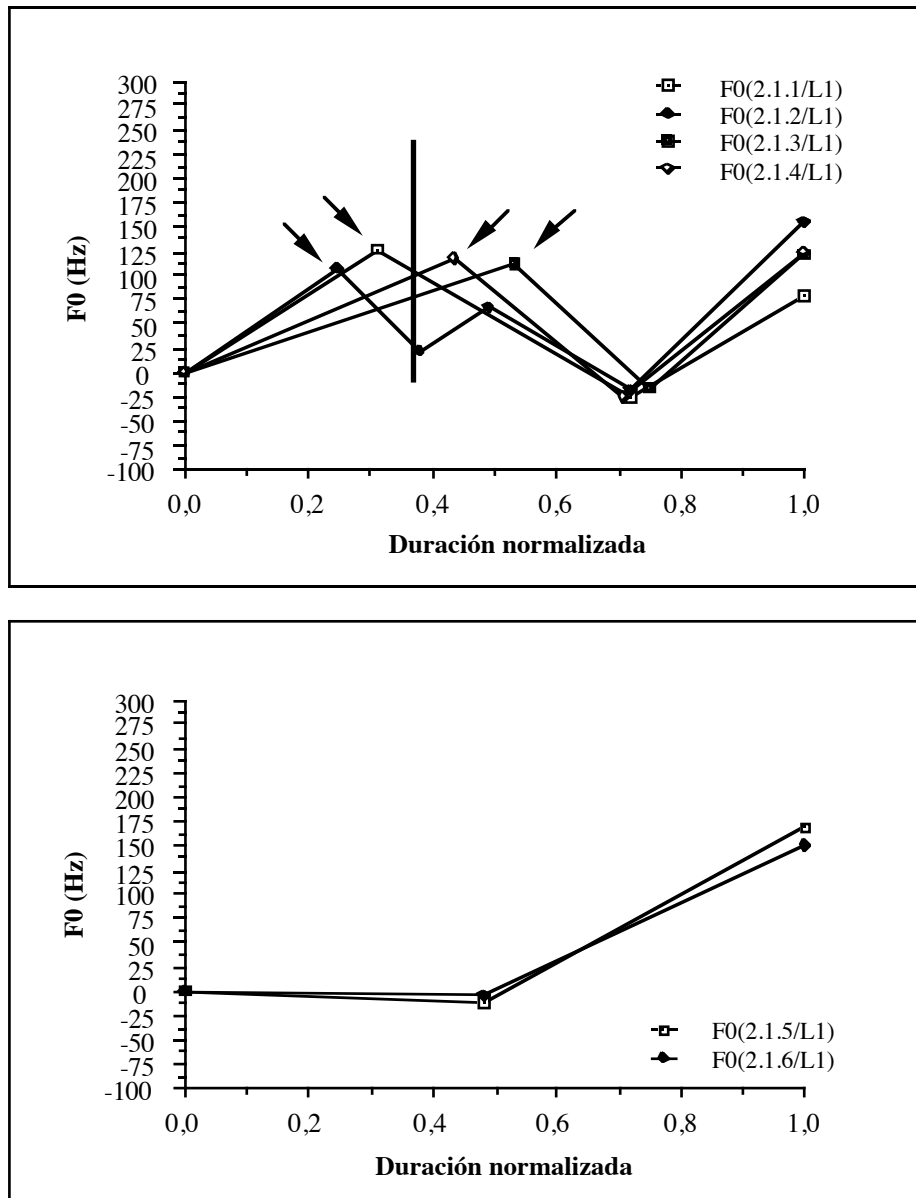


Figura 4.12.- Diferencias existentes en la posición del primer pico según cuál sea la primera sílaba acentuada en las frases interrogativas absolutas. El cuadro superior muestra cómo las frases con el acento en la primera sílaba presentan el primer pico antes que las frases con el acento en la segunda. El cuadro inferior ilustra la ausencia de pico inicial en las frases cortas.

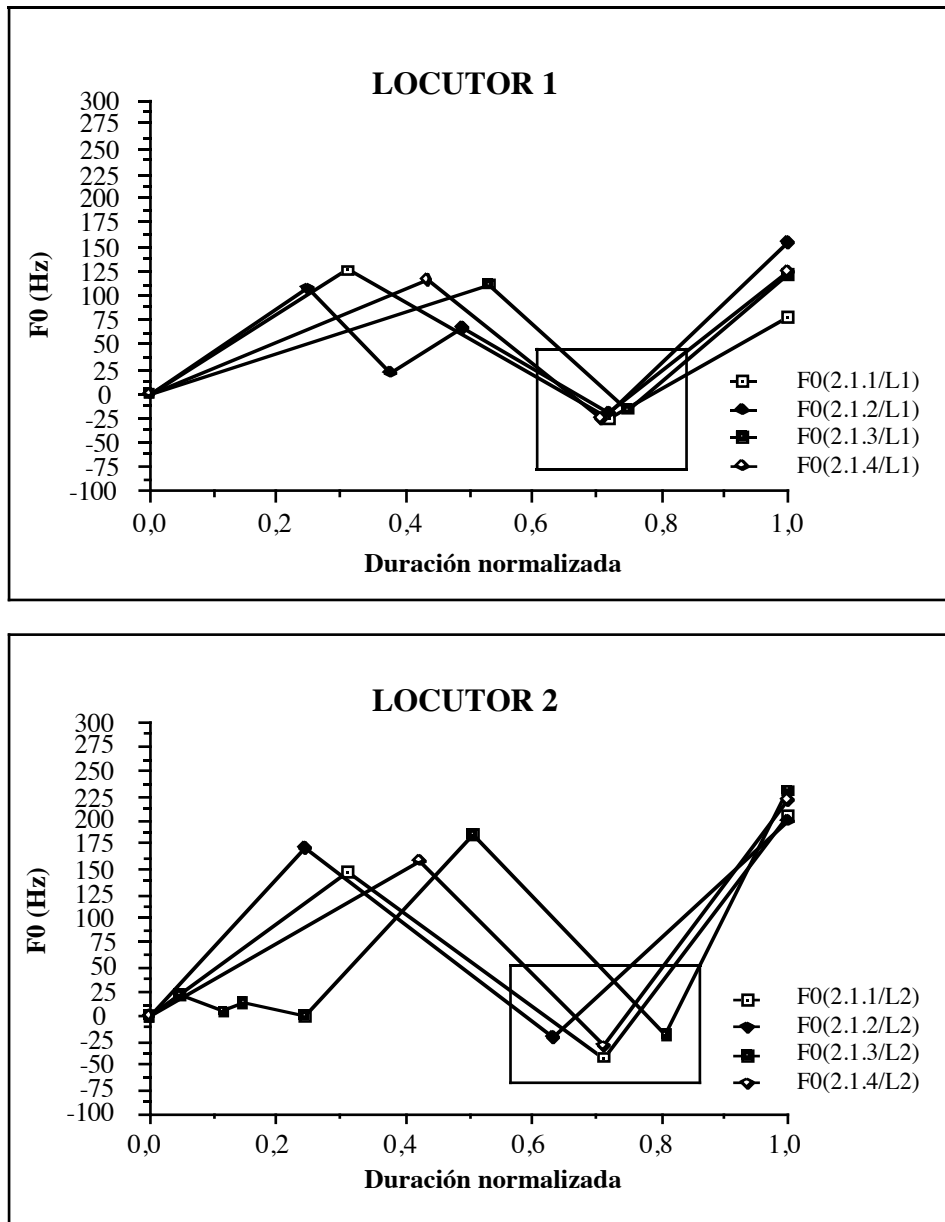


Figura 4.13.- En la presente figura se muestran las semejanzas en el punto de inicio del segmento final en las interrogativas absolutas. La posición de este punto se localiza aproximadamente al comienzo de la última sílaba, sea ésta tónica o átona.

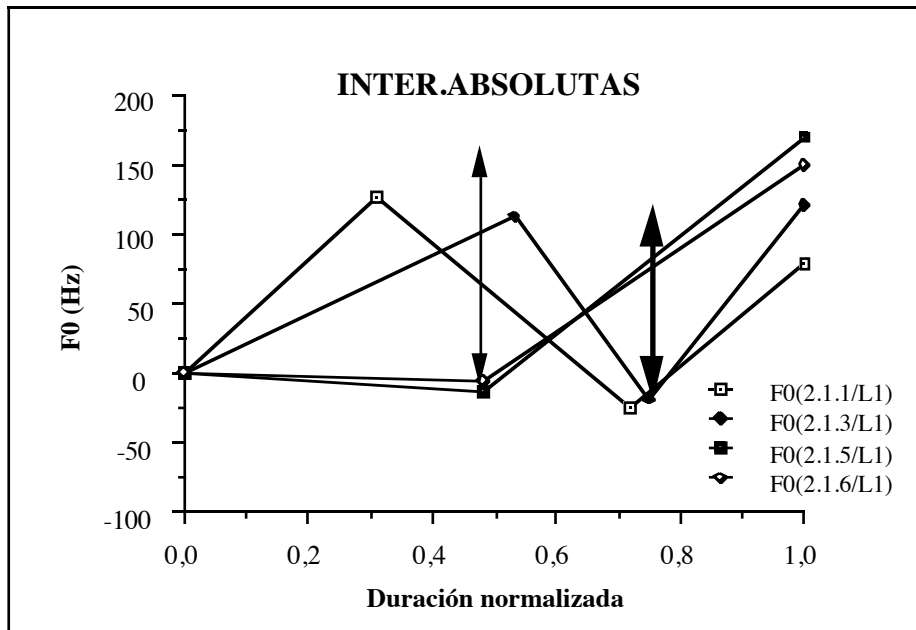


Figura 4.14.- La presente figura ilustra las diferencias existentes en el rango frecuencial, para las interrogativas absolutas, entre las frases largas y las frases cortas. Al igual que en 4.11, la línea más gruesa indica el rango en las frases largas (que en el cuadro son 2.1.1/L1 y 2.1.3/L1), y la más delgada el rango en las frases cortas (2.1.5/L1 y 2.1.6/L1). Se puede apreciar cómo el rango incluso es mayor en las frases cortas que en las largas.

El rango frecuencial comprendido por este tipo de frases es marcadamente más elevado, en términos generales, que en las enunciativas: 206 Hz por término medio, si se tienen sólo en cuenta las frases con final ascendente. Sin embargo, el histograma refleja una dispersión bastante acusada de los valores. Tampoco se han encontrado las diferencias debidas a la longitud de la frase que se observaban en las enunciativas; incluso en ocasiones el rango es mayor en las frases cortas que en las largas (figura 4.14).

4.4.3. Frases interrogativas pronominales (2.2.1 - 2.2.6)

Tal como ya señalaban Navarro Tomás y otros autores, los resultados de este estudio indican que los hablantes del español pueden escoger entre dos modelos melódicos diferentes en el momento de formular preguntas de este tipo.

Los esquemas correspondientes a las frases pronominales del *corpus* pueden separarse en dos grupos claramente diferenciados:

1) Esquema A:

Es un esquema que presenta ciertas semejanzas con el de las enunciativas, puesto que se caracteriza por un **segmento final descendente**.

Las frases que presentan este esquema en nuestro *corpus* son:

| | | | |
|----------|----------|----------|----------|
| 2.2.2/L1 | 2.2.1/L2 | 2.2.1/L3 | 2.2.5/L5 |
| 2.2.4/L1 | 2.2.5/L2 | 2.2.2/L3 | |
| 2.2.5/L1 | 2.2.6/L2 | 2.2.3/L3 | |
| 2.2.6/L1 | | 2.2.4/L3 | |
| | | 2.2.6/L3 | |

Es decir, un total de 13 sobre 24 casos analizados, prácticamente el 50 %. Hay que destacar además que la elección de uno u otro patrón no parece forzada por alguno de los esquemas propuestos, sino que parece quedar totalmente en manos del locutor (no hay ninguna frase que presente el mismo patrón en los 4 locutores).

El valor final de F_0 en el segmento descendente varía desde - 7 Hz con respecto al valor de la F_0 en el inicio (2.2.6/L3), hasta los - 74 Hz (2.2.5/L2). La media de los valores obtenidos es - 42 Hz, algo menor que en las enunciativas, aunque hay bastante dispersión en los resultados obtenidos. Como puede comprobarse, a diferencia de las interrogativas absolutas con final descendente, el valor de la F_0 al final de la curva se sitúa por debajo del 0, correspondiente al valor de la F_0 al inicio. En cuanto a la pendiente, oscila desde los - 1,19 hasta los - 0,161 Hz/ms, y su valor medio fue - 0,52 Hz/ms., aunque

los valores obtenidos no están demasiado agrupados, según muestra el histograma.

Dentro de este grupo se observan diferencias, sin embargo, en cuanto al punto en que se inicia el segmento final descendente. En función de esto, podrían establecerse dos subgrupos diferentes:

a) Uno, formado por aquellas frases en las que el segmento final se inicia en el pico de la primera sílaba acentuada. A este grupo pertenecerían las frases 2.2.3/L1, 2.2.4/L1, 2.2.5/L1, 2.2.6/L1, 2.2.5/L2 y 2.2.6/L2.

b) Otro, en el que se incluirían las frases restantes, y que se caracterizaría por la aparición, en las proximidades del último (o penúltimo; caso de 2.2.1/L2) acento, de una segunda subida, además de la principal, a partir de la cual comenzaría el segmento final descendente.

2) Esquema B:

Es el esquema propio de las interrogativas absolutas, descrito en el apartado anterior, y caracterizado por un **segmento final marcadamente ascendente**. Al igual que en las absolutas, este segmento final presenta en las frases analizadas un inicio bastante independiente de la posición del acento (figura 4.15): lo encontramos situado entre el 69 % (2.2.2/L2) y el 76 % (2.2.4/L2) de la longitud total de la frase, cuando la frase tiene más de un acento (valores prácticamente iguales a los de las interrogativas absolutas), y entre el 85,4 % (2.2.5/L3) y el 97 % (2.2.6/L5), en las que sólo tienen una sílaba tónica (aquí las cifras difieren bastante con respecto a las interrogativas absolutas). También se asemejan a las absolutas en la altura tonal que este punto suele alcanzar, un poco por debajo de la Fo inicial.

También como en las interrogativas absolutas, el valor de la Fo al final de este segmento es extremadamente variable: desde los - 8 hasta los + 261 Hz con respecto al valor de la Fo al principio de la curva, con un valor medio de + 146 Hz, algo inferior al de las interrogativas absolutas. El análisis del histograma correspondiente confirma esta idea.

La pendiente oscila entre los 0,52 y los 2,60 Hz/ms. (1,06 Hz/ms. de media), aunque éste último se trata de un valor aislado, alejado del resto. Al igual que en el parámetro anterior, el histograma muestra una gran dispersión de los valores.

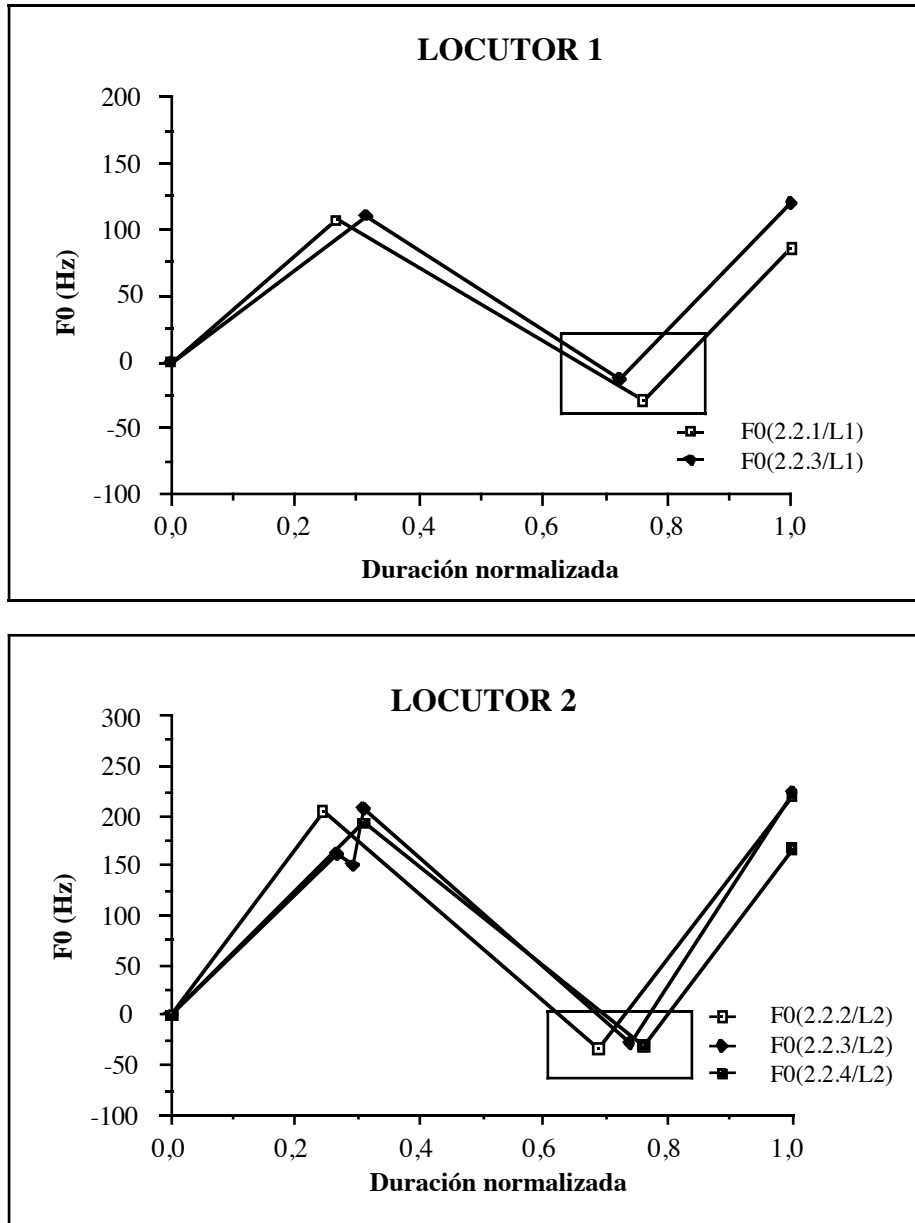


Figura 4.15.- En la presente figura se muestra la posición del punto de inicio del segmento final en las interrogativas pronominales. Al igual que en las absolutas, esta posición es bastante semejante en las distintas frases y se localiza aproximadamente al comienzo de la última sílaba, sea ésta tónica o átona.

