



UNIVERSITAT_{DE}
BARCELONA

Estructura de los modelos entonativos: interpretación fonológica del acento y la entonación en castellano

Francisco José Cantero Serena



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution 4.0. Spain License.**

Universitat de Barcelona
Departament de Filologia Romànica / Secció de Lingüística General

Programa de doctorado: Aplicaciones de la Lingüística
Bienio 1987 - 1989

Tesis Doctoral

**ESTRUCTURA DE LOS MODELOS ENTONATIVOS:
INTERPRETACIÓN FONOLÓGICA DEL ACENTO Y LA ENTONACIÓN
EN CASTELLANO**

presentada por

Francisco José Cantero Serena

en 1995

para optar al título de *Doctor en Filología Románica*

Dirigida por el Dr. *Eugenio Martínez Celdrán*

AGRADECIMIENTOS Y DEDICATORIA

Para mí es muy agradable mirar hacia atrás y comprobar cómo tantas personas valiosas han coincidido en la trayectoria vital de este trabajo y le han aportado una brizna de su sabiduría, de su experiencia o de su amor:

Eugenio Martínez Celdrán ha sabido ser un maestro y un amigo, y todo lo que sé de fonética y fonología se lo debo a él.

Valeria Salcioli jalonó el inicio de este trabajo con largas conversaciones, llenas de ideas y de entusiasmo, y me brindó algunos ejemplos tan interesantes.

José Enrique Gargallo me puso en contacto con mis informantes más exóticos: Yun-yuoung Choi, Midori Hara, Holmfridur Mathiasdottir, Bodö Mihaly, Salmán Raheel, Chichu Scanu, Agata Siwierska, Nicoleta G. Tudor y Annemarie Wagner. Las pruebas perceptivas fueron posibles gracias a la colaboración de un centenar de alumnos de Filología y de Magisterio. Igualmente, múltiples informantes anónimos han prestado su voz para mis análisis.

Mis alumnos del Máster de E/LE han puesto a prueba, año tras año, la bondad teórica de los conceptos más novedosos empleados en este trabajo, obligándome a formularlos con la mayor precisión.

Guillermo Andrés Toledo tuvo la bondad de entusiasmarse con mis explicaciones, y me llenó de confianza.

Marta Balcells ha tenido la generosidad de confiar en el método de análisis de la entonación que propongo en este trabajo, y de emplearlo en el suyo propio.

M^a Carmen Ferriz me prestó una gran ayuda en la búsqueda y captura de material bibliográfico.

Teresa Marbà veló porque mi horario fuera muy flexible, y seguramente es suya la culpa de que este trabajo conste de tantas páginas.

Antonio Mendoza veló por mí.

Francisco De Arriba y Sofía García me brindaron la tranquilidad necesaria para comenzar a escribir: me envolvieron en algodón, y las palabras surgían gozosas.

Su hija, Clara, puso en orden mi vida y mi mente, tiempo atrás. Este trabajo contiene mis palabras y mis ideas, es cierto, pero tiene su aliento. Un trabajo de investigación lingüística, entonces, también puede ser un poema de amor.

Por eso se lo dedico.

A Clara

CONTENIDO

0. PRESENTACIÓN Y OBJETIVOS	9
1. INTRODUCCIÓN	15
1.1. Aspectos acústicos y articulatorios del sonido	18
1.2. Cuestiones previas de fonética y fonología	35
1.3. Aspectos fonéticos de la entonación	47
1.4. El concepto de entonación	69
2. PRINCIPALES ENFOQUES Y MÉTODOS DE ANÁLISIS DE LA ENTONACIÓN	83
2.1. Introducción	85
2.2. La tradición británica	97
2.3. La tradición norteamericana	133
2.4. La escuela holandesa	157
2.5. La escuela española	169
2.6. Conclusión: cuestiones abiertas en el estudio de la entonación . . .	213
3. EL ACENTO Y LA ENTONACIÓN PRELINGÜÍSTICA	225
3.1. Acento paradigmático y ritmo	227
3.2. El acento como fenómeno segmental	259
3.3. Acento sintagmático y grupo fónico	289
3.4. Acento y entonación	305