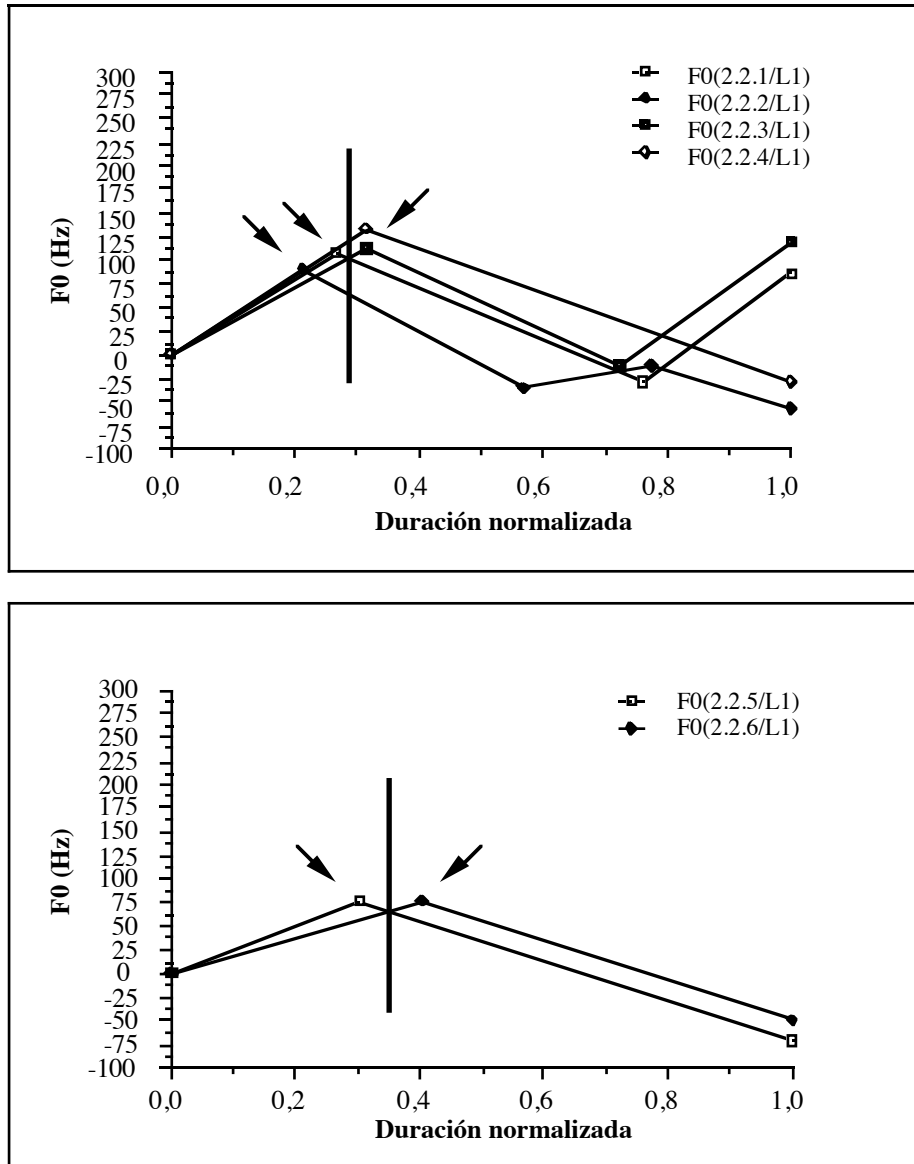


Figura 4. 16.- La figura ilustra las diferencias existentes en la posición del primer pico según la posición de la primera sílaba acentuada en las frases interrogativas pronominales. En los dos gráficos puede observarse cómo, tanto en los esquemas ascendentes como en los descendentes, aquellas frases en las que la primera sílaba es tónica presentan el primer pico más cerca del inicio de la frase que aquellas con el acento en la segunda sílaba.

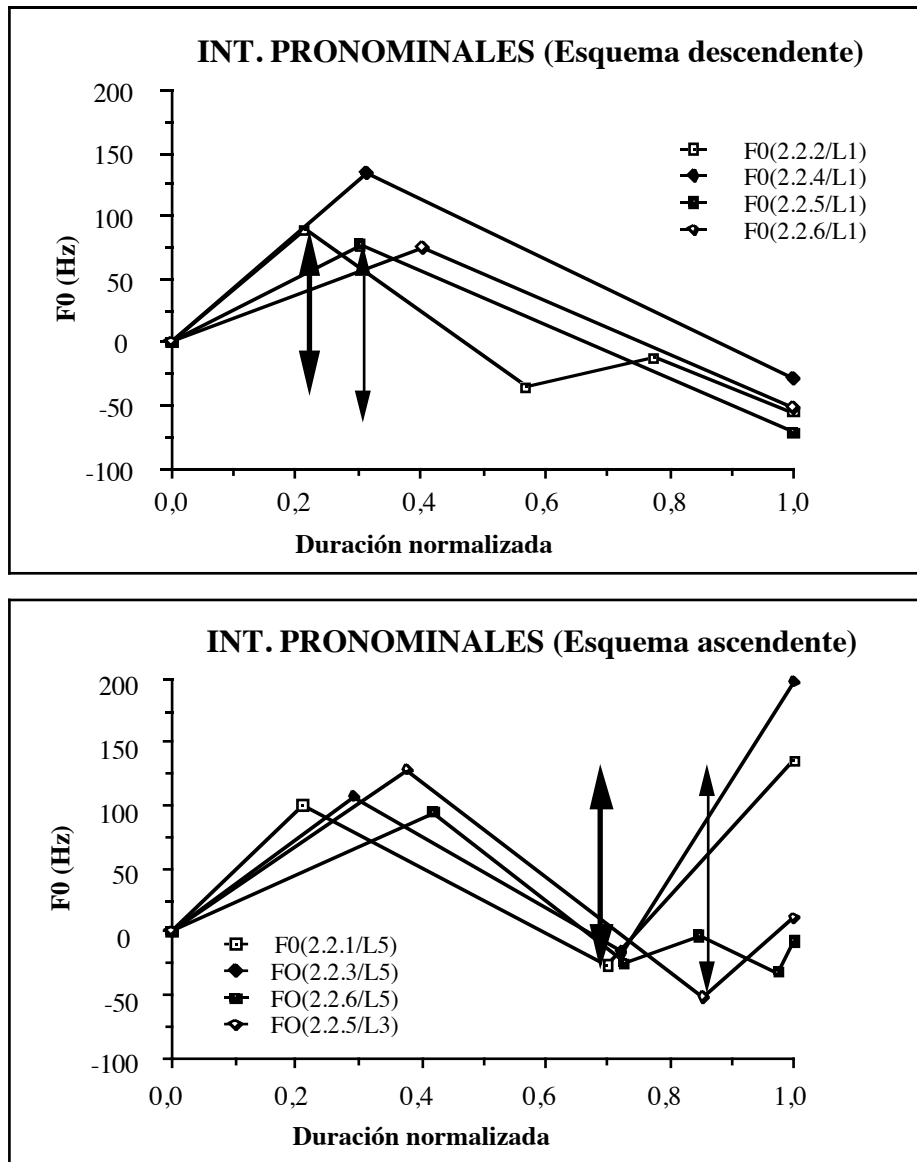


Figura 4.17.- La figura ilustra las diferencias existentes en el rango frecuencial, para las frases interrogativas pronominales, entre las frases largas y las frases cortas. Como en figuras anteriores, la línea más gruesa indica el rango en las frases largas (2.2.2/L1, 2.2.3/L1, 2.2.1/L5 en los cuadros), y la más delgada el rango en las frases cortas (2.2.5/L1, 2.2.6/L1, 2.2.6/L5 y 2.2.5/L3). Se observa que en ambos cuadros las frases cortas presentan un mayor rango que las largas.

Aspecto común a ambos esquemas parece ser la aparición de la subida inicial, cerca del primer acento de la frase, según se puede apreciar en la figura 4.16. La altura de esta subida, como ya se ha señalado para grupos anteriores, varía considerablemente en los casos estudiados: se registran valores desde + 43 hasta + 204 Hz normalizados en las frases con final ascendente (valor medio + 123 Hz), y desde + 12 hasta + 164 Hz normalizados (valor medio + 80 Hz) para las frases con final descendente. Es decir, que los valores parecen más altos en las frases con final ascendente, pero en conjunto son semejantes a los de las interrogativas absolutas. En cuanto a la pendiente de la subida, el valor medio de ésta es de 0,57 Hz/ms. en las frases con final ascendente, un valor algo inferior al de las interrogativas absolutas, pero superior al de las enunciativas, con una distribución de los valores también bastante agrupada. En las frases con final descendente, se ha obtenido un valor medio de 0,75 Hz/ms., sensiblemente superior al del otro grupo y al de las interrogativas absolutas. Los valores, de acuerdo con el histograma, se presentan bastante agrupados. Estos resultados parecen indicar que el primer pico en las frases con final descendente es menos alto, pero alcanza su máxima altura más rápidamente que en las frases con final ascendente.

El rango medio en el caso de las frases con final descendente es 135 Hz, próximo al de las enunciativas; en las frases con final ascendente es 203 Hz, un valor bastante semejante al de las interrogativas absolutas. En ambos casos, sin embargo, los valores no están demasiado agrupados. Tampoco se observan diferencias significativas según la longitud de la frase. Como en las absolutas, los valores en las frases cortas son superiores a los de las frases largas, tanto en los esquemas ascendentes como en los descendentes (*vid.* figura 4.17).

4.4.4. Frases interrogativas relativas (2.3.1 - 2.3.12)

Las representaciones obtenidas en el estudio de este tipo de frases son, en su mayoría, prácticamente idénticas a las de las absolutas: primer pico en las proximidades del primer acento de la frase - *vid.* figura 4.18 -, arranque del segmento final en un punto muy semejante al de las interrogativas absolutas (figura 4.19), y **final ascendente** en todos los casos estudiados (48), menos en uno (2.3.6/L3).

La pendiente media de la subida del primer pico es similar a la de los otros tipos de interrogativas: 0,7 Hz/ms. La distribución de los valores obtenidos está muy próxima igualmente a la normal, lo que hace pensar que la pendiente de la subida inicial puede ser un parámetro relevante de cara a la identificación de las interrogativas de todo tipo. En cuanto a la altura que alcanza la Fo en este pico, los valores registrados oscilan entre los 0 y los + 248 Hz normalizados, siendo + 110 Hz el valor medio, un valor semejante al de las interrogativas ya comentadas, y superior al de las enunciativas. La distribución de los valores es bastante normal, según se aprecia en el histograma.

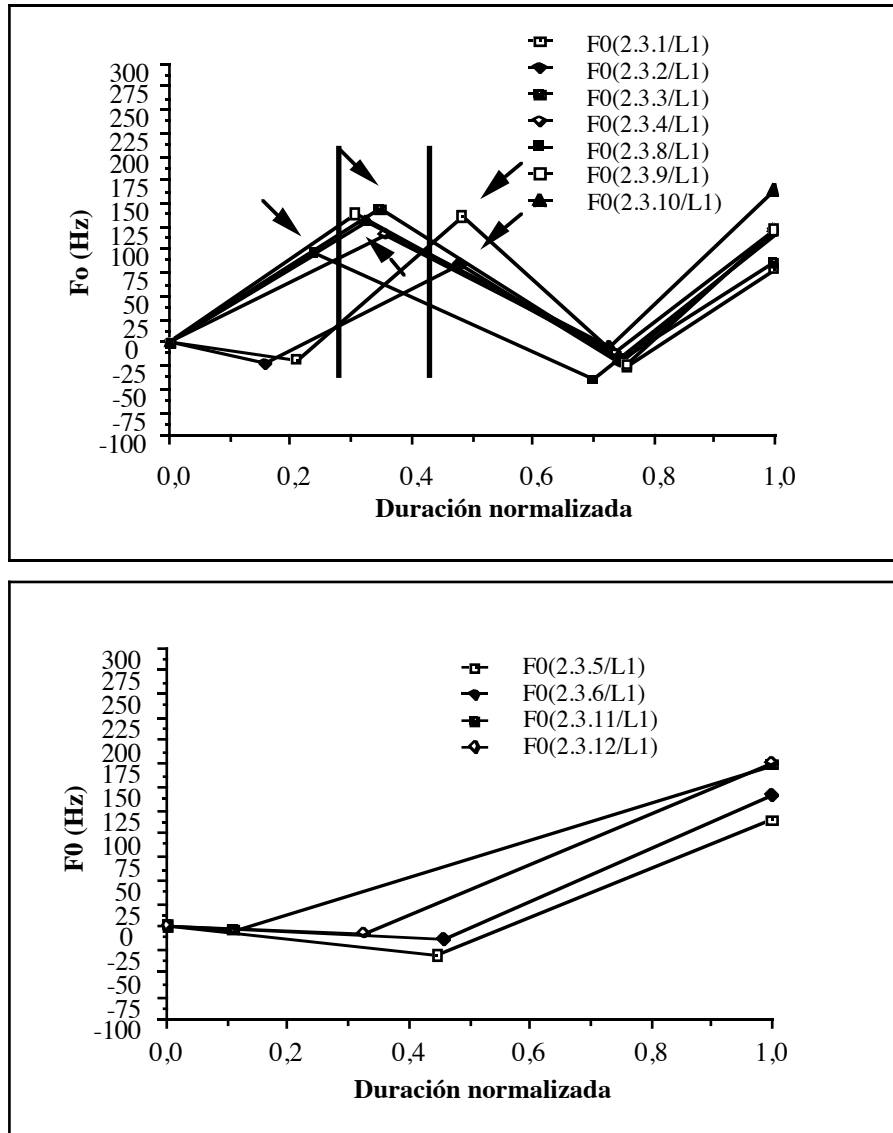


Figura 4.18.- La figura ilustra las diferencias existentes en la posición del primer pico según la posición de la primera sílaba acentuada en las frases interrogativas relativas. En el primer gráfico puede observarse cómo, en las frases largas, se cumple en general (aunque aquí con alguna excepción) la norma sobre la posición del primer pico enunciada para tipos anteriores. En el cuadro inferior se aprecia cómo, al igual que las interrogativas absolutas, las frases interrogativas cortas no presentan un pico inicial.

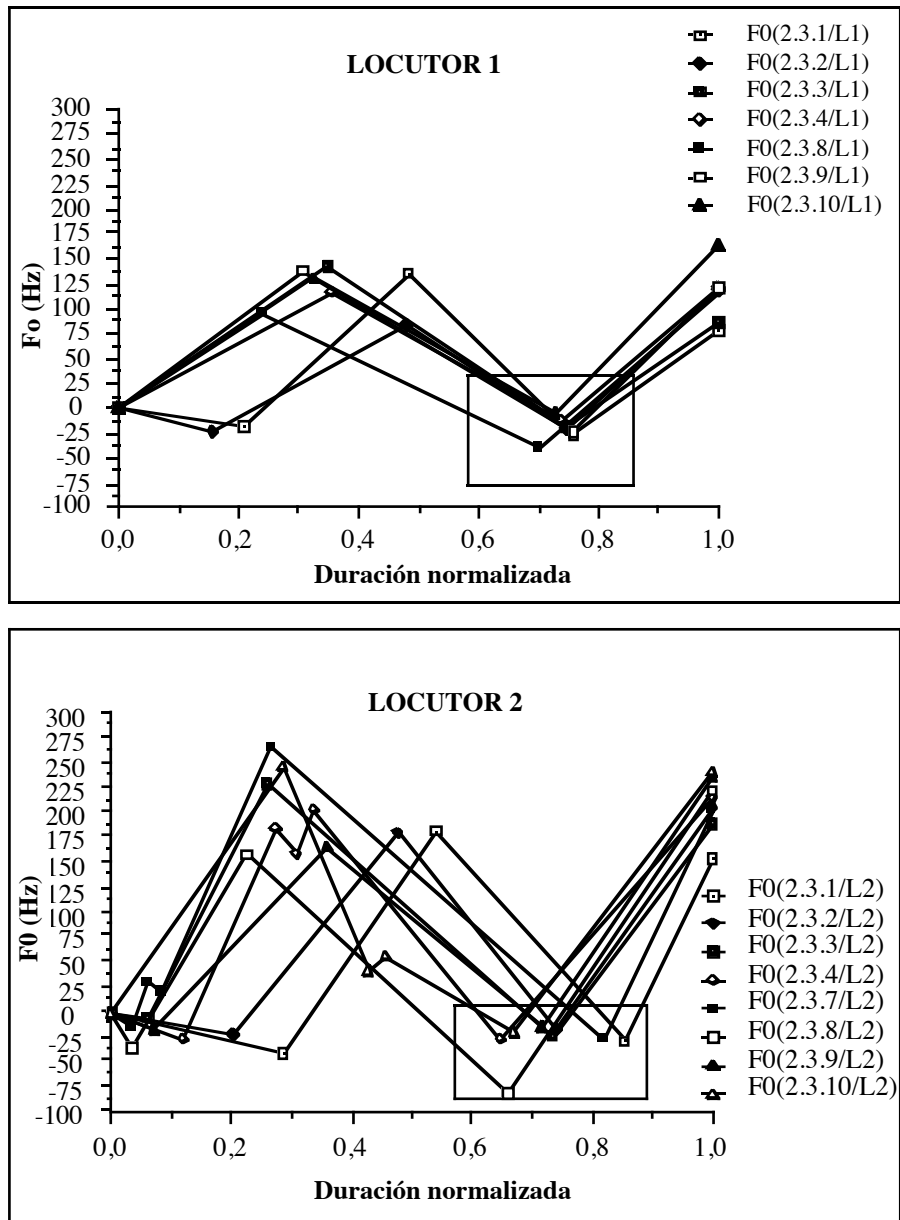


Figura 4.19.- En la presente figura se muestra la posición del punto de inicio del segmento final en las interrogativas relativas. Al igual que en absolutas y pronominales, esta posición es bastante semejante en las distintas frases y se localiza aproximadamente al comienzo de la última sílaba, sea ésta tónica o átona.

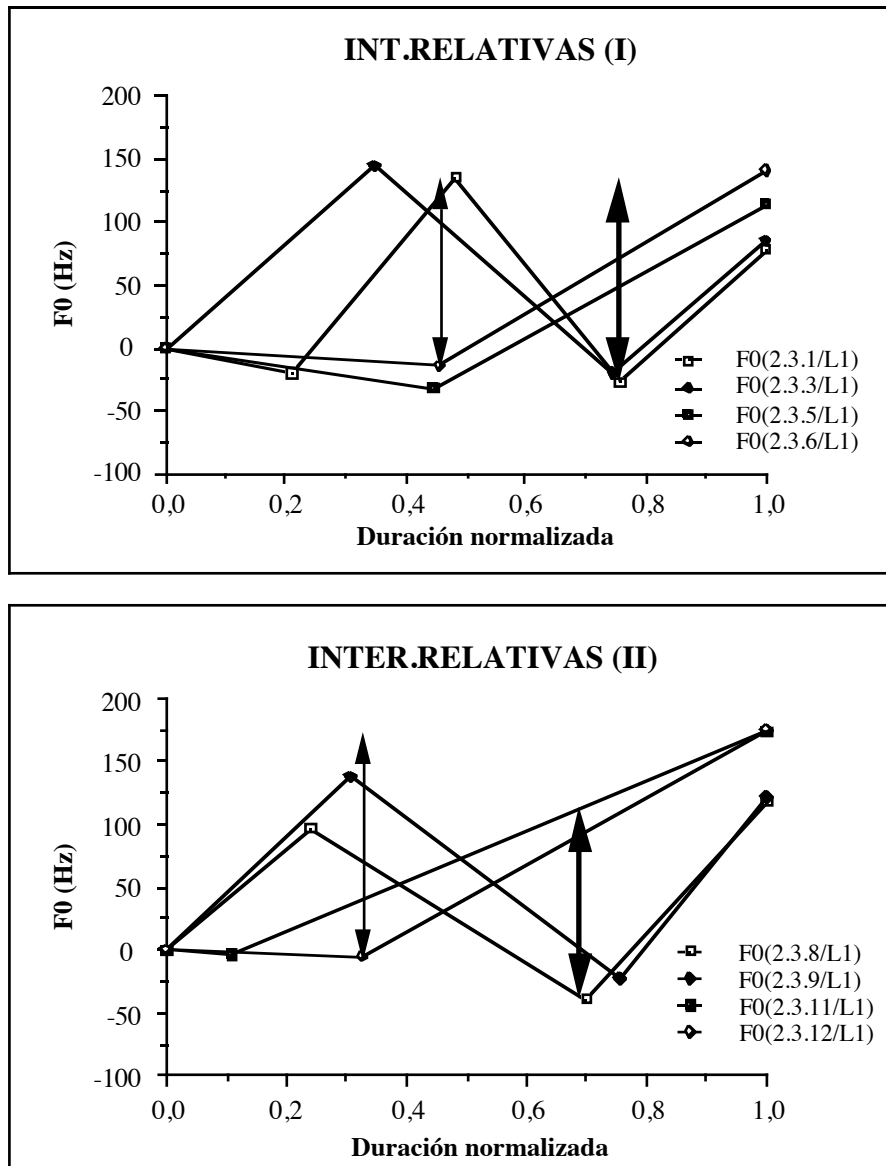


Figura 4.20.- La figura ilustra las diferencias existentes en el rango frecuencial, para las frases interrogativas relativas, entre las frases largas y las frases cortas. Como en figuras anteriores, la línea más gruesa indica el rango en las frases largas (2.3.1/L1, 2.3.3/L1, 2.3.8/L1, 2.3.8/L1 en los cuadros), y la más delgada el rango en las frases cortas (2.3.5/L1, 2.3.6/L1, 2.3.11/L1 y 2.3.12/L1). Se observa que en ambos cuadros las frases cortas presentan un igual o mayor rango que las largas.

Esta semejanza del primer pico en los tres tipos de interrogativas, tanto por lo que hace referencia a la pendiente como a la altura que alcanza la F_0 , lleva a pensar que este primer pico sería un elemento identificador, en términos generales, de las frases interrogativas. Además, serviría para distinguirlas de las enunciativas, que presentan generalmente valores inferiores para ambos parámetros. El solapamiento entre los valores de ambos grupos, sin embargo, podría llevar a errores de identificación.

El punto de inicio del segmento final se sitúa entre unos márgenes muy semejantes a los de las absolutas, como ya se ha avanzado: correspondería, al igual que en las absolutas y pronominales, al inicio de la última sílaba de la frase. Pueden encontrarse ejemplos de este fenómeno en la figura 4.19.

El segmento final, siempre ascendente, presenta por término medio una pendiente de 0,79 Hz/ms., algo menor que en el caso de las interrogativas absolutas. Su distribución se aproxima mucho a la normal, con valores entre 0 y 1,82 Hz/ms. El valor de la F_0 al final de este segmento varía bastante, entre los 9 y los 332 Hz normalizados. El valor medio obtenido es 157 Hz, semejante al de las otras interrogativas. También en este parámetro la distribución se acerca bastante a la normal.

Las interrogativas relativas presentan un rango medio muy semejante al de las absolutas, 200 Hz. Los valores oscilan entre los 97 y los 343 Hz. Con una distribución próxima a la normal, el histograma muestra la ausencia de diferencias entre frases de distinta longitud. Incluso, al igual que ocurre en las otras interrogativas, las frases más cortas llegan a alcanzar un rango mayor que las largas, tal como puede comprobarse en las representaciones de la figura 4.20.

4.4.5. Frases exclamativas (3.1-3.12)

En el *corpus* se distinguieron dos tipos diferentes de exclamativas, según la connotación expresiva que contuviesen. Así, las frases 3.1 a 3.6 estaban incluidas en contextos que implicaban **admiración**, en tanto que las frases 3.7 a 3.12 estaban en contextos con matiz implícito de **desaprobación**. Sin embargo, tras el análisis, no se han apreciado diferencias sistemáticas entre un grupo y otro, por lo cual los resultados correspondientes a ambos grupos se exponen aquí conjuntamente.

Destaca ante todo la presencia de dos tipos de segmentos finales, uno ascendente y otro descendente, con predominio del segundo sobre el primero (36 casos frente a 9). A continuación se especifican las frases que presentan una y otra clase de final:

a) Segmento descendente:

| | | | |
|--------|--------|---------|---------|
| 3.1/L1 | 3.3/L2 | 3.7/L1 | 3.10/L2 |
| 3.2/L1 | 3.4/L2 | 3.8/L1 | 3.12/L2 |
| 3.4/L1 | 3.6/L2 | 3.9/L1 | |
| 3.5/L1 | | 3.10/L1 | |
| 3.6/L1 | | 3.11/L1 | |
| 3.1/L2 | | 3.12/L1 | |
| 3.2/L2 | | | |
| 3.1/L3 | 3.1/L5 | 3.8/L3 | 3.10/L5 |
| 3.2/L3 | 3.2/L5 | 3.9/L3 | 3.11/L5 |
| 3.3/L3 | 3.4/L5 | 3.10/L3 | 3.7/L5 |
| 3.4/L3 | 3.6/L5 | | 3.8/L5 |
| 3.5/L3 | | | |

b) Segmento ascendente:

| | | | |
|--|--------|---------|---------|
| | | | 3.9/L2 |
| | | | 3.11/L2 |
| | 3.3/L5 | 3.7/L3 | 3.9/L5 |
| | 3.5/L5 | 3.11/L3 | 3.12/L5 |
| | | 3.12/L3 | |

Como se desprende de esta relación, si bien el final descendente aparece en ambas clases de exclamativas en un número semejante de ocasiones, el final ascendente se encuentra, en la mayor parte de los casos, en frases del segundo grupo, las que implican un matiz de desaprobación.

En ambos grupos se observa la subida inicial en las proximidades de la primera sílaba tónica - figura 4.21. La altura que alcanza esta primera subida varía de un locutor a otro, pero es en general mayor que en las enunciativas. El valor mínimo se sitúa, para las frases con final ascendente, en 14 Hz por encima de la F_0 inicial, y el máximo, 254 Hz por encima del valor inicial. En las que presentan final descendente, se registran valores entre los 5 y los 210 Hz normalizados. Los valores medios en cada caso son, respectivamente, 119 y 93 Hz, más altos que en las enunciativas y semejantes a los de las interrogativas.

La pendiente en esta primera subida es, por término medio, de 0,72 Hz/ms. en las curvas con final ascendente, y de 0,51 Hz/ms. en las que presentan un final descendente. Los valores son algo inferiores a los de las interrogativas, pero superiores a los de las enunciativas.

Los resultados del análisis del primer pico revelan, por tanto, un mayor parecido con el de las interrogativas que con el de las enunciativas, aunque, como ya se ha indicado, los valores de los diferentes grupos se solapan.

A la primera subida le siguen otros picos cuando la frase contiene más de una sílaba tónica. El número de estos picos es por regla general mayor que en las frases enunciativas correspondientes, tal como puede apreciarse en la figura 4.22, y suele coincidir con el resto de sílabas tónicas de la frase. En las enunciativas, sólo una de las sílabas acentuadas restantes suele presentar una subida de F_0 , de manera que mientras que en las enunciativas se observan normalmente dos picos por frase, en las exclamativas lo más usual son tres.

Este hecho puede relacionarse con el esquema "ondulado", que, como ya se ha señalado, propone Navarro (1948) como una de las formas de expresividad en español.

El segmento final, si es descendente, comienza habitualmente en la última sílaba acentuada, y al inicio de la última sílaba, si es ascendente. En el caso de las descendentes, la pendiente media es de $-0,49$ Hz/ms., y su altura media al final del mismo, -35 Hz normalizados. Ambos valores son más bajos que los correspondientes a las enunciativas, lo que indica un segmento final con un descenso más acusado en las exclamativas que en las interrogativas. En las ascendentes, el valor medio de la pendiente es $0,48$ Hz/ms., y el de la altura de la F_0 , $+13$ Hz normalizados, ambos significativamente más bajos que en las interrogativas con final ascendente, lo que indica un segmento ascendente menos pronunciado que en éstas. Los valores de las pendientes, sin embargo, presentan una cierta dispersión en los dos casos. En cambio, los valores correspondientes a la altura de la F_0 están un poco más agrupados.

En algunos de los casos estudiados, como se puede observar en la figura 4.23, el último pico alcanza una altura inusual, bastante mayor que en los picos finales de las enunciativas, que incluso supera la altura del primer pico de la frase. Así ocurre en las frases 3.4/L1, 3.4/L3 y 3.6/L1. En estos casos, el pico suele situarse en la última sílaba, con independencia de que ésta sea tónica o no. Este hecho podría relacionarse con el **tonema circunflejo** ya mencionado en el capítulo 1, considerado por Navarro Tomás como una de las formas de expresión de la emotividad en español.

El valor del rango medio en las frases exclamativas es más elevado en las frases con final ascendente (170 Hz) que en las que presentaban un final descendente (145 Hz). En conjunto, son valores algo mayores que en las enunciativas, pero menores que en las interrogativas. No hay apenas variación del rango en función de la duración, según se aprecia en la figura 4.24. Sí hay, en cambio, bastante variación de un locutor a otro, como puede comprobarse en la figura 4.25. De hecho, aunque esta variación es más acusada en las frases exclamativas, puede apreciarse también en los demás tipos de frases: en general, es L2 el locutor que presenta los rangos más elevados, aspecto que se correspondería con una mayor **expresividad** en la pronunciación de las frases. La presencia de dos picos en el histograma de este tipo de frases corrobora esta idea.

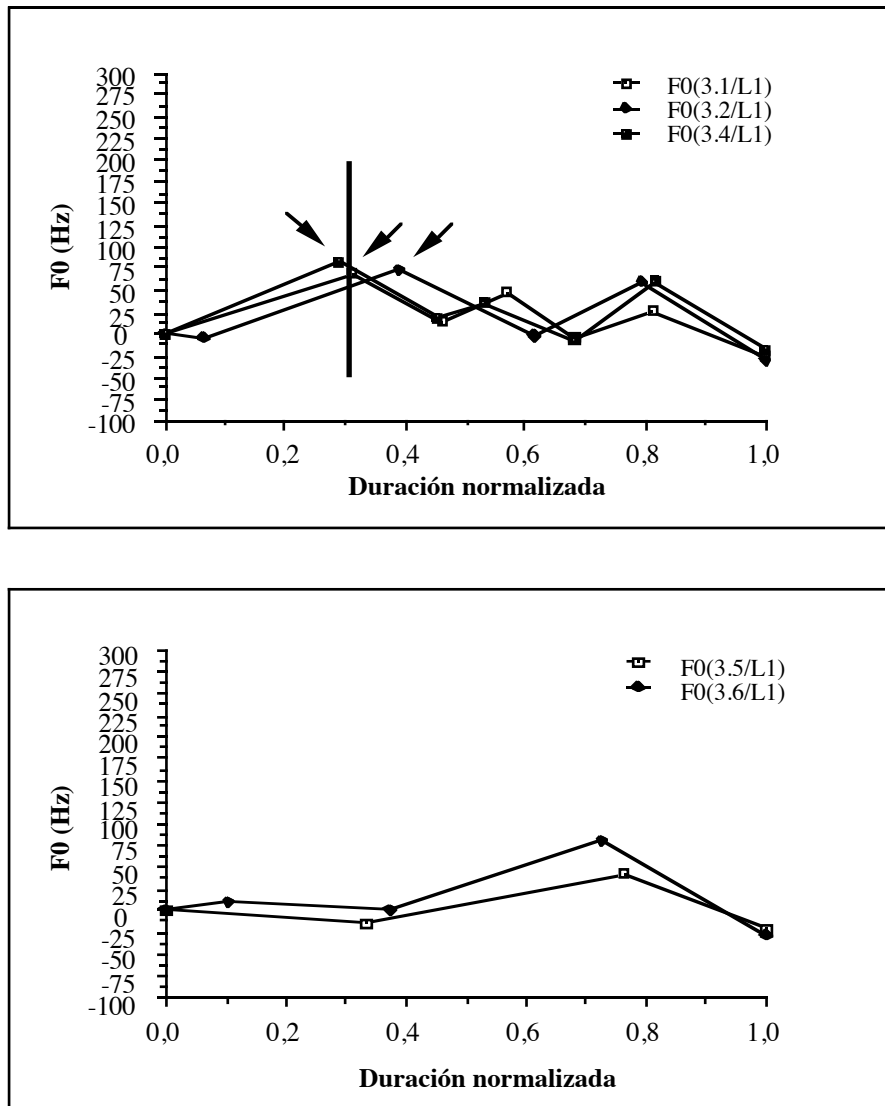
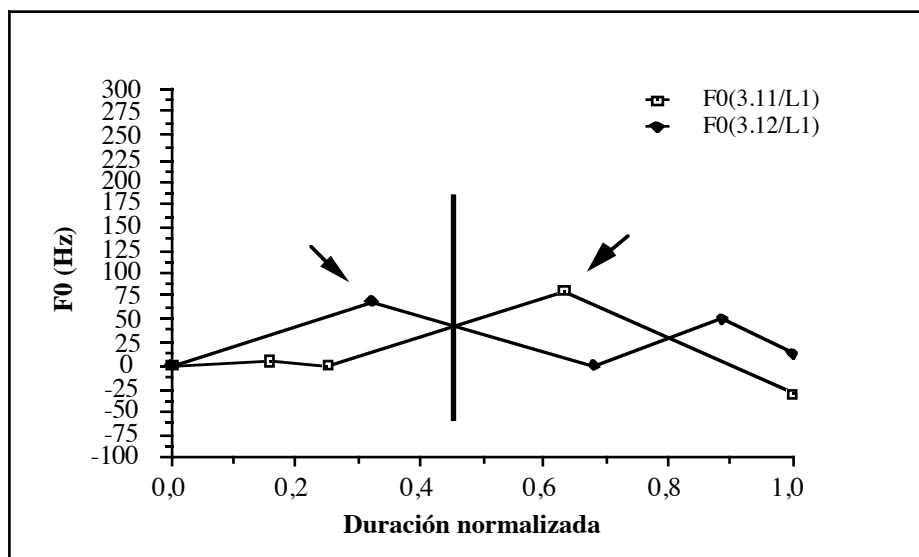
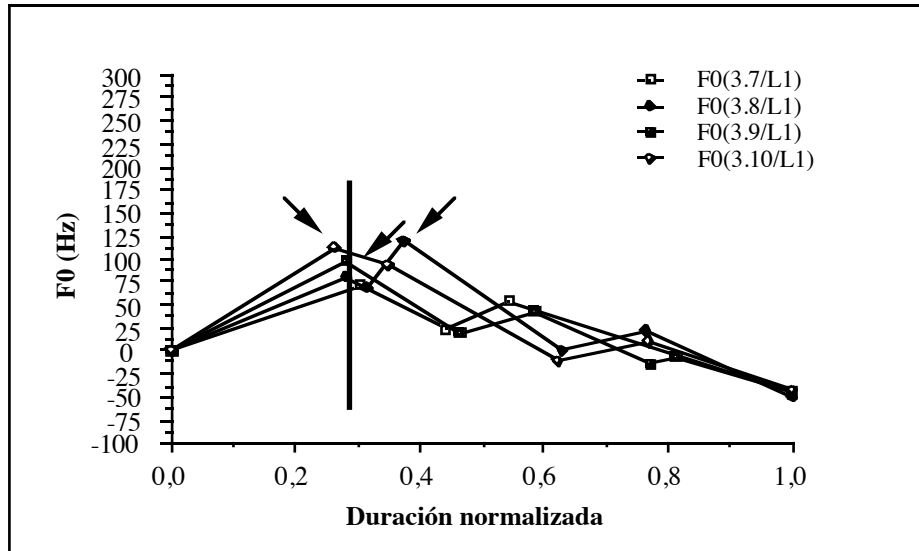


Figura 4.21.- La presente figura ilustra las diferencias existentes en la posición del primer pico según cuál sea la primera sílaba acentuada en las frases exclamativas. En los diferentes gráficos puede observarse cómo, tanto en las exclamativas de admiración como en las de desaprobación, aquellas frases en las que la primera sílaba es tónica presentan el primer pico más cerca del inicio de la frase que aquellas cuya primera sílaba tónica se encuentra en la segunda sílaba.



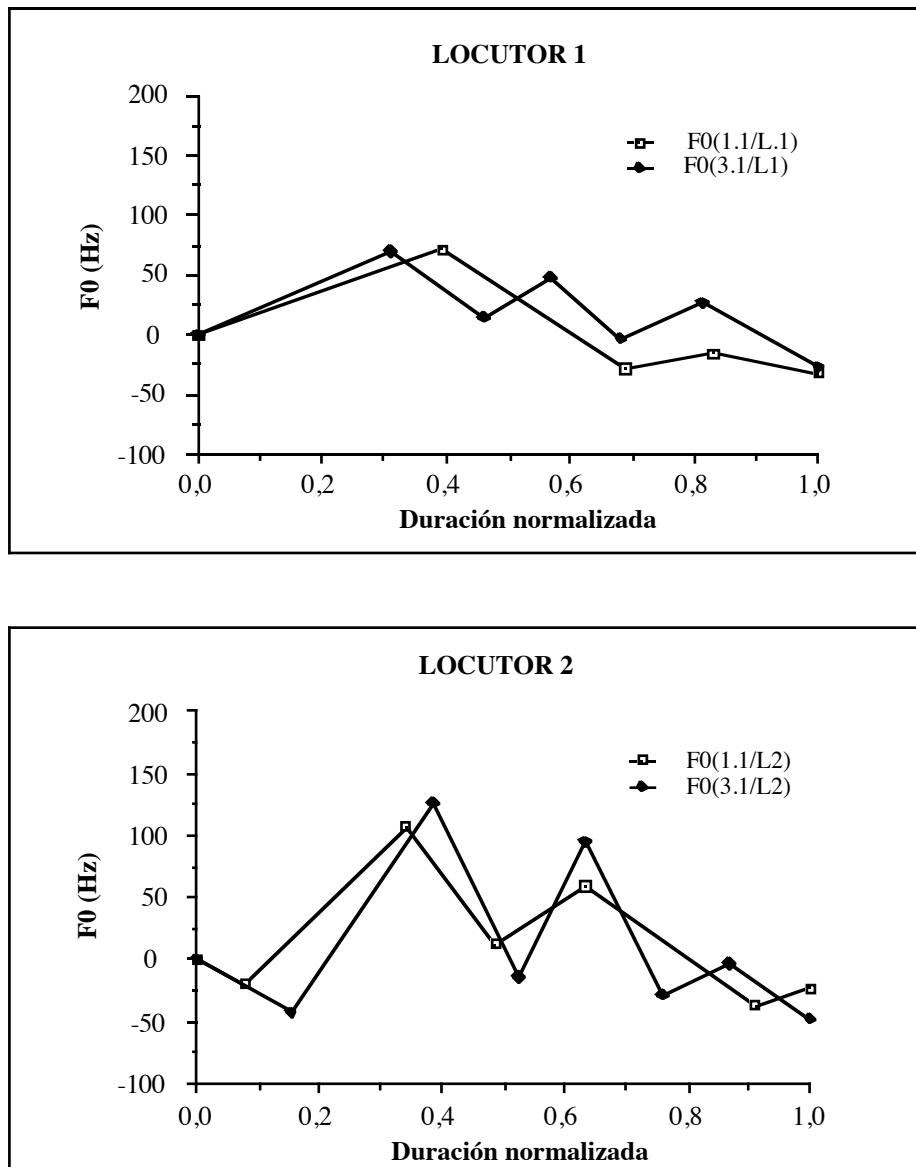
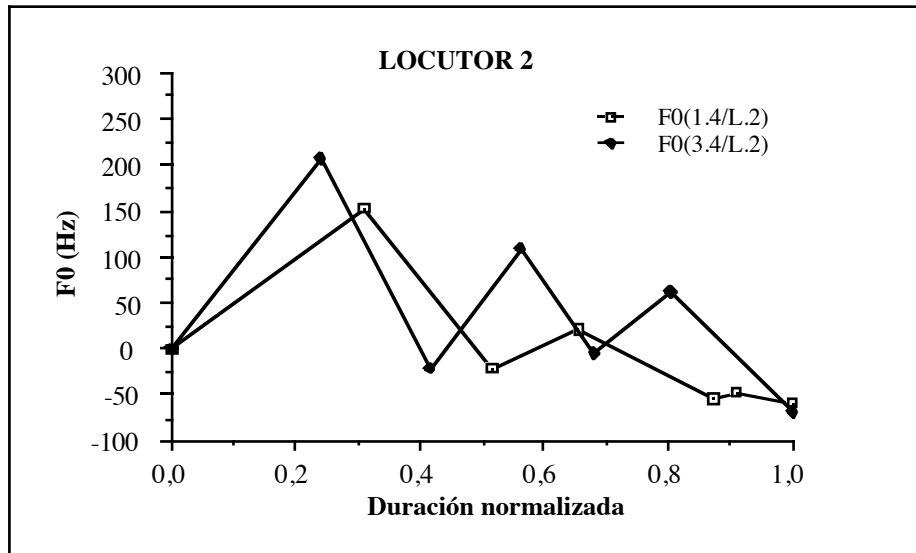
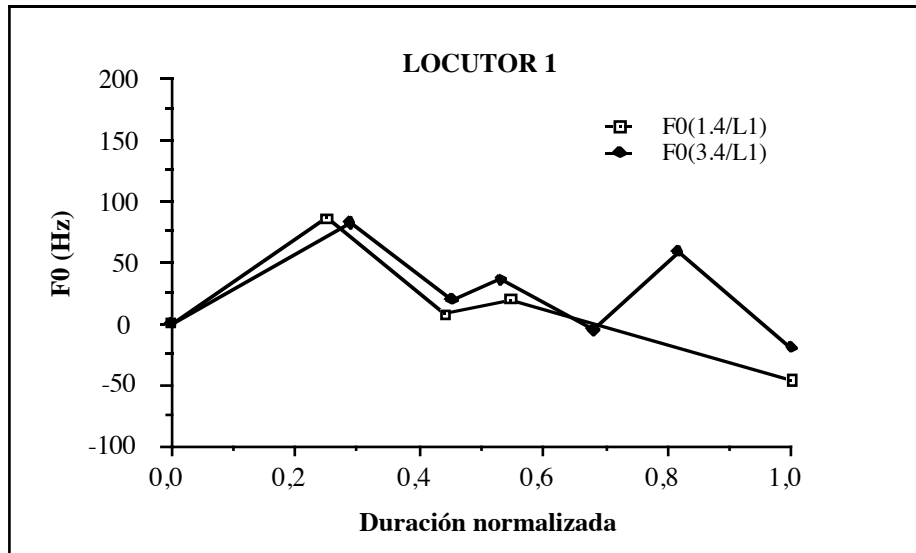


Figura 4.22.- En la presente figura se muestran las diferencias en el número de picos de la curva entre las frases enunciativas y las exclamativas. En cada cuadro se presenta la misma frase pronunciada con entonación enunciativa y con entonación exclamativa.