4. LA ENTONACIÓN LINGÜÍSTICA	335
4.1. Cuestiones previas	337
4.2. El significado entonativo	367
4.3. El método de análisis formal	411
4.4. Modelos entonativos	451
5. LA ENTONACIÓN NO LINGÜÍSTICA	481
5.1. Expresividad y márgenes de dispersión	483
5.2. Motivación y arbitrariedad	503
6. CONCLUSIONES	525
7. REFERENCIAS BIBLIOGRÁFICAS	541
APÉNDICE I: EXPERIMENTO SOBRE LA PAUSA	555
APÉNDICE II: ANALIZADOR MELÓDICO	581
APÉNDICE III: «UN ENSAYO DE CUANTIFICACIÓN DE LAS ENTONACIONES LINGÜÍSTICAS»: RESULTADOS	595
ÍNDICE DE AUTORES	615
ÍNDICE DE FIGURAS	627
ÍNDICE GENERAL	635

0. PRESENTACIÓN Y OBJETIVOS

Durante el curso 1985-86, en mi último año de carrera, cursé la asignatura optativa *Fonética Experimental*, impartida por el Dr. E. Martínez Celdrán, la actividad central de la cual era la realización de un pequeño trabajo de investigación experimental.

Aprovechando la reciente adquisición, por parte del *Laboratorio de Fonética*, del *Visi-Pitch* (un visualizador del tono en tiempo real, controlado por ordenador, que permitía un análisis muy detallado de los fenómenos tonales del habla), el Dr. Martínez Celdrán me sugirió enfocar el trabajo de la asignatura en el estudio de algún aspecto concreto de la entonación. Mi primera intención, pues, fue la de encarar el estudio de algún fenómeno entonativo muy limitado, fácilmente abarcable, desde una perspectiva acústica, con la única finalidad de realizar un pequeño experimento aislado. Con esta idea comencé a introducirme en la bibliografía científica sobre la entonación: el tema, obviamente, me apasionó de inmediato, y me raptó.

Pronto me di cuenta, sin embargo, de que no iba a ser fácil dedicarme al estudio de algún tópico concreto en el mundo múltiple y difuso de la entonación, que se aparecía ante mí como un bosque frondoso, casi selvático, del que nadie había trazado aún el mapa definitivo. Ciertamente, los tratados más completos y sistemáticos sobre la entonación aparecieron poco tiempo después, a partir de 1986, y mis fuentes documentales eran entonces escasas. Supuse que no era lógico conformarse con unas cuantas mediciones aisladas, sin un marco teórico completo en el tuvieran sentido. Así, una cierta ignorancia documental me llevó a afrontar en su conjunto una tarea que se ha demostrado

ardua, conceptualmente agotadora, como es la interpretación fonológica de la entonación.

Desde entonces, han sido diez años de estudio y reflexión continuos, de cientos de análisis acústicos y docenas de pruebas perceptivas, en un itinerario de zig-zag difícil de reconstruir, que me han suministrado suficientes datos concretos como para proceder a un ordenamiento de buena parte de los hechos tonales relevantes lingüísticamente. Si no he conseguido ajardinar el bosque, pues no era mi intención dar cuenta exhaustiva de todos los fenómenos tonales, relevantes o no relevantes, sí creo haber abierto, al menos, sendas limpias de maleza que permiten adentrarse en él y recorrerlo en toda su amplitud, sin riesgo de perderse.

Tras cursar el programa de doctorado *Aplicaciones de la Lingüística*, el Dr. Martínez Celdrán aceptó dirigirme el proyecto y puso a mi disposición el *Laboratorio de Fonética*, del que es director. En él he realizado buena parte de mis análisis, hasta que en 1992 se creó el *Laboratorio de Fonética Aplicada* del Departamento de Didáctica de la Lengua y la Literatura de la Universidad de Barcelona, del que soy fundador y responsable, y al que trasladé toda mi actividad científica.