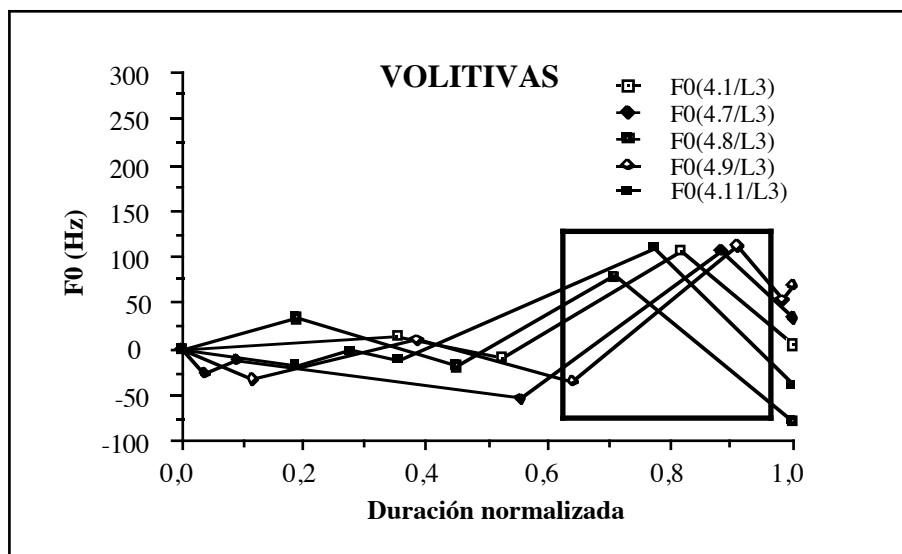
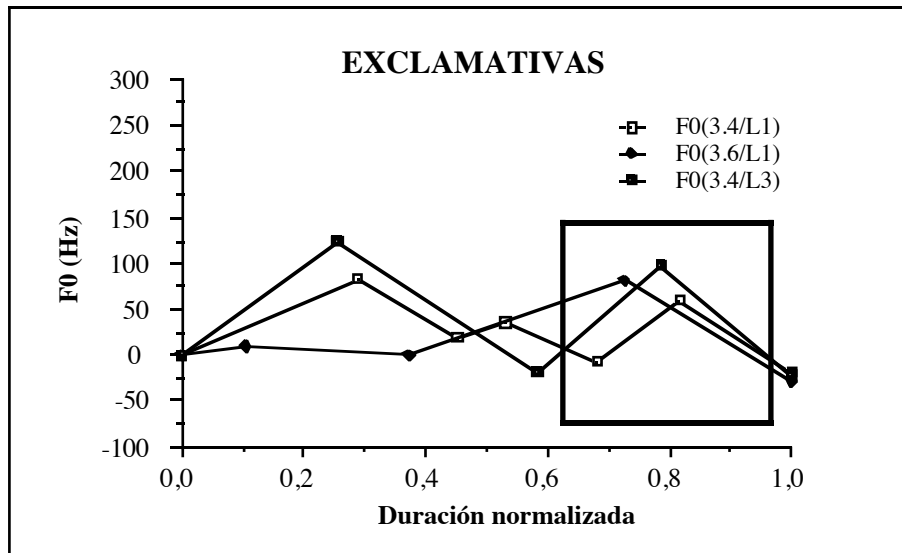


Figura 4.23.- Los cuadros de esta figura muestran los casos de "tonema circunflejo" hallados en las curvas analizadas. Este esquema consistiría, como se observa en estas curvas, en la presencia de un pico de altura elevada al final de las mismas, en general coincidiendo con la última sílaba.

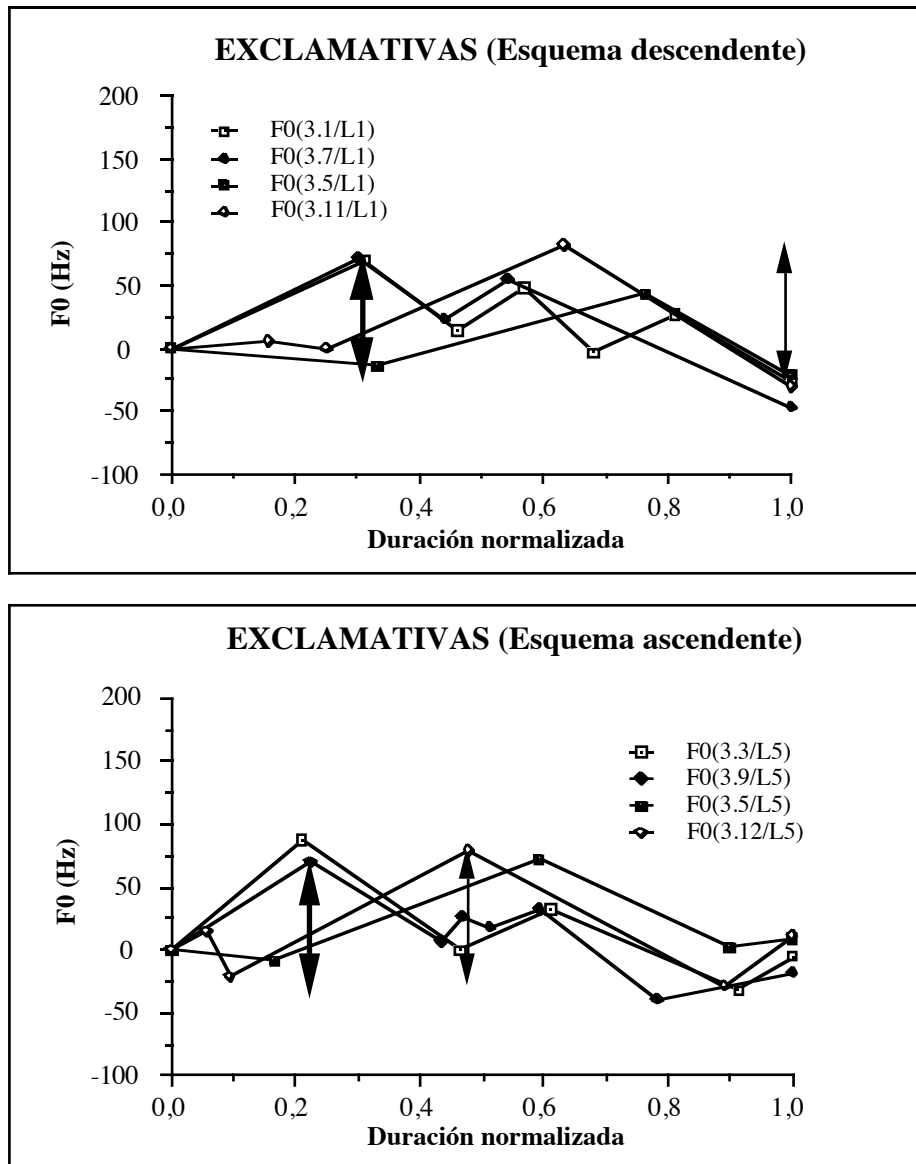


Figura 4.24.- La figura ilustra las diferencias existentes en el rango frecuencial, para las frases exclamativas, entre las frases largas y las frases cortas. Como en figuras anteriores, la línea más gruesa indica el rango en las frases largas (3.1/L1, 3.7/L1, 3.3/L5, 3.9/L5 en los cuadros), y la más delgada el rango en las frases cortas (3.5/L1, 3.11/L1, 3.5/L5 y 3.12/L5). Se observa que parece haber variabilidad en cuanto al rango en función del locutor.

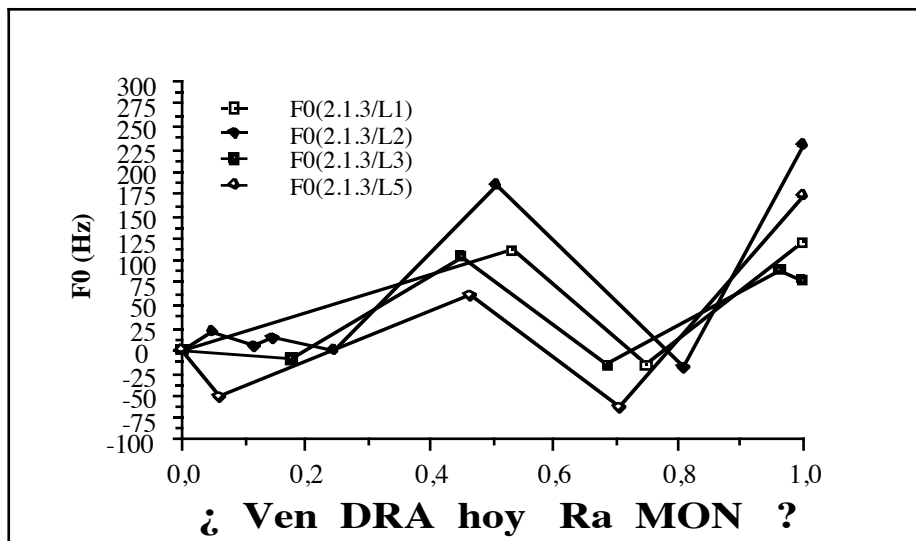
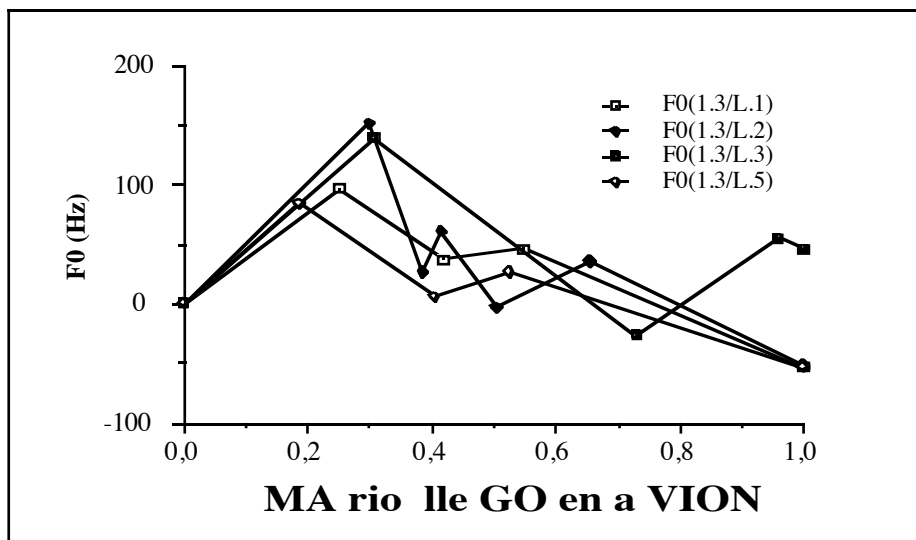
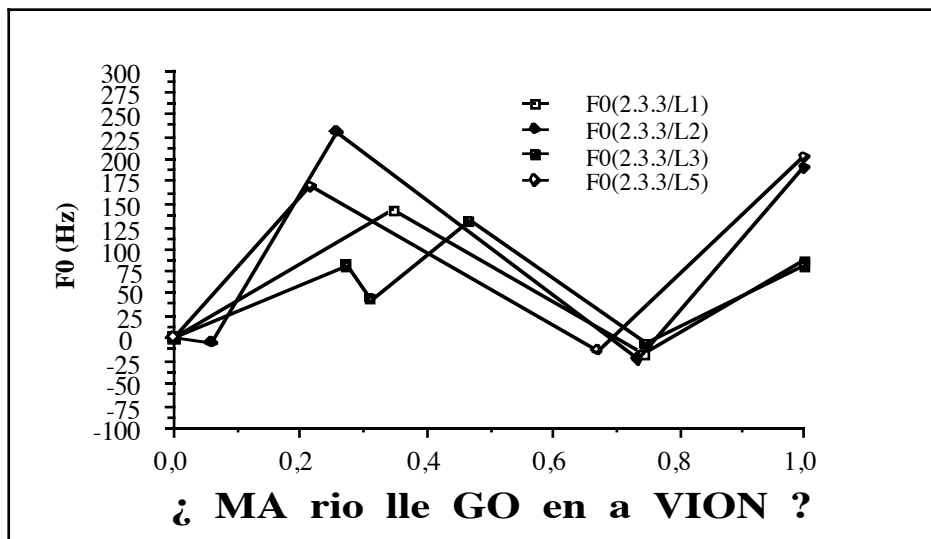
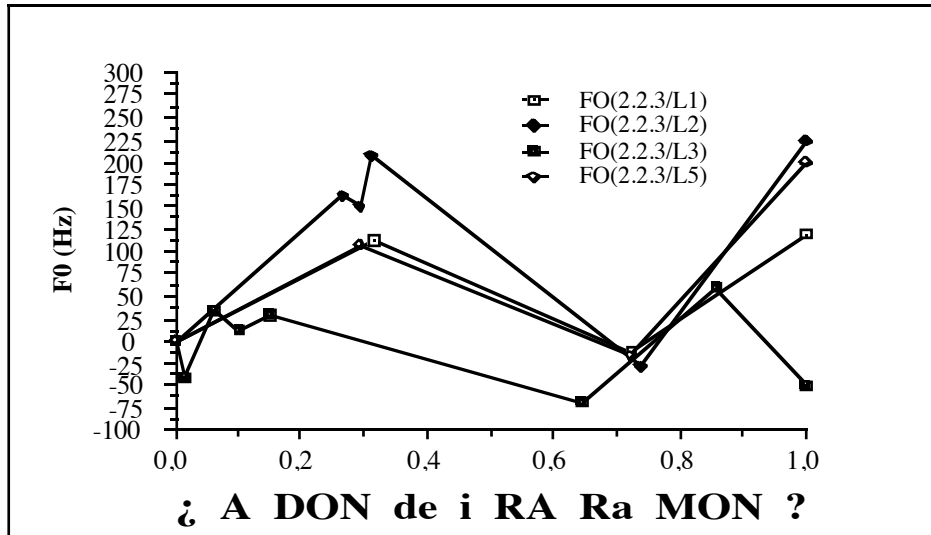
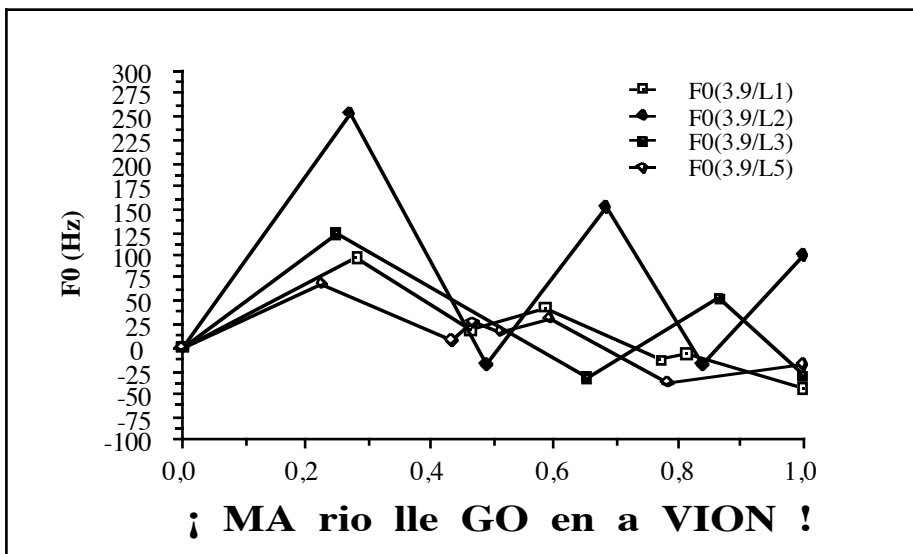
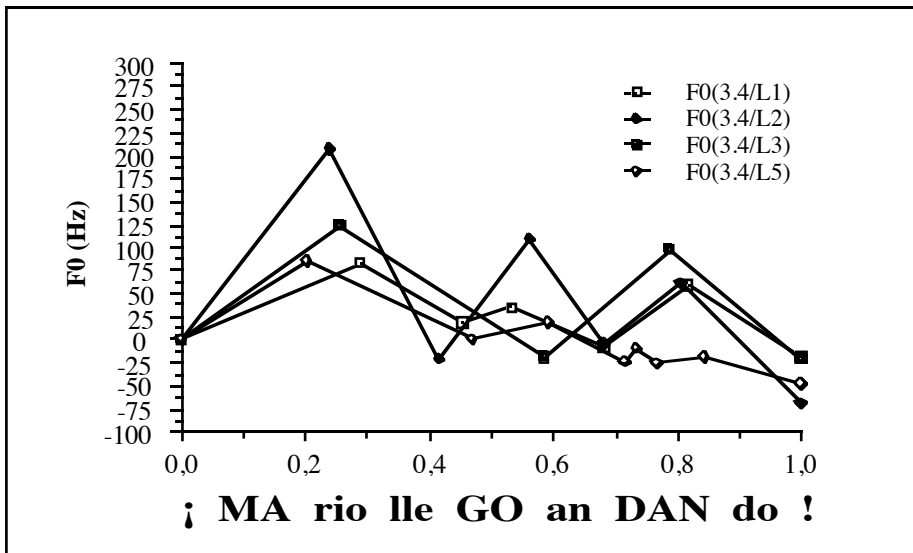
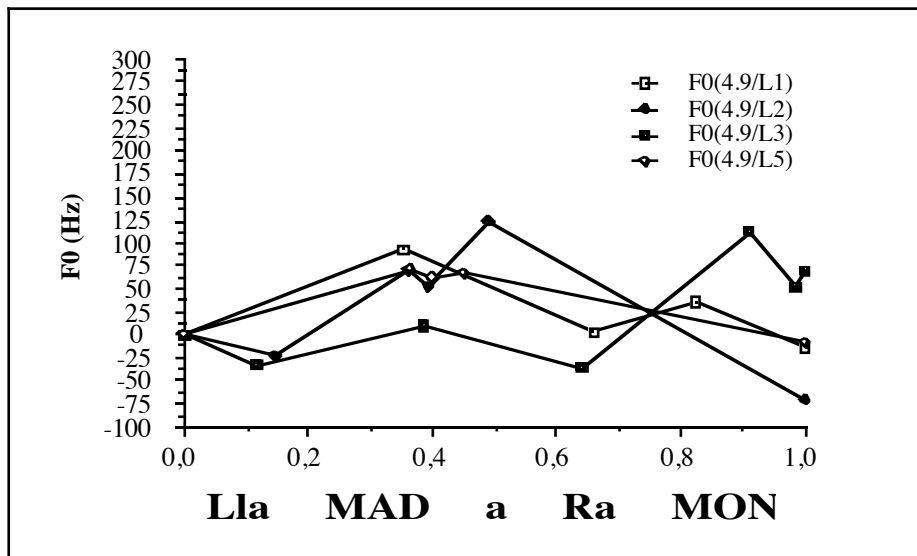
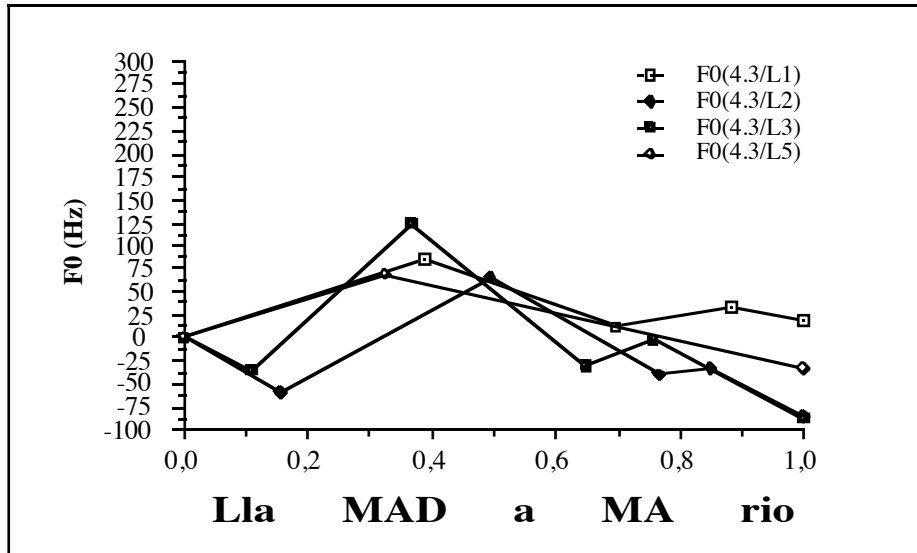


Figura 4.25.- La presente figura ilustra las diferencias existentes en el rango frecuencial, dentro de cada tipo, entre los distintos locutores. En general, el locutor cuyas frases presentan un rango más amplio es L2.







4.4.6. Frases volitivas (4.1-4.12)

Al igual que se ha hecho con las frases exclamativas, se han agrupado los resultados de las volitivas de orden (4.1- 4.6) y los de las de ruego (4.7- 4.12), puesto que su análisis no ha revelado diferencias significativas entre ambos grupos.

En función del segmento final, se distinguen dos tipos de esquemas diferentes, uno con segmento final ascendente y otro con descendente, con predominio del segundo sobre el primero.

Las frases en que se ha detectado un segmento final ascendente son:

| | | | |
|--------|---------|---------|---------|
| 4.5/L1 | 4.12/L1 | | |
| 4.6/L1 | | | |
| 4.5/L3 | 4.6/L5 | 4.9/L3 | 4.7/L5 |
| | 4.12/L3 | 4.10/L5 | |
| | | | 4.11/L5 |
| | | | 4.12/L5 |

El resto, lógicamente, presentó un segmento final con pendiente descendente.

Como puede comprobarse, hay una clara tendencia a preferir el esquema ascendente en las frases cortas (4.5, 4.6, 4.11 y 4.12) y una preferencia de dicho esquema por parte de determinados locutores (el L5, especialmente), en tanto que algún otro (L2, en concreto) no lo emplea en ninguna ocasión.

Al igual que las demás frases analizadas, los casos estudiados de frases volitivas presentan una subida inicial que tiene tendencia a situarse en las proximidades de la primera sílaba acentuada (figura 4.26). La altura que alcanza esta primera subida es semejante a la de las enunciativas: los valores obtenidos se mueven entre los - 13 y los + 124 Hz sobre el nivel tonal inicial, para las frases con final descendente, y entre 0 y los 100 Hz, para las frases con final ascendente (valores medios 59 y 43 Hz, respectivamente). Su pendiente es algo superior en las primeras (0,61 Hz/ms.) que en las segundas (0,36 Hz/ms.), y también semejante en general a la de las enunciativas.

En las frases con más de un acento, aparece normalmente marcado el segundo con otro pico. También se ha detectado en ocasiones (4.1/L3, 4.7/L3, 4.8/L3, 4.9/L3, 4.11/L3) la presencia del esquema circunflejo (*vid.* figura 4.23), como en las frases volitivas, con unas características semejantes a los casos hallados en éstas: altura tonal importante, incluso mayor a veces que el primer pico, y proximidad al final de la curva. Se da, por otra parte, una clara preferencia por parte del locutor 3 en el uso de este esquema.

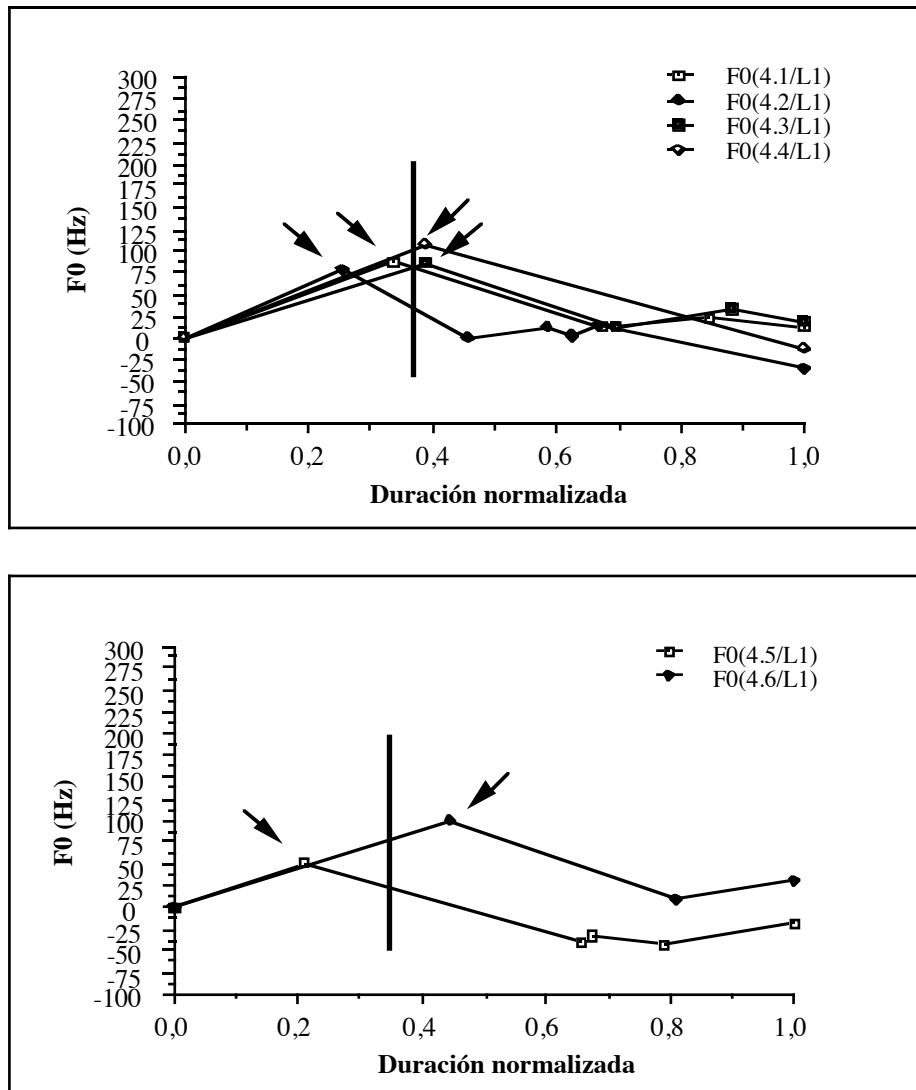
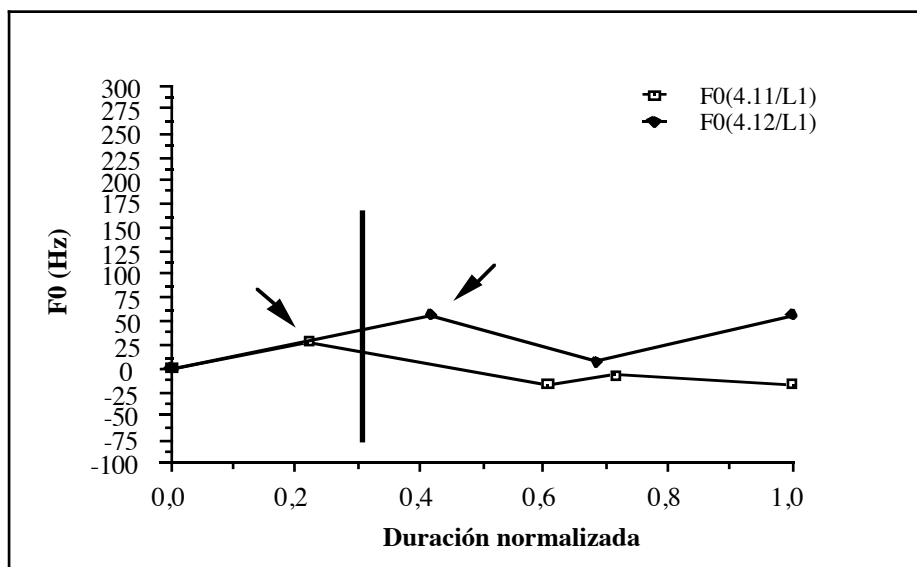
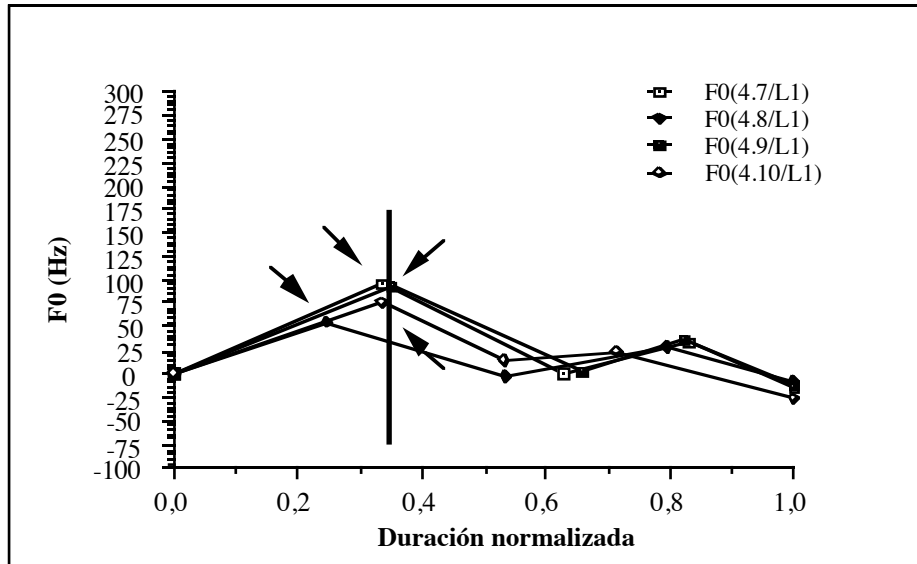


Figura 4.26.- La presente figura ilustra las diferencias existentes en la posición del primer pico según cuál sea la primera sílaba acentuada en las frases volitivas. En los diferentes gráficos puede observarse cómo, tanto en las volitivas de orden como en las de ruego, aquellas frases en las que la primera sílaba es tónica presentan el primer pico más cerca del inicio de la frase que aquellas cuya primera sílaba tónica se encuentra en la segunda sílaba.



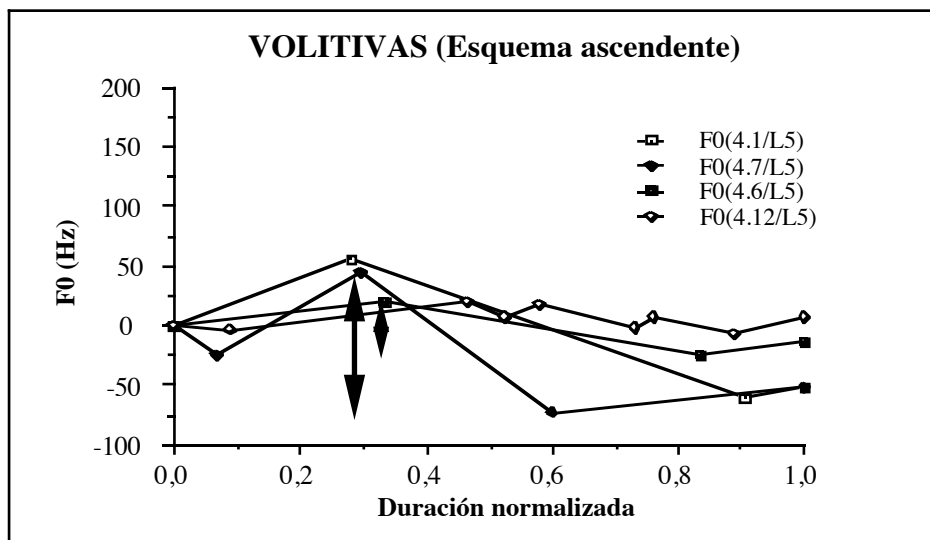
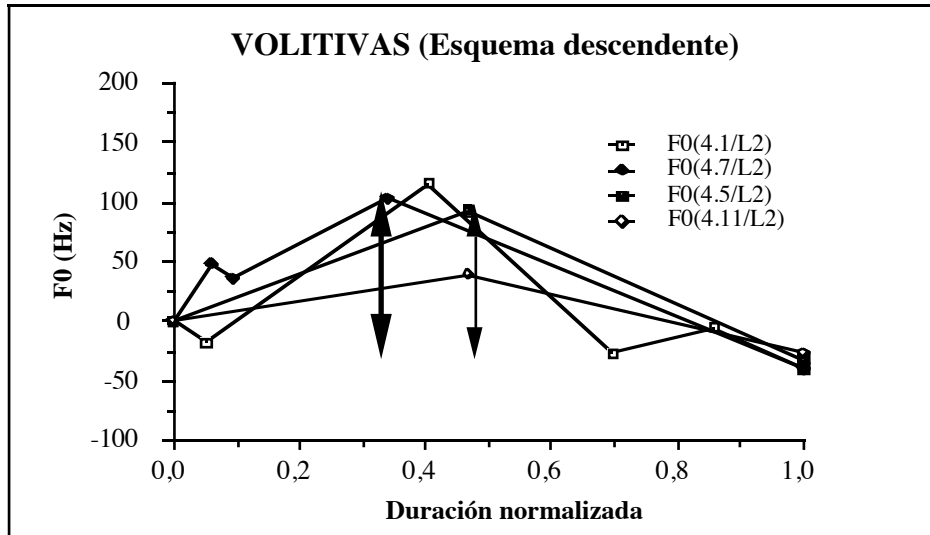


Figura 4.27.- La figura ilustra las diferencias existentes en el rango frecuencial, para las frases volitivas, entre las frases largas y las frases cortas. Las flechas indican lo mismo que en figuras anteriores. Se observan diferencias de rango entre ambos tipos de frases, pero no parecen muy sistemáticas.

La pendiente de los finales descendentes es bastante semejante en la mayoría de las ocasiones (valores sorprendentemente agrupados: del total de casos analizados, 13 se encuentran entre los - 0,36 y los - 0,22 Hz/ms.), con un promedio de - 0,42 Hz/ms.

Los finales ascendentes también presentan una pendiente bastante semejante, con una distribución de los valores próxima a la normal. El valor medio es en este caso de 0,40 Hz/ms.

Tanto los segmentos ascendentes como los descendentes presentan una pendiente media significativamente próxima a las de las exclamativas, lo cual, unido a la ya señalada semejanza de la pendiente del pico inicial, y a la aparición del tonema circunflejo, hace pensar en una relación bastante estrecha entre ambos grupos.

Por lo que hace referencia al valor de la F_0 al final de la curva, los resultados obtenidos son los siguientes: en las frases con final descendente, el valor medio es - 38 Hz normalizados, con valores que oscilan entre los - 94 y los + 34 Hz; en las frases con final ascendente, la media se sitúa en los 15 Hz normalizados, con valores entre los - 51 y los + 132 Hz. En ambos casos son valores muy próximos a los de las exclamativas, lo que parece confirmar la idea de una relación entre ambos grupos.

El rango frecuencial en este tipo de frases se muestra más elevado en las volitivas con final descendente (120 Hz) que en las que presentan final ascendente (89 Hz). El valor correspondiente a las frases con final descendente es semejante al de las enunciativas, y el de las frases con final ascendente es incluso inferior; ambos valores son claramente inferiores a los de los rangos en interrogativas y exclamativas. Han aparecido también pequeñas variaciones según la longitud de la frase, aunque no se dan en las frases de todos los locutores - *vid.* figura 4.27. Este hecho parece confirmarse con la presencia de pequeñas separaciones entre los dos picos en los histogramas de ambos subgrupos.

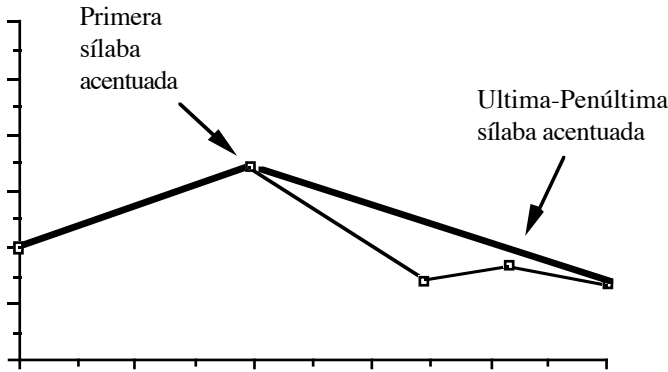
4.5. Clasificación de los esquemas entonativos obtenidos

Los esquemas expuestos en los apartados anteriores se pueden resumir en una serie de esquemas melódicos básicos. Sin embargo, hay que distinguir entre los **patrones melódicos** propiamente dichos, que son los que se exponen en el apartado 4.5.1, y las **formas superpuestas**, que se pueden encontrar, como su nombre indica, añadidas a los diferentes patrones, modificando su forma original. Éstas se describen en el apartado 4.5.2.

4.5.1. Patrones melódicos

Las curvas melódicas de las diferentes frases analizadas en este estudio pueden incluirse dentro de uno de los siguientes patrones:

Figura 4.28.- PATRON DESCENDENTE



CARACTERISTICAS:

1) Subida inicial hasta la primera sílaba acentuada. Se toman como referencia los valores de las enunciativas:

Val. medios PENDIENTE (Hz/ms) ALTURA TONAL (Hz)

Enun. 0,38

2) Segmento final descendente, que comienza en la última o penúltima sílaba acentuada. La altura tonal al final de este segmento se sitúa siempre por debajo del valor inicial de la Fo al inicio de la curva. La pendiente del segmento varía un poco de un tipo a otro de frase, aunque se puede tomar como referencia la de las enunciativas:

Val. medios PENDIENTE (Hz/ms) ALTURA TONAL (Hz)

Enun. - 0,27

3) Presencia de declinación cuando aparece más de un pico en la curva.