En principio, pues, mi intención era llevar a cabo un análisis acústico sistemático de los fenómenos tonales del castellano, con el objeto de ofrecer una descripción completa y objetiva, instrumental, de los mismos. Sin embargo, volví a tropezar con el mismo escollo del marco teórico insuficiente: cuando me disponía a sistematizar mis observaciones, siempre surgía algún problema de carácter teórico que impedía contemplar los hechos descritos dentro de un esquema elaborado por otro autor. Tenía que optar, entonces, entre realizar una descripción instrumental dentro de la concepción teórica de alguna de las escuelas tradicionales (v. Cap. 2), que a mi juicio no abordaban desde una perspectiva estrictamente lingüística, según los parámetros de la fonología moderna, el estudio de los hechos tonales, o diseñar una nueva concepción teórica de los fenómenos, dentro de un nivel estricto de análisis fonológico.

En general, los problemas teóricos de las concepciones tradicionales de la entonación se dividen en tres grandes temas:

a) *Las relaciones entre el acento y la entonación.* Por una parte, la mayoría de los autores presenta el acento como un fenómeno de intensidad, a pesar de que hay la evidencia experimental de que se trata de un fenómeno tonal (3.1.1.3.); por otra parte, dicha evidencia no ha dado lugar aún a una integración teórica de los dos fenómenos tonales, acento y entonación, que se han tratado, y siguen tratándose, como dos fenómenos independientes, a pesar de estar informados por el mismo parámetro físico, la frecuencia fundamental (F_0).

b) *El funcionamiento lingüístico de la entonación.* Los autores que la han considerado un fenómeno de relevancia lingüística indudable han vacilado entre distintos niveles de descripción lingüística: el nivel gramatical, el nivel discursivo, el nivel pragmático; también, un nivel de transición fonológico-morfológico-léxico (como el análisis de niveles: v. 2.3.1.). En ningún caso, sin embargo, se determina el nivel de significación estrictamente entonativo (v. 4.2.), ni qué tipo de signo lingüístico constituye la unidad de la entonación, cuyo carácter suprasegmental la ha relegado a un segundo plano dentro de la fonología clásica, marcadamente segmental.

c) *El papel de la entonación no lingüística.* Otros autores, en cambio, han preferido concebir la entonación como un fenómeno no lingüístico, no sistematizable, precisamente a raíz de su carácter no segmental, y no segmentable. Así, la tradición más rica en el análisis de la entonación (la tradición británica: v. 2.2.) en realidad ni siquiera se plantea su relevancia lingüística. La similitud entre la entonación y otros códigos no lingüísticos (como el complejo gestual), por lo demás, establece su carácter supuestamente no simbólico sino motivado (v. 5.2.), es decir, de relevancia, en todo caso, paralingüística o estilística, pero no plenamente, no estrictamente lingüística. Esta es la razón fundamental, junto con el carácter de fenómeno no segmentable, de que la entonación no encuentre más que un lugar marginal en la teoría fonológica.

Con la intención de salvar, en la medida de nuestras posibilidades, tales problemas teóricos, en el presente trabajo se propone una *interpretación fonológica* de los fenómenos tonales de la lengua castellana (el acento y la entonación) que los integre plenamente en la teoría fonológica estructural, entendida esta como la fonología no dependiente de otros niveles de análisis lingüístico (gramatical, lógico-semántico, etc.). Ciertamente, dentro de la fonología estructural la concepción teórica de los fenómenos del acento y la entonación no parece haber encontrado aún su sitio.

Los objetivos que nos planteamos en el presente trabajo, por tanto, son:

- a) Establecer de qué modo los fenómenos de acento y entonación actúan solidariamente, si lo hacen, ya que se trata de dos fenómenos igualmente tonales: determinar a qué nivel funcionan complementariamente, cómo se conjugan en el discurso y qué funciones cumplen juntos.
- b) Determinar el funcionamiento lingüístico de la entonación, si constituye un fenómeno plenamente fonológico, qué tipo de significado aporta, cuáles son los rasgos fonológicos y fonéticos que entran en juego en su descripción y cuáles son las unidades de la entonación.
- c) Determinar el papel que juega la entonación no lingüística, a qué nivel actúa y cómo se conjuga con la entonación lingüística; determinar cómo el carácter de fenómeno motivado puede conjugarse con el carácter de fenómeno lingüístico.

El trabajo se estructura en seis capítulos, de los cuales los dos primeros hacen las veces de introducción teórica al tema: el capítulo 1, "Introducción", constituye una presentación de algunos de los conceptos básicos que se emplearán a lo largo del trabajo, como el propio concepto de "entonación", los aspectos fonéticos que se implican en él, etc.; por su parte, el capítulo 2, "Principales enfoques y métodos de análisis de la entonación", pretende ser una introducción sistemática a los principales enfoques empleados en el estudio de la entonación, con el ánimo no de ser exhaustivo sino de ser comprensivo, destacando exclusivamente la aportación de los autores más relevantes y dedicando una atención especial a las aportaciones realizadas en nuestra propia tradición lingüística y en nuestra lengua.

El núcleo central del trabajo, que corresponde al desarrollo teórico de los tres temas enunciados como objetivo de la investigación, ocupa los capítulos 3, 4 y 5 del trabajo: "El acento y la entonación prelingüística", "La entonación lingüística" y "La entonación no lingüística".

Finalmente, el capítulo 6, "Conclusiones", pretende ofrecer una visión de conjunto, simple pero sistemática, de las principales aportaciones teóricas del trabajo.

1. INTRODUCCIÓN

1. INTRODUCCIÓN

La entonación es un fenómeno complejo en el que no conviene entrar directamente sin realizar una serie de consideraciones teóricas previas sobre algunos aspectos fonéticos y fonológicos que son de alcance general pero que afectan de un modo importante a la descripción y a la concepción misma del fenómeno.

En concreto, al tratarse de un fenómeno tonal nos interesa establecer las características del sonido (tono, intensidad, timbre y duración) y las relaciones que mantienen entre ellas, para poder determinar después con la mayor precisión el papel que puedan desempeñar en la entonación. Por otra parte, al tratarse de un fenómeno suprasegmental nos interesa también examinar la realidad fonética de los segmentos, el carácter fonológico de la segmentación del continuum sonoro y la relación que pueda tener, de entrada, la entonación con los demás fenómenos suprasegmentales.

Así, en el presente capítulo comenzaremos desde el principio, por la definición del sonido, la descripción de sus características y la producción de la voz articulada (1.1.), y desplazaremos nuestra atención selectivamente por los aspectos de la Fonética y la Fonología que luego tendrán alguna relevancia en la concepción del fenómeno (1.2.).

El estudio de la entonación comienza propiamente en el punto 1.3., en el que se tratan la fisiología de la producción del tono, la percepción, así como la descripción acústica del rasgo tonal.

Finalmente, en el punto 1.4. definimos el fenómeno, examinamos los parámetros que intervienen en él y distinguimos los niveles de abstracción fonológica que lo componen.

1.1. Aspectos acústicos y articulatorios del sonido.

1.1.1. Definición de sonido

Llamamos *sonido* a las variaciones de presión producidas en un medio natural por la vibración de un cuerpo. A esta definición clásica de *sonido*, esencialmente correcta, le faltan, sin embargo, algunos datos, pues no todas las variaciones de presión pueden ser *oídas*: en efecto, el sonido es una sensación producida por la decodificación cerebral de tales vibraciones, cuando son captadas por el oído. Así, podría decirse que el oído es una especialización del sentido del tacto, que nos permite tener una sensación especial, que llamamos *sonora*, de una realidad que es eminentemente táctil: las variaciones de presión producidas por la vibración de un cuerpo.