4.5.1.1. Patrón descendente

Este patrón, definido en la figura 4.28, sería propio de frases enunciativas, de algunas interrogativas pronominales, de bastantes exclamativas y de la mayoría de las volitivas. De forma general, presentaría los siguientes elementos:

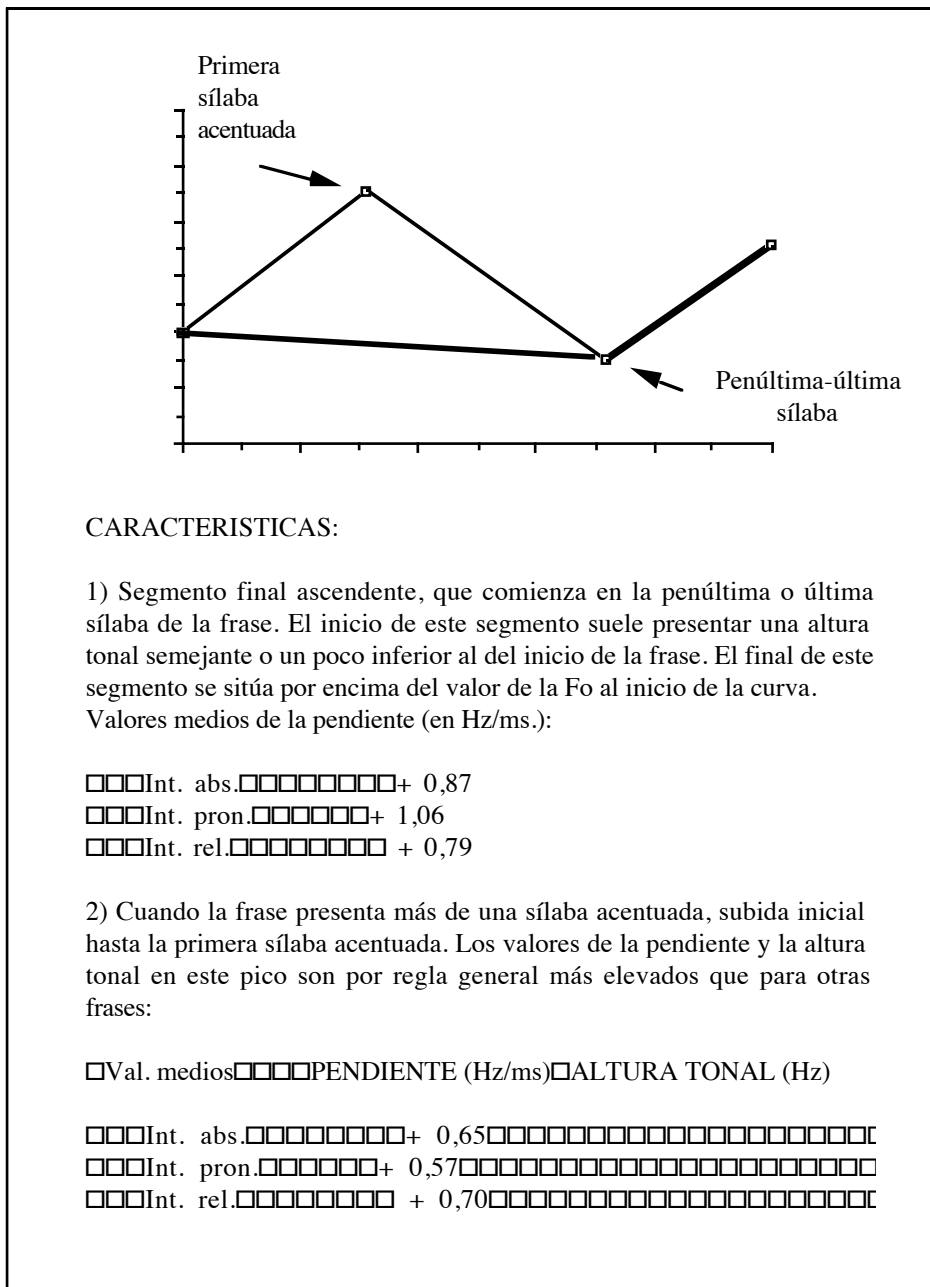
- a) Una **subida de tono en las proximidades de la primera sílaba acentuada**, que en la mayoría de los casos será la que alcance mayor altura. La pendiente y la altura final de esta subida varían según el tipo de frase de que se trate, como se indica más adelante.
- b) Una serie de **picos secundarios**, de menor altura tonal, que tienden también a coincidir con otros acentos de la frase. La situación de estos picos no es fija, y depende de varios factores: el tipo de frase de que se trate, su número de sílabas tónicas o la palabra en la que el locutor quiera hacer recaer el énfasis. En estos picos se da, por regla general, el fenómeno de la **declinación**, caracterizado por un descenso progresivo, desde el inicio de la frase hasta el final, de la altura tonal en los distintos picos.
- c) Un **descenso final** del tono a partir de la última inflexión. También aquí hay variaciones, en función del tipo de frase, por lo que se refiere al valor de la F_0 al final del segmento y a la pendiente de éste. Sin embargo, podrían considerarse como valores de base los obtenidos para las enunciativas. Una explicación más detallada de estas variaciones se encuentra más adelante.

En este patrón podrían incluirse igualmente los pocos casos de frases enunciativas con final ascendente, si se considera condición suficiente que el valor de la F_0 al final de la curva sea inferior al inicial, característica que cumplen tanto las enunciativas con final descendente como las que presentan final ascendente. Queda pendiente un estudio perceptivo para confirmar o desechar esta hipótesis.

4.5.1.2. Patrón ascendente-interrogativo

Es el patrón típico de las frases interrogativas absolutas, de algunas pronominales y de las interrogativas relativas. Sus características básicas serían (figura 4.29):

Figura 4.29.- PATRON ASCENDENTE-Interrogativo

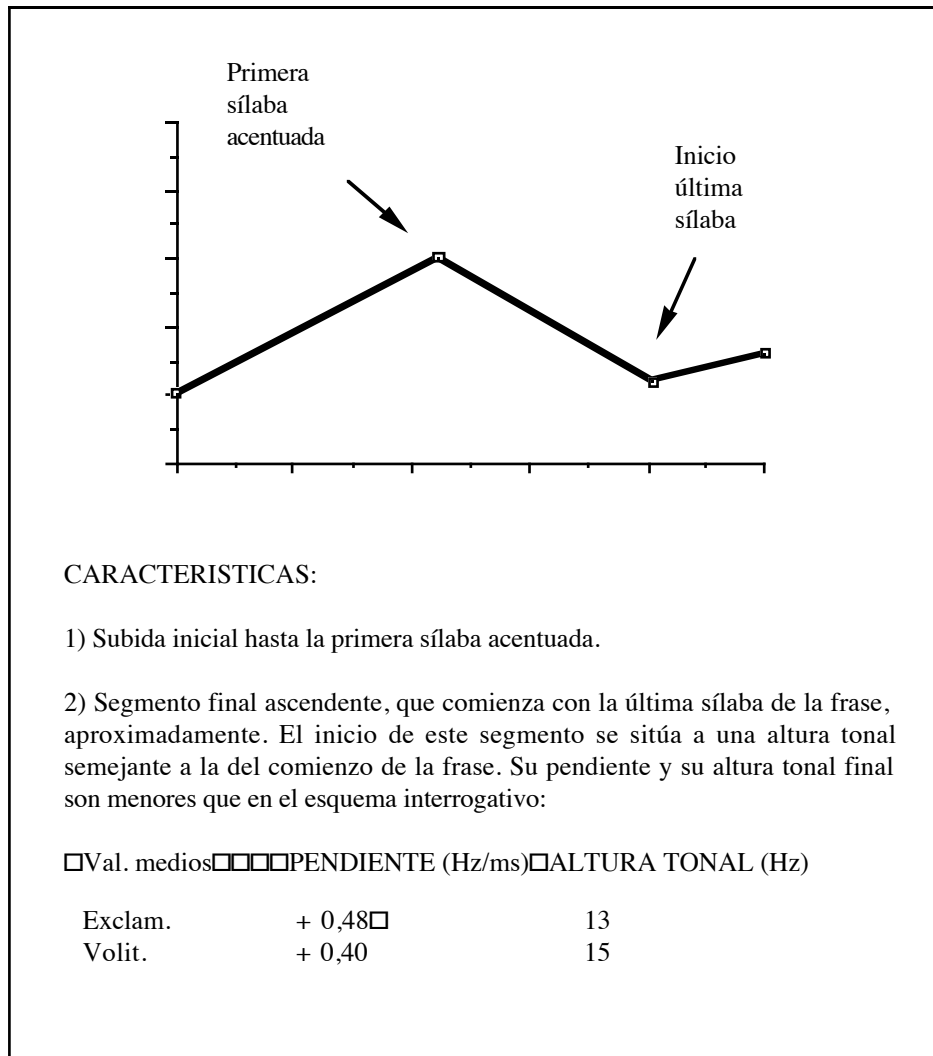


a) Un **segmento final claramente ascendente**, con una pendiente bastante pronunciada: los valores medios oscilan entre los 0,79 (interrogativas relativas) y los 1,06 Hz/ms. (interrogativas pronominales). La altura tonal al final de este segmento es igualmente muy elevada: +171 Hz por término medio en las interrogativas absolutas, +157 Hz en las interrogativas relativas, y +146 Hz en las interrogativas pronominales.

Los resultados de los t-tests indican que tanto los valores de las pendientes como los de la altura tonal están lo suficientemente próximos como para que puedan considerarse un solo grupo: en todos los casos menos uno - comparación de la pendiente en pronominales y relativas -, el valor de p es mayor que 0,05. Sin embargo, los valores obtenidos son mayores para la altura tonal que para las pendientes, lo que parece indicar que el primero es un parámetro más significativo en este caso.

Al igual que ocurre en el patrón anterior, las frases interrogativas absolutas y relativas con final descendente halladas en este estudio podrían considerarse variantes del patrón ascendente-interrogativo si se toma como condición general no el signo de la pendiente del segmento final (positivo o negativo) sino la altura de la F_0 al final de la curva. Al contrario que en las enunciativas, todos los valores de F_0 al final de estas curvas están muy por encima del valor de la misma en su inicio. También serían necesarios más estudios, sin embargo, para confirmar la importancia perceptiva de este hecho.

b) Según los casos - cuando la frase tiene más de una sílaba acentuada -, puede darse una **subida importante en la F_0** al comienzo de la frase, en las proximidades de la primera sílaba tónica. Sin embargo, si la frase tiene un solo acento, esta primera subida se minimiza, e incluso llega a desaparecer. Cuando aparece, la pendiente de esta primera subida es bastante semejante en los tres tipos de frases: 0,65, 0,57 y 0,70 Hz/ms. por término medio en las interrogativas absolutas, pronominales y relativas, respectivamente. La altura tonal que alcanza este pico es también bastante semejante en los tres grupos: +96 Hz normalizados por término medio (interrogativas absolutas), +123 Hz (interrogativas pronominales) y +110 Hz (interrogativas relativas). Los resultados de los t-tests realizados indican, en ambos casos, que las semejanzas registradas son significativas: todos los valores de significación están por encima de 0,05 menos uno (comparación entre la altura tonal en pronominales con final ascendente y descendente). Sin embargo, a diferencia de lo que ocurre con el segmento final, la pendiente parece ser aquí más relevante que la altura de la F_0 , a la vista de los valores de una y otra.

Figura 4.30.- PATRON ASCENDENTE-Enunciativo

4.5.1.3. Patrón ascendente-enunciativo

Sería un patrón propio de frases con contenido expresivo, pero que también guardan alguna relación con la enunciación. Así ocurre con las exclamativas y volitivas, concretamente. Este esquema presentaría como características generales (figura 4.30):

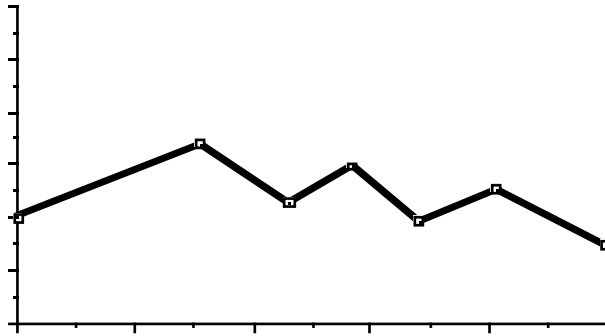
- a) Una **subida de Fo** en las proximidades de la primera sílaba tónica. Los valores de la pendiente y de la altura en este pico varían según el locutor y el tipo de frase, pudiendo llegar, en las exclamativas, a valores semejantes a los de las interrogativas.
- b) Un **segmento final ascendente**, pero con una pendiente menos acusada que en el caso de las interrogativas. Los valores medios de pendiente en estos casos son -0,42 Hz/ms. para las volitivas de este tipo, y -0,49 Hz/ms. para las exclamativas. Igualmente, el valor de Fo al final de este segmento es inferior al de las interrogativas: +13 Hz normalizados en las exclamativas, por término medio, y +15 Hz en las volitivas. La comparación, por medio de t-tests, entre los valores de ambos parámetros en exclamativas y enunciativas revela un alto grado de semejanza: $p=0,941$ en la altura tonal y $0,689$ para la pendiente, lo que hace pensar que la altura de la Fo es más relevante en este caso, como en el segmento final del patrón ascendente-interrogativo.

4.5.2. Formas superpuestas

Bajo este nombre podríamos reunir a una serie de elementos que de alguna manera se añaden a los esquemas básicos descritos en el apartado anterior, agregando matices o incluso transmitiendo la mayor parte de la información entonativa. Cinco son los recursos de este tipo detectados durante el presente estudio (*vid.* figura 4.31):

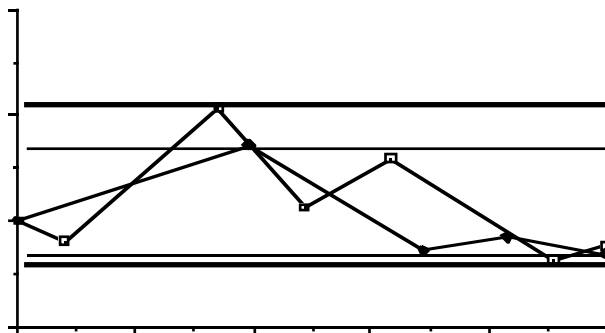
4.5.2.1. "Esquema ondulado"

Este recurso consistiría en el aumento del número de picos de una frase determinada, con respecto al esquema básico sobre el que se superpone. Es un recurso que transmite información de tipo expresivo, y que se añade al esquema descendente, generalmente en las frases exclamativas.

Figura 4.31.- FORMAS SUPERPUESTAS**1) Esquema ondulado**

CARACTERISTICAS:

1) Se marcan con picos un mayor número de sílabas que en la frase enunciativa correspondiente.

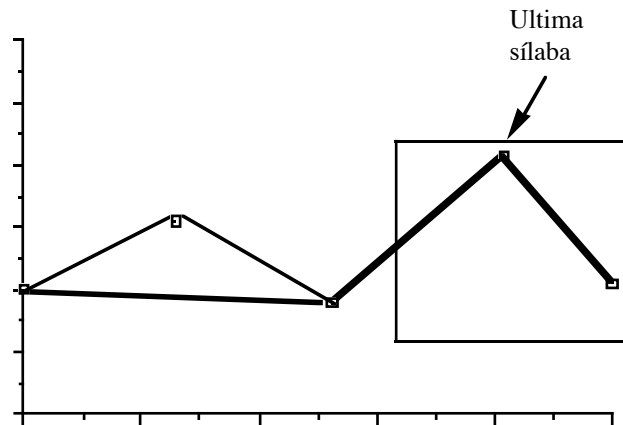
2) Aumento del rango frecuencial

CARACTERISTICAS:

1) Se aumentan las diferencias entre el punto más elevado y el más bajo de la curva.

2) Se utiliza como marca general de expresividad; en la exclamativas, se utiliza como un rasgo que refuerza el contenido expresivo.

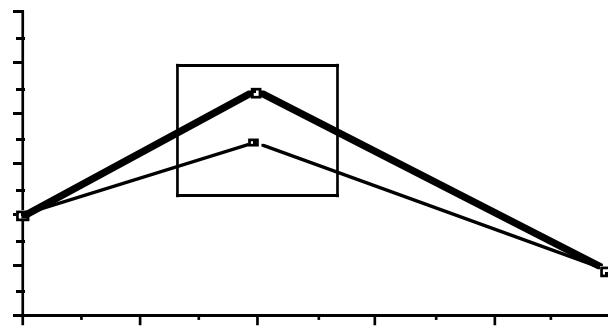
3) Esquema circunflejo



CARACTERISTICAS:

- 1) Presencia de un final circunflejo en la última sílaba de la frase, sea acentuada o no, consistente en un segmento ascendente seguido de un segmento descendente dentro de la misma sílaba.

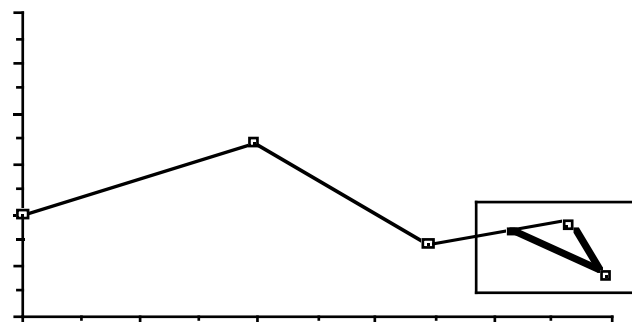
4) Elevación de la altura tonal o la pendiente del primer pico



CARACTERISTICAS:

- 1) Aumento de la pendiente, de la altura tonal o de ambas a la vez en el primer pico de la frase.

5) Variación de la pendiente en el segmento final



CARACTERISTICAS:

- 1) Aumento de la pendiente del segmento final. También puede ir acompañado de un descenso del tono al final del segmento.

4.5.2.2. Aumento del rango frecuencial

A diferencia del recurso anterior, que al menos en este estudio queda bastante limitado a las frases exclamativas, el aumento del rango frecuencial es un elemento que puede darse en cualquier tipo de frase, y que, en principio, sería bastante propio de cada locutor. En cierta forma, sería una medida de lo que normalmente se denomina "expresividad" de un hablante. Tal como se ha señalado, cada locutor parece presentar una tendencia propia a variar más o menos el rango frecuencial al hablar, sea cual sea el tipo de frase que esté pronunciando.

No se ha considerado aquí finalmente el rango frecuencial de los distintos tipos de frases como un elemento distintivo, intrínseco para cada uno. En algún caso, como en las interrogativas, el rango medio se sitúa claramente por encima del de algunos otros tipos, pero se presta también a confusiones con el de las exclamativas - aunque en éstas el rango varía bastante en función del locutor. Además, las grandes fluctuaciones que se dan en el rango dentro de cada tipo provocan muchos solapamientos que llevan a pensar en la poca relevancia de este parámetro para la distinción de la modalidad oracional.

4.5.2.3. Esquema circunflejo

Aparece en algunas frases exclamativas y volitivas, y consistiría en la aparición de una subida y bajada rápida del tono en la última sílaba de la frase, con independencia de que ésta sea acentuada o inacentuada, tal como ya describe Navarro (1948). La altura de F_0 en este pico suele ser bastante elevada, de manera que en ocasiones es el punto de frecuencia más alta de toda la frase.

4.5.2.4. Elevación de la altura tonal o la pendiente del primer pico

La utilización de este recurso respondería a una doble motivación:

a) Para marcar el carácter interrogativo de frases con patrón entonativo descendente. Así ocurriría en las frases interrogativas pronominales con patrón descendente, para distinguirlas de enunciativas y volitivas:

| | Altura tonal (Hz) | Pendiente (Hz/ms.) |
|-----------|-------------------|--------------------|
| Enun. | 59 | 0,38 |
| Int. pro. | 80 | 0,75 |
| Vol. | 59 | 0,61 |

Altura tonal y pendiente parecen ir aquí ligados, dado que en ambos casos el valor medio más elevado es el de las interrogativas pronominales. Este aumento iría en consonancia con el de las otras interrogativas, con lo que podría pensarse que la elevación en el primer pico sería una marca común a todas las interrogativas, como ya se ha sugerido anteriormente. Sin embargo, los valores de significación obtenidos en los t-tests no muestran una diferencia clara entre interrogativas pronominales, por un lado, y enunciativas y volitivas, por otro: en unos casos, p se sitúa por debajo de 0,05 (por ejemplo, en la comparación de la pendiente en enunciativas e interrogativas pronominales: $p=0,007$), mientras que en otros supera ese umbral (comparación entre la pendiente de volitivas e interrogativas pronominales: $p=0,598$). Por ello, parece un rasgo que necesitaría un estudio más profundo - a nivel perceptivo, por ejemplo.

b) Como marca de expresividad, ligada a la elevación general del rango que se da en algunas frases exclamativas.

| | Altura tonal (Hz) | Pendiente (Hz/ms.) |
|-------|-------------------|--------------------|
| Enun. | 59 | 0,38 |
| Excl. | 93 | 0,51 |
| Vol. | 59 | 0,61 |

A diferencia del caso anterior, sólo la altura parece ser relevante, puesto que no se aprecian diferencias significativas en cuanto a la pendiente. Sin embargo, al igual que con las interrogativas pronominales, la comparación por medio de t-tests no muestra diferencias significativas en todos los casos ($p=0,492$). Es posible que la elevación del primer pico esté relacionada con el aumento general del rango que se da en las exclamativas de algunos locutores.

4.5.2.5. Variación de la pendiente en el segmento final

Entre enunciativas, por un lado, y exclamativas, volitivas e interrogativas pronominales, por otro, se dan diferencias en la pendiente del segmento final, de manera que ésta es menos pronunciada en las primeras y más pronunciada en las segundas:

| | Alt. tonal final (Hz) | Pendiente (Hz/ms.) |
|-----------|-----------------------|--------------------|
| Enun. | -29 | -0,27 |
| Int. pro. | -42 | -0,52 |
| Excl. | -35 | -0,49 |
| Vol. | -38 | -0,42 |

En este caso, las diferencias parecen más significativas en los valores correspondientes a las pendientes que en los de las alturas tonales, aspecto que se confirma al analizar los resultados de los t-tests: en las pendientes, los valores obtenidos en la comparación de las enunciativas con interrogativas pronominales, volitivas y exclamativas no presentan diferencias significativas en los tres casos, mientras que la comparación de la altura tonal arroja siempre valores de p superiores a 0,05.

La variación de la pendiente en el segmento final se utilizaría también para distinguir los dos patrones ascendentes. Aquí, sin embargo, los valores obtenidos no parecen tan separados como para considerar dos patrones distintos en función de estas diferencias.

4.6. Incidencia de los resultados en la descripción de los patrones melódicos del español

Los resultados de este estudio vienen a confirmar y a completar, de alguna manera, las conclusiones a las que han llegado otros trabajos anteriores, especialmente el de Navarro Tomás, que es el que se tomó como referencia al inicio de este estudio. Así, se ha podido aquí confirmar:

- 1) La importancia del **segmento final descendente** como principal marca de identificación de algunos tipos de frases, enunciativas principalmente.
- 2) La importancia del **segmento final ascendente** como marca característica de la interrogación en sus diferentes formas.
- 3) La alternancia existente en las frases **interrogativas pronominales** entre un patrón con **final ascendente**, igual al del resto de interrogaciones, y otro con **final descendente**, muy semejante al de las enunciativas.
- 4) La alternancia también en las **frases exclamativas** de diferentes patrones, uno con **final descendente** y otro con **final ascendente**, además de la utilización del **patrón circunflejo** y del **esquema ondulado** como marcas típicas de expresividad.
- 5) La existencia para las frases **volitivas** de un patrón con **final ascendente** y otro con **final descendente**.

6) Los resultados de este estudio indican que en los grupos melódicos terminales del español pueden aparecer tres tipos de segmentos finales diferentes:

- un segmento que podría llamarse de **semicadencia** siguiendo la terminología tradicional, que sería el del patrón descendente. Los valores de la pendiente de este segmento oscilarían, aproximadamente, entre 0 y -1 Hz/ms. En algunos casos también podría definirse el segmento de acuerdo con la altura de la Fo al final del mismo, incluso aunque la pendiente sea positiva. La semicadencia se marcaría así con finales por debajo de la altura de la Fo al inicio de la curva.
- un segmento de **semianticadencia**, correspondiente al patrón ascendente-enunciativo. La pendiente oscilaría normalmente en este caso entre 0 y 0,5 Hz/ms. Si se tomase la altura tonal del final del segmento como punto de referencia, ésta se situaría en la mayoría de los casos por encima de los +75 Hz con respecto a la Fo en el inicio de la curva.
- un segmento de **anticadencia**, que aparece en el patrón ascendente-interrogativo, y que correspondería, aproximadamente, a pendientes superiores a los 0,5 Hz/ms. Los límites en cuanto a altura tonal serían 0 y 75 Hz por encima de la Fo al inicio de la curva.

La existencia de un segmento de cadencia no queda confirmada en este estudio, pero hay indicios de que puede aparecer en algunas frases no enunciativas con patrón descendente. Serían aquellas frases cuya pendiente final está por debajo de -1 Hz/ms., aunque los datos presentados aquí no permiten establecer una frontera en cuanto a la altura final de Fo. En ese caso, salvo el tonema continuativo, que en principio sería propio de grupos no terminales, los demás tonemas finales establecidos por Navarro Tomás para el español estarían presentes en las frases analizadas aquí. Las tesis de Navarro Tomás parecen confirmarse, frente a las de Quilis y otros autores, en cuanto al número de tonemas finales en español.

7) El hecho de que el primer pico también pueda tener importancia a la hora de diferenciar entre patrones melódicos diferentes parece confirmar la idea de que en español son relevantes tanto el segmento final como el inicial de las curvas melódicas.

Por contra, existen algunos aspectos en los que los resultados de este estudio difieren de lo establecido en trabajos anteriores, y que parece interesante destacar:

- 1) Por lo que se refiere a la caracterización de las curvas, tradicionalmente se ha considerado que las frases enunciativas comenzaban su segmento final en la última sílaba acentuada. Algunos casos de nuestro *corpus* revelan, como ya se ha señalado, que también puede comenzar este segmento en las proximidades de la **penúltima sílaba** tónica.

2) Algunos autores - Navarro (1948) o Canellada y Kulhmann (1987), por ejemplo - han relacionado las **interrogativas relativas** con **esquemas entonativos circunflejos**. En este estudio, el esquema típico de las interrogativas relativas ha resultado ser el mismo que en de las interrogativas absolutas, con **segmento final ascendente**. También es posible, sin embargo, que el tipo de interrogativa relativa utilizada en este trabajo, que podría definirse como una “interrogativa confirmativa”, no coincida con las interrogativas relativas analizadas en los otros estudios.

3) Navarro Tomás distingue entre esquemas entonativos propios de las frases exclamativas y esquemas propios de las volitivas. A la vista de los resultados de nuestro estudio, se hace difícil mantener esta distinción: la semejanza de las pendientes y de la altura de la Fo en el segmento final, la utilización en ambos casos tanto de patrones ascendentes como descendentes, la aparición de finales circunflejos en uno y en otro tipo, son argumentos a favor de incluirlos en un mismo grupo. Tan sólo muestran diferencias en cuanto a la altura y la pendiente del primer pico, mayores en las exclamativas.

5. CONCLUSIONES GENERALES

En este último apartado se exponen las conclusiones finales del trabajo. Se analizan en primer lugar (apartado 5.1), las posibilidades de aplicación de los resultados obtenidos a sistemas automáticos de síntesis y reconocimiento, tanto por lo que se refiere al procedimiento de estilización como a la información obtenida acerca de los patrones melódicos. En el apartado 5.2, finalmente, se hace un repaso de las cuestiones que quedan pendientes tras este estudio.

5.1. Aplicación de los resultados a la síntesis y el reconocimiento

El método de estilización y los patrones obtenidos en este estudio pueden emplearse para la elaboración de sistemas automáticos de reconocimiento y síntesis de entonación. Este estudio proporciona información sobre la estructura de las curvas melódicas del español de la que no se disponía hasta ahora, o sólo se tenía conocimiento parcial (Olabe, 1983):

- 1) **Valores numéricos** sobre la **pendiente** y la **altura tonal** de los segmentos finales en español.
- 2) **Valores numéricos** sobre la **pendiente** y la **altura tonal** en la primera subida de los diferentes patrones.
- 3) **Valores numéricos** sobre el **rango frecuencial** de las curvas en español.

En la figura 5.1 se presenta un esquema con los rasgos que, de acuerdo con este estudio, caracterizarían la curva melódica de los diferentes tipos de frase. En los dos apartados que siguen, se analiza brevemente cómo incorporar esta información tanto en la síntesis como en el reconocimiento.

| TIPO DE FRASE | PATRON | FORMAS SUPERPUESTAS |
|--------------------------------------|---|---|
| Enunciativas | Descendente | |
| Interrogativas absolutas y relativas | Ascendente-interrogativo | |
| Interrogativas pronominales | Descendente Ascendente-interrogativo | Elevación primer pico |
| Exclamativas y volitivas | Descendente Ascendente-enunciativo | (Aumento rango) (Patrón circunflejo) (Descenso pendiente final) |

Figura 5.1.- Una propuesta de reglas para la generación y el reconocimiento automáticos de patrones melódicos terminales utilizando la clasificación presentada en este estudio. Los paréntesis indican que se trata de un patrón opcional; cuando se indica más de un patrón para un tipo determinado de frase, significa que se elige cada vez uno solo.

5.1.1. Aplicación a la síntesis

Los patrones melódicos estilizados obtenidos en este trabajo pueden emplearse como modelos para la generación de curvas de F_0 en un módulo de síntesis. Los tres patrones definidos se utilizarían para configurar el esquema básico de la curva melódica, haciendo coincidir sus puntos de inflexión con las sílabas señaladas para cada uno. Los valores intermedios de F_0 entre los puntos de inflexión se calcularían por interpolación, bien siguiendo líneas rectas, bien utilizando funciones más complejas.

Sobre la curva de F_0 resultante se superpondrían las variaciones dependientes del número de sílabas tónicas del enunciado, que aquí no se han tratado de sistematizar, pero que se han revelado importantes: a cada sílaba tónica no le corresponde siempre un pico de F_0 , ni todos los picos coinciden con sílabas tónicas; sin embargo, el número de picos en la frase sí parece tener una relación con la duración de la misma, y con el número de sílabas tónicas.

A continuación, dicho esquema básico se modificaría con las formas superpuestas que correspondiesen al tipo de frase que se estuviese generando.

En el caso de que la oración estuviese constituida por más de un grupo fónico, entrarían además en juego las reglas que definirían las curvas en los grupos no terminales. Está por determinar todavía si serían reglas que se superpondrían a los esquemas generales a la oración, o si bien servirían a su vez como patrones básicos para generar las curvas de grupos no terminales.

5.1.2. Aplicación al reconocimiento

El sistema de estilización definido aquí podría automatizarse y servir para extraer representaciones simplificadas sobre las que se realizarían los procesos posteriores de análisis. Como paso previo sería necesario incorporar un procedimiento de **alisamiento** que de alguna manera minimizase los errores introducidos durante la fase de estimación de la F_0 .

Por lo que se refiere a la estrategia de reconocimiento en sí, la variabilidad en el número de picos, ya mencionada en el apartado anterior, parece desechar un enfoque que implique el reconocimiento global de la curva. El proceso de análisis debería centrarse sobre partes determinadas de la misma, como son el último segmento y el primer pico detectado. Una serie de reglas de decisión basadas, por ejemplo, en el valor de la F_0 al final del último segmento o en la altura de la F_0 en el primer pico, servirían para reconocer los patrones melódicos definidos en este estudio, e igualmente podrían definirse reglas similares para el reconocimiento de los diferentes esquemas superpuestos. La identificación de una curva determinada con una modalidad oracional en concreto se realizaría en función del esquema básico y de las formas superpuestas detectadas en la misma.

Un sistema de estilización del tipo empleado para este estudio, que utiliza contornos y no niveles, tiene en su favor que permite eliminar alguno de los problemas con que se han encontrado los diferentes sistemas desarrollados hasta el momento (*vid.* apartado 2.2, "Entonación y reconocimiento automático"), al no necesitar de una segmentación previa de la señal en sílabas o en alófonos.

Es evidente que el establecimiento de una reglas de decisión a partir de los datos obtenidos aquí no es tarea fácil. Las variaciones observadas en los distintos parámetros parecen aconsejar un enfoque estadístico, más que el establecimiento de umbrales que distingan entre las diferentes categorías. Los resultados estadísticos obtenidos en este estudio pueden emplearse para el desarrollo de un sistema de reconocimiento tanto probabilístico como categorial.

5.2. Cuestiones pendientes

Los resultados de este trabajo dejan una serie de cuestiones pendientes que abren la vía para estudios posteriores, tanto por lo que se refiere al procedimiento de estilización (apartado 5.2.1) como a los patrones melódicos obtenidos (apartado 5.2.2).

5.2.1. Sistema de estilización

El análisis de las representaciones obtenidas en este estudio mediante el procedimiento de estilización definido anteriormente indica que éste debe refinarse aún más. Algunos de los puntos de inflexión mantenidos en ciertas curvas correspondían claramente a variaciones micromelódicas dentro de una misma sílaba (diferencia vocal -consonante) o del mismo sonido. El umbral definido para eliminar los puntos de inflexión innecesarios ha de ser revisado en posteriores trabajos.

Está también pendiente de realización un estudio perceptivo sobre la validez de las representaciones estilizadas obtenidas mediante este método. Serían tareas posibles en este sentido:

- A) Comprobar que la información entonativa se mantiene en dichas representaciones.
- B) Comprobar si todos los puntos de inflexión mantenidos en la curva mediante este procedimiento son necesarios para preservar dicha información.

5.2.2. Patrones melódicos

Las cuestiones por resolver, en lo que se refiere a los patrones melódicos del español, pueden clasificarse en dos grupos principales. Así, algunos estudios posibles que se derivarían del presentado aquí serían:

A) A nivel acústico:

- 1) La validación de estas conclusiones con un grupo mayor de locutores, o en un *corpus* de habla espontánea.
- 2) La realización de estudios acústicos sobre cuestiones pendientes de la entonación del español, como la declinación.
- 3) La ampliación del estudio iniciado con este trabajo a los patrones de grupos no terminales del español.

B) A nivel perceptivo:

- 1) Una validación perceptiva de los patrones obtenidos en este estudio. Sería especialmente interesante de cara a validar la utilidad de dichos patrones en un sistema de síntesis.
- 2) La realización de estudios perceptivos que permitan aclarar algunas de las distinciones entre tipos de oraciones que han quedado algo oscuras en este estudio:
 - En primer lugar, la diferenciación entre las oraciones enunciativas y las interrogativas pronominales. Aunque se apunta la posibilidad de que sea la altura del pico inicial, o su pendiente, la que distinga ambos tipos, sería interesante comprobar si esta diferencia tiene valor distintivo en el plano lingüístico. Es posible también que la diferencia resida simplemente en la presencia o ausencia de la partícula interrogativa en la oración, con lo que la distinción se realizaría a nivel léxico y no fonético.
 - La distinción entre las interrogativas absolutas y las relativas, con esquemas prácticamente iguales en este estudio, y sin elementos semánticos que en principio puedan distinguirlos, especialmente en el *corpus* empleado, donde las frases eran exactamente iguales en los dos grupos.
 - Finalmente, habría que comprobar si las oraciones exclamativas y volitivas, que en este estudio presentan unos patrones muy semejantes, se diferencian perceptivamente de alguna manera. Es posible que la presencia de palabras con matiz de ruego o de mandato, como los verbos en forma imperativa, influya también en el etiquetado de la frase.

Tales cuestiones, o al menos algunas de ellas, serán objeto de próximos trabajos. De hecho, como ya se ha indicado en la introducción, el estudio que se ha presentado en estas páginas ha servido para abrir una línea de trabajo con muchísimas ramificaciones, puesto que la caracterización de los fenómenos entonativos del español plantea todavía muchos interrogantes.

ANEXO

I. FRASES AISLADAS

1. ENUNCIATIVAS

- 1.1. Ramón llegó en avión.
- 1.2. Ramón llegó andando.
- 1.3. Mario llegó en avión.
- 1.4. Mario llegó andando.
- 1.5. Ramón.
- 1.6. Mario.

2. INTERROGATIVAS

2.1. Absolutas

- 2.1.1. ¿Viene hoy Ramón?
- 2.1.2. ¿Viene hoy Mario?
- 2.1.3. ¿Vendrá hoy Ramón?
- 2.1.4. ¿Vendrá hoy Mario?
- 2.1.5. ¿Viene?
- 2.1.6. ¿Vendrá?

2.2. Pronominales

- 2.2.1. ¿Dónde vivió Ramón?
- 2.2.2. ¿Dónde vivió Mario?
- 2.2.3. ¿Adónde irá Ramón?
- 2.2.4. ¿Adónde irá Mario?
- 2.2.5. ¿Dónde?
- 2.2.6. ¿Adónde?

2.3. Relativas

- 2.3.1. ¿Ramón llegó en avión?
- 2.3.2. ¿Ramón llegó andando?
- 2.3.3. ¿Mario llegó en avión?
- 2.3.4. ¿Mario llegó andando?
- 2.3.5. ¿Ramón?
- 2.3.6. ¿Mario?

- 2.3.7. ¿Dónde vivió Ramón?
- 2.3.8. ¿Dónde vivió Mario?
- 2.3.9. ¿Adónde irá Ramón?
- 2.3.10. ¿Adónde irá Mario?
- 2.3.11. ¿Dónde?
- 2.3.12. ¿Adónde?

3. EXCLAMATIVAS

- 3.1. ¡Ramón llegó en avión!
- 3.2. ¡Ramón llegó andando!
- 3.3. ¡Mario llegó en avión! *(admiración)*
- 3.4. ¡Mario llegó andando!
- 3.5. ¡Ramón!
- 3.6. ¡Mario!

- 3.7. ¡Ramón llegó en avión!
- 3.8. ¡Ramón llegó andando!
- 3.9. ¡Mario llegó en avión! *(desaprobación)*
- 3.10. ¡Mario llegó andando!
- 3.11. ¡Ramón!
- 3.12. ¡Mario!

4. VOLITIVAS

- 4.1. Llama a Ramón.
- 4.2. Llama a Mario.
- 4.3. Llamad a Ramón. *(orden)*
- 4.4. Llamad a Mario.
- 4.5. Llámale.
- 4.6. Llamadle.

- 4.7. Llama a Ramón.
- 4.8. Llama a Mario.
- 4.9. Llamad a Ramón. *(ruego)*
- 4.10. Llamad a Mario.
- 4.11. Llámale.
- 4.12. Llamadle.

II. DIALOGOS

1. - ¿Sabes? Ramón llegó en avión. (1.1)
 - ¿Ramón llegó en avión? (2.3.1)
 ¡Pero si a él le aterroriza volar!
 - Llámale. El te lo confirmará. (4.11)
 - ¡Ramón llegó en avión! ¡Casi no puedo creerlo! (3.1)

2. - ¿Sabes? Ramón llegó andando. (1.2)
 - ¿Ramón llegó andando? (2.3.2)
 ¡Pero si hay más de una hora de camino!
 - Llamadle. El os lo confirmará. (4.12)
 - ¡Ramón llegó andando! ¡Casi no puedo creerlo! (3.2)

3. - ¿Sabes? Mario llegó en avión. (1.3)
 - ¿Mario llegó en avión? (2.3.3)
 ¡Pero si a él le aterroriza volar!
 - Llama a Ramón. El te lo confirmará. (4.7)
 - ¡Mario llegó en avión! ¡Casi no puedo creerlo! (3.3)

4. - Mario llegó andando. (1.4)
 - ¿Mario llegó andando? (2.3.4)
 ¡Pero si hay más de una hora de camino!
 - Llamad a Ramón. El os lo confirmará. (4.9)
 - ¡Mario llegó andando! ¡Casi no puedo creerlo! (3.4)

5. - ¿Sabéis? Mario y Ramón ya han llegado.
 - ¿Ramón? ¿Cuándo ha llegado? (2.3.5)
 - No estoy seguro. Ramón ha llegado... a las tres,
 me parece, un poco antes que Mario.
Llamad a Mario. El os lo confirmará. (4.10)
6. - ¿Sabes? Mario y Ramón ya han llegado.
 - ¿Mario? ¿Cuándo ha llegado? (2.3.6)
 - No estoy seguro. Mario llegó... a las tres,
 me parece, un poco antes que Ramón.
Llama a Mario. El te lo confirmará. (4.8)
7. - ¿Viene hoy Ramón? (2.1.1)
 - Sí, vendrá esta tarde.
8. - ¿Viene hoy Mario? (2.1.2)
 - Sí, vendrá esta tarde.
9. - ¿Vendrá hoy Ramón? (2.1.3)
 - No, vendrá mañana. Hoy no puede.
10. - ¿Vendrá hoy Mario? (2.1.4)
 - No, vendrá mañana. Hoy no puede.
11. - ¿Viene? (2.1.5)
 - ¿Quién?
 - Mario. (1.6)
 - No, todavía no se le ve.
12. - ¿Vendrá? (2.1.6)
 - ¿Quién?
 - Ramón. (1.5)
 - No, me dijo que hoy no podría.

13. - ¿Dónde vivió Ramón? (2.2.1)
 - En París.
 - ¿Dónde? (2.3.11)
 - ¿No me has oído? En París.
14. - ¿Dónde vivió Mario? (2.2.2)
 - En París.
 - ¿Qué has dicho? ¿Dónde vivió Mario? (2.3.8)
 - He dicho que en París.
15. - ¿Adónde irá Ramón? (2.2.3)
 - A París.
 - ¿Adónde? (2.3.12)
 - ¿No me has oído? A París.
16. - ¿Adónde irá Mario? (2.2.4)
 - A París.
 - ¿Qué has dicho? ¿Adónde irá Mario? (2.3.10)
 - He dicho que a París.
17. - ¿Pero qué haces ahí parado? Llama a Mario. (4.2)
 ¿No me has oído?
18. - ¿Pero qué haces ahí parado? Llama a Ramón. (4.1)
 ¿No me has oído?
19. - ¿Pero qué hacéis ahí parados? Llamad a Mario. (4.4)
 ¿No me habéis oído?
20. - ¿Pero qué hacéis ahí parados? Llamad a Ramón. (4.3)
 ¿No me habéis oído?
21. - ¿Pero qué haces ahí parado? Llámale. (4.5)
 ¿No me has oído?

22. - ¿Pero qué hacéis ahí parados? Llamadle. (4.6)
 ¿No me habéis oído?
23. - ¿Sabes? Ramón vivió en París dos años. (2.3.7)
 - No te he oído bien. ¿Dónde vivió Ramón?
 - En París.
 - ¿En París? ¡Ramón! ¡Qué suerte la tuya! (3.5)
24. - ¿Sabes? Ramón irá a París este verano. (2.3.9)
 - No te he oído bien. ¿Adónde irá Ramón?
 - A París.
 - ¿A París? ¡Ramón! (3.11)
 ¿Por qué no nos has dicho nada?
25. - ¿Sabes? Mario se va hoy. (2.2.6)
 - ¿Ah, sí? ¿Adónde?
 - A París.
 - ¿A París? ¡Mario! (3.12)
 ¿Por qué no nos has dicho nada?
26. - ¿Sabes? Mario vivió fuera dos años. (2.2.5)
 - ¿Ah, sí? ¿Dónde?
 - En París.
 - ¿En París? ¡Mario! ¡Qué suerte la tuya! (3.6)
27. - ¿Sabes? ¡Ramón llegó en avión! (3.7)
 Qué ostentación, ¿verdad?
28. - ¿Sabes? ¡Ramón llegó andando! (3.8)
 Qué vulgaridad, ¿verdad?
29. - ¿Sabes? ¡Mario llegó en avión! (3.9)
 Qué ostentación, ¿verdad?

30. - ¿Sabes? ¡Mario llegó andando!
Qué vulgaridad, ¿verdad? (3.10)

REFERENCIAS

- AKERS, G. - LENNIG, M. (1985).- "Intonation in text-to-speech synthesis: Evaluation of algorithms", *Journal of the Acoustical Society of America*, 77, pp. 2157-65.
- ALARCOS LLORACH, E. (1960).- "Esquemas fonológicos de la frase", *Lengua y enseñanza. Perspectivas*, Madrid, Ministerio de Educación Nacional, pp. 47-52.
- ALCINA, J. - BLECUA, J.M. (1975).- *Gramática española*, Barcelona, Ariel, 1983⁴.
- BLACK, J. (1949).- "Natural frequency, duration, and intensity of vowels in readings", *Journal of Speech and Hearing Disorders*, 14, pp. 216-221.
- BOE, L.J. (1973).- "Étude de l'interaction source laryngienne-conduit vocal dans la détermination des caractéristiques intrinsèques des consonnes du français", *Bulletin de l'Institut de Phonétique de Grenoble*, 2, pp. 1-24.
- BOLINGER, D.L. (1951).- "Intonation: levels versus configurations", *Word*, 7, 3, pp. 199-210.
- BOLINGER, D.L. (1955).- "Intersections of stress and intonation", *Word*, 11, 2, pp. 195-205.
- BOLINGER, D.L. (1989).- *Intonation and its uses. Melody in grammar and discourse*, Londres, Edward Arnold.
- CAELEN, G. (1981).- *Structures prosodiques de la phrase énonciative simple et étendue*, Hamburgo, Hermut Buske.
- CANELLADA, M.J. (1941).- "Notas de entonación extremeña", *Revista de Filología Española*, XXV, pp. 79-91.
- CANELLADA, M.J. - KUHLMANN, J. (1987).- *Pronunciación del español*, Madrid, Castalia.
- COHEN, A. - T HART, J. (1967).- "On the anatomy of intonation", *Lingua*, 19, pp.

177-192.

COHEN, A. - COLLIER, R.- T HART, J. (1982).- "Declination: construct or intrinsic feature of speech pitch?", *Phonetica*, 39, pp. 254-73.

COLLIER, R. - TERKEN, J. (1987).- "Intonation by rule in text-to-speech applications", *European Conference on Speech Technology. Edinburgh, September 1987*, Edimburgo, CEP Consultants Ltd., pp. 165-68.

COLLIER, R. (1989).- "Intonation analysis: the perception of speech melody in relation to acoustics and production", en TUBACH, J.P. - MARIANI, J.J. (Eds.) (1989).- *Eurospeech 89*, Vol. 1, pp. 38-44.

COLLIER, R. *et al.* (1989).- "On the perceptual salience of melodic variations and its consequences for intonation synthesis", en TUBACH, J.P. - MARIANI, J.J. (Eds.) (1989).- *Eurospeech 89*, Vol. 2, pp. 108-111.

CONTRERAS, H. (1976).- *A Theory of Word Order with Special Reference to Spanish*, Amsterdam, North Holland.

COOPER, W.E.- SORENSEN, J.M. (1981).- *Fundamental Frequency in Sentence Production*, Nueva York, Springer Verlag.

CRUTTENDEN, A. (1986).- *Intonation*, Cambridge, Cambridge University Press (Cambridge Textbooks in Linguistics); traducción al español de I. MASCARO (1990): *Entonación. Teoría general y aplicación al inglés*, Barcelona, Teide, (Serie Lingüística).

DANES, F. (1960).- "Sentence intonation from a functional point of view", *Word*, 16, pp. 34-54.

DE BOT, K. (1986).- "The transfer of Intonation and the Missing Data Base", en KELLERMAN, E. - SHARWOOD-SMITH, M. (Eds.)- *Crosslinguistic Influence in Second Language Acquisition*, Oxford, Pergamon Press, pp. 110-19.

DEBELLO, N. *et al.* (1987).- "Design of an Italian Text-To-Speech System", en LAVER, J.- JACK, M.A. (Eds.)- *European Conference on Speech Technology. Edinburgh, September 1987*, Edimburgo, CEP Consultants Ltd., pp. 17-20.

DENES, P. (1959) "A preliminary investigation of certain aspects of intonation", *Language and Speech*, 2, pp. 106-22.

DENES, P. - MILTON-WILLIAMS, J. (1962) "Further studies in intonation", *Language and Speech*, 5, pp. 1-14.

DI CRISTO, A. (1978).- "L'intonation est congruente à la syntaxe: une confirmation", en ROSSI *et al.* (1981).- *L'intonation, de l'acoustique à la sémantique*, París, Klincksieck, pp. 272-89.

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

ENRIQUEZ, E.V. - CASADO, C. - SANTOS, A. (1989).- "La percepción del acento en español", *Lingüística Española Actual*, 11, pp. 241-269.

FANT, L. (1984).- *Estructura informativa en español. Estudio sintáctico y entonativo*, Uppsala, Acta Universitatis Upsalensis (Studia Romanica Upsaliensis, 34).

FLANAGAN, J. - SASLOW, M. (1958).- "Pitch discrimination of synthetic vowels", *Journal of the Acoustical Society of America*, 30, pp. 435-42.

FLANAGAN J. L. - LANDGRAF, F. (1968).- "Self-Oscillating Source for Vocal - Tract Synthesizers", *IEEE Transactions in Audio and Electroacustics*, 16, 1, pp. 57-65.

FUJISAKI - HIROYA - NAGASHIMA (1969).- "A model for synthesis of pitch contours of connected speech", *Annual Report, Engineering Research Institute, University of Tokyo*, 28, pp. 53-60.

FUJISAKI - HIROYA - SUDO (1971).- "Synthesis by rule of prosodic features of connected Japanese", *Proceedings of 7th International Congress on Acoustics*, vol. 3, pp. 133-36.

FUJISAKI - HIROYA - KEIKICHI HIROSE (1982).- "Modelling the dynamic characteristics of voice fundamental frequency with applications to analysis and synthesis of intonation", *Preprints of Papers, Working Group in Intonation, 13th International Congress of Linguists, Tokyo*, pp. 57-70.

GANDOUR, J. - WEINBERG, B. (1980).- "On the relationship between vowel height and fundamental frequency: evidence from esophageal speech", *Phonetica*, 37, 5-6, pp. 344-54.

GARRO, L. - PARKER, F. (1983).- "Relative Clauses in Spanish: Some Suprasegmental Characteristics", *Journal of Phonetics*, 11, pp. 85-99.

GILI GAYA, S. (1924).- "Influencia del acento y de las consonantes en las curvas de entonación", *Revista de Filología Española*, 11, pp. 154-177.

GOLDSTEIN, J.L. (1973).- "An optimum processor theory for the central formation of the pitch of complex tones", *Journal of the Acoustical Society of America*, 54, pp. 1496-1516.

GW INSTRUMENTS INC. (1988).- *MacSpeech Lab II. User's Manual. Version 1.0*, Cambridge, MA, GW Instruments.

HOMBERT, J. M.(1978) "Consonantal types, vowel quality and tone", en FROMKIN (ed.)- *Tone: a linguistic survey*, New York, Academic Press, pp. 71-111.

HOUSE, A. - FAIRBANKS, G. (1953).- "The influence of consonant environment upon

the secondary acoustical characteristics of vowels", *Journal of the Acoustical Society of America*, 25, pp. 105-113.

JASSEM, W. (1952).- *Intonation of Conversational English*, Wrocław.

JASSEM, W.- DEMENKO, G. (1986).- "Extracting Linguistic Information from Fo traces", en JOHNS-LEVIS, C. (ed.) (1986).- *Intonation in Discourse*, Londres, Croom Helm, pp. 1-17.

KINGDON, R. (1958).- *The Groundwork of English Intonation*, Londres.

KLATT, D. H. (1973).- "Discrimination of fundamental frequency contours in synthetic speech: Implications for models of speech perception", *Journal of the Acoustical Society of America*, 53, pp. 8-16.

KULLOVA, J. (1982).- "Acerca del contorno melódico de la oración hispánica", *Ibero-america praguensia*, XVI, pp. 69-76.

KULLOVA, J. (1988).- *Modulaciones de la cadena hablada en español*, Praga, Univerzita Karlova.

KVAVIK, K.H. - OLSEN, C.L. (1974).- "Theories and methods in Spanish intonational studies. Survey", *Phonetica*, 30, pp. 65-100.

LARREUR, D. - BOE, L.J. (1973).- "Étude de l'influence des variations de la fréquence laryngienne sur l'intelligibilité et la qualité des consonnes generées par vocodeur", *Bulletin de l'Institut de Phonétique de Grenoble*, 2, pp. 103-26.

LEHISTE, I. - PETERSON, G. (1961).- "Some basic considerations in the analysis of intonation", *Journal of the Acoustical Society of America*, 33, pp. 419-425.

LEHISTE, I. (1970).- *Suprasegmentals*, Cambridge, The MIT Press.

LÉON, P.R.- MARTIN, Ph. (1969).- *Prolegomènes a l'étude des structures intonatives*, Ottawa, Didier.

LÉON, P.R. (1970).- "Systématique des fonctions expressives de l'intonation", en LÉON, P. - FAURE, G. - RIGAULT, A. (eds.) (1970).- *Analyse des faits prosodiques*, Ottawa, Didier, pp. 57-72.

LÉON, P.R. (1971).- *Essais de phonostylistique*, Montreal, Didier.

LÉON, P.R. (1971).- "Où en sont les études sur l'intonation", *Proceedings of the Seventh International Congress of Phonetic Sciences, 1971*, La Haya, Mouton, pp. 113-56.

- LIEBERMAN, Ph. (1967).-*Intonation, perception and language*, Cambridge, MIT Press.
- MAEDA, S. (1976).- *A characterization of American English Intonation*, PhD dissertation, MIT.
- MARTIN Ph. (1970).- "Reconnaissance des patrons intonatifs", en LÉON, P. - FAURE, G. - RIGAULT, A. (eds.) (1970).- *Analyse des faits prosodiques*, Ottawa, Didier, pp. 175-91.
- MARTIN, Ph. (1973).- "Les problèmes de l'intonation: recherches et applications", *Langue française*, 19, pp. 4-32.
- MARTIN, Ph. (1982).- "Phonetic Realisations of Prosodic Contours in French", *Speech Communication*, 1, pp. 283-294.
- MARTINET, A. (1956).- *La description phonologique avec application au parler franco-provençal d'Hauteville (Savoie)*, París.
- MATEO, A. (1988).- "Experimento sobre el tono intrínseco de las vocales castellanas", *Estudios de Fonética Experimental*, III, pp. 157-179.
- MERTENS, P. (1989).- "Automatic Recognition of Intonation in French and Dutch", en TUBACH, J.P. - MARIANI, J.J. (Eds.) (1989).- *Eurospeech 89*, Vol. 1, pp. 46-50.
- MORENO, P.J. *et al.* (1989).- "Improving Naturalness in a Text to Speech System with a New Fundamental Frequency Algorithm", en TUBACH, J.P. - MARIANI, J.J. (Eds.) (1989).- *Eurospeech 89*, Vol. 1, pp 360-63.
- NAVARRO TOMAS, T. (1948).- *Manual de entonación española*, Madrid, Guadarrama, 1974⁴.
- O'CONNOR, J.D. - ARNOLD, G.F. (1961).- *Intonation of Colloquial English*, Singapur, Longman, 1973².
- O'SHAUGHNESSY, D. (1979).- "Linguistic features in fundamental frequency patterns", *Journal of Phonetics*, 7, pp. 119-145.
- O'SHAUGHNESSY, D. (1987).- "The fundamental frequency generator", en ALLEN *et al.* (1987).- *From text to speech: the MITalk system*, Cambridge, Cambridge University Press.
- OLABE, J.C. (1983).- *Sistema para la conversión de un texto ortográfico a hablado en tiempo real*, Tesis doctoral, Escuela Técnica Superior de Ingenieros de Telecomunicación, Universidad Politécnica de Madrid.

OLIVE, J. -LIBERMAN, M. (1979).- "A set of concatenative Units for Speech Synthesis", en WOLF, J. - KLATT, D. (eds.) (1979).- *Speech Communication Papers Presented at the 97th Meeting of the Acoustical Society of America*, New York, Acoustical Society of America, pp. 515-518.

PARDO, J.M. *et al.* (1987).- "Improving Text to Speech Conversion in Spanish Linguistic Analysis and Prosody", en LAVER, J.- JACK, M.A. (Eds.).- *European Conference on Speech Technology. Edinburgh, September 1987*, Edimburgo, CEP Consultants Ltd., Vol 2, pp. 173-76.

PETERSON, G.- BARNEY, H. (1952).- "Control methods used in a study of the vowels", *Journal of the Acoustical Society of America*, 24, pp. 175-184.

PIERREHUMBERT, J. (1979).- "The perception of fundamental frequency declination", *Journal of the Acoustical Society of America*, 66, pp. 363-69.

PIERREHUMBERT, J. (1987).- *The Phonology and Phonetics of English Intonation*, Bloomington, Indiana University Linguistics Club.

QUILIS, A.- FERNÁNDEZ, J.A. (1979).- *Curso de fonética y fonología españolas*, Madrid, CSIC.

QUILIS (1981).- *Fonética acústica de la lengua española*, Madrid, Gredos (Biblioteca Románica Hispánica, Manuales 49).

REAL ACADEMIA ESPAÑOLA (1973).- *Esbozo de una nueva gramática de la lengua española*, Madrid, Espasa-Calpe, 1982⁸.

ROSSI, M. (1977).- "L'intonation et la troisième articulation", *Bulletin de la société de linguistique de Paris*, LXXII (1), pp. 55-68.

ROSSI, M. (1981).- "Intonation, énonciation, syntaxe", en ROSSI *et al.* (1981).- *L'intonation, de l'acoustique à la sémantique*, Paris, Klincksieck, pp. 200-249.

ROSSI, M. (1985).- "L'intonation et l'organisation de l'énoncé", *Phonetica*, 42, pp. 135-53.

ROSSI *et al.* (1981).- *L'intonation, de l'acoustique à la sémantique*, Paris, Klincksieck.

SCHOUTEN, J. F. (1940).- "The residue and the mechanism of hearing", *Proc. K. Ned. Akad. Wet.*, 43, pp. 991-99.

SCHOUTEN, J. F. (1970).- "The residue revisited", en PLOMP - SMOORENBURG (eds.).- *Frequency Analysis and Periodicity Detection in Hearing*, Leiden, Sijthoff.

SORIN, C. (1982).- "Evaluation de la contribution de Fo à l'intelligibilité", *Recherches/Acoustique*, VII, Lannion, CNET, pp. 141-155.

STOCKWELL, R. P. *et al.* (1956).- "Spanish juncture and intonation", *Language*, 32, pp. 641-65.

STUDDERT-KENNEDY, M. - HADDING, K. (1973).- "Auditory and linguistic processes in the perception of intonation contours", *Language and Speech*, 16, pp. 293-313.

SZMIDT, Y. (1968).- "Étude de la phrase interrogative en français canadien et en français standard", *Studia Phonetica*, 1, pp. 192-209.

SZMIDT, Y. (1979).- "Niveaux de voix caractéristiques des questions totales", en LÉON, P.R.- ROSSI, M (eds.) (1979).- *Problèmes de prosodie. Vol II. Experimentations, modèles et fonctions*, Ottawa, Didier, pp. 17-28.

T HART, J. (1974).- "Discriminability of the size of pitch movements in speech", *IPO Annual Progress Report*, 9, pp.56-63.

T HART, J. (1979).- "Explorations in automatic stylization of F0 curves", *IPO Annual Progress Report*, 14, pp. 61-65.

T HART, J. - COLLIER, R. (1975).- "Integrating different levels of intonation analysis", *Journal of Phonetics*, 3, pp. 235-255.

T HART, J. - COLLIER, R. - COHEN, A. (1990).- *A perceptual Study of Intonation. An Experimental - Phonetic Approach to Intonation*, Cambridge, Cambridge University Press.

TAKEFUTA, Y. (1975).- "Method of Acoustic Analysis of Intonation" en SINGH, S. (Ed.), *Measurement Procedures in Speech, Hearing and Language*, Baltimore, University Park Press, pp. 363-378.

TERHARDT, E. (1972a).- "Zur Tonhöhenwahrnehmung von Klängen. I. Psychoakustische Grundlagen", *Acustica*, 26 pp. 173-86.

TERHARDT, E. (1972b).- "Zur Tonhöhenwahrnehmung von Klängen. II. Ein Funktionsschema", *Acustica*, 26 pp. 187-99.

TERKEN, J.M.B. - COLLIER, A. (1989).- "Automatic synthesis of natural-sounding intonation for text-to-speech conversion in Dutch", en TUBACH, J.P. - MARIANI, J.J. (eds.) (1989).- *Eurospeech 89*, Vol. I, pp. 357-9.

THORSEN, N. (1978).- "An acoustical investigation of Danish intonation", *Journal of Phonetics*, 6, pp. 151-175.

THORSEN, N. (1979).- "Interpreting Raw Fundamental-Frequency Tracings of Danish", *Phonetica*, 36, pp. 57-78.

THORSEN, N. (1980).- "A study of the perception of sentence intonation - Evidence from

Danish", *Journal of the Acoustical Society of America*, 67, 3, pp. 1014-1030.

THURLOW, W. R. (1963).- "Perception of low auditory pitch: a multicue mediation theory", *Psychological Review*, 70, pp. 515-19.

VAISSIÈRE, J. (1983).- "Language-Independent Prosodic Features", en CUTLER, A. - LADD, D.R. (1983).- *Prosody: Models and Measurements*, Berlín, Springer, pp. 53-66.

VAISSIÈRE, J. (1989).- "On automatic extraction of prosodic information for automatic speech recognition system", en TUBACH, J.P. - MARIANI, J.J. (eds.) (1989).- *Eurospeech 89*, Vol. I, pp. 202-205.

WAIBEL, A. (1988).- *Prosody and Speech Recognition*, Londres, Pitman.

WHITFIELD, I.C. (1967).- *The Auditory Pathway*, Londres, Arnold.

WHITFIELD, I.C. (1970).- "Central nervous processing in relation to spatio-temporal discrimination of auditory patterns", en PLOMP - SMOORENBURG (eds.)- *Frequency Analysis and Periodicity Detection in Hearing*, Leiden, Sijthoff.