Aun así, la definición resulta demasiado abstracta. Imaginemos, pues, la cuerda de una guitarra: cuando la presionamos y la soltamos, provocamos una vibración que es fácilmente observable a simple vista, pues se trata de un material muy flexible. La vibración de la cuerda, en la que se repiten ciclos idénticos, es una vibración *periódica*, y produce un *sonido*; sin embargo, también puede haber vibraciones *aperiódicas*, a cuyo producto llamamos *ruido*: un golpe contra un cuerpo sólido como la mesa o la pared hará que este vibre, pero de un modo irregular. En realidad, cualquier cuerpo puede vibrar, pues todos los materiales tienen un punto de flexibilidad, aunque se trate de un cuerpo sólido y compacto: en música, por ejemplo, la percusión no es otra cosa que el aprovechamiento de este fenómeno.

Pero volvamos a nuestra cuerda de guitarra: al presionarla, la desplazamos de su posición de reposo; cuando la soltamos, esa misma presión hace que se

desplace hacia su posición de reposo y que siga en dirección contraria, compensando con la inercia el desplazamiento anterior; una vez llegada al punto

opuesto más lejano correspondiente a la presión ejercida, vuelve atrás y el ciclo se repite hasta que la fuerza empleada en la presión inicial se consume. Una vibración viene a ser como el movimiento oscilatorio de un péndulo, desplazándose a un lado y a otro: en la figura 1 se ha representado esa correspondencia entre la vibración de una cuerda y el movimiento del péndulo.

Si la cuerda vibra en el vacío, no ocurre nada más que la vibración descrita; pero si la cuerda vibra en un medio natural, por ejemplo en un medio gaseoso como el aire, los desplazamientos en que consiste la vibración hacen que la cuerda empuje a infinidad de partículas que encuentra en su camino, chocando con ellas y provocando en ellas un movimiento oscilatorio similar al de la vibración inicial de la cuerda, o como el movimiento del péndulo: como se muestra en la figura 2, cada partícula que

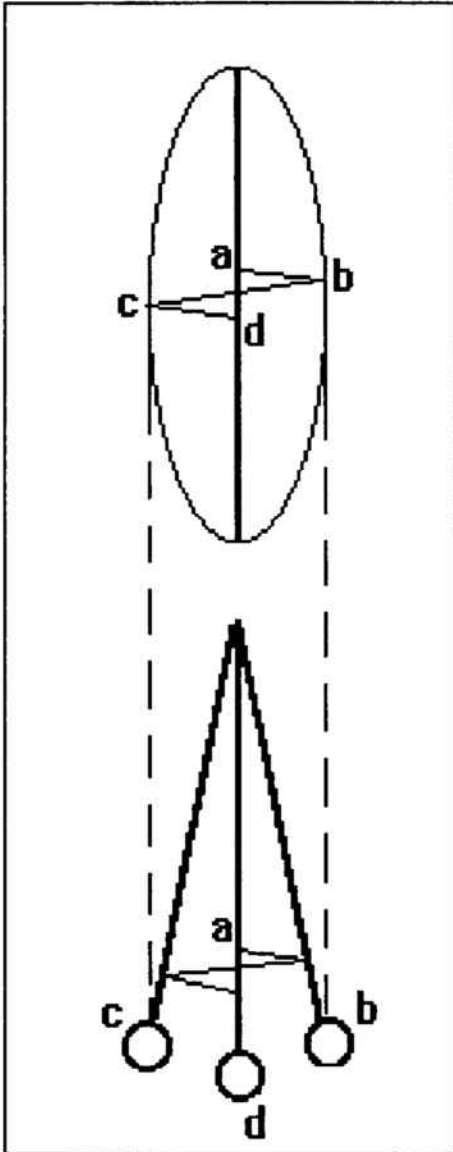


Figura 1 - Correspondencia entre una vibración y el movimiento de un péndulo.

se ve desplazada hacia la posición *b* empuja a su vez a otras partículas, a las que también desplaza, las cuales a su vez empujan a otras, etc.; desde la posición *b*, cada partícula, agotada la presión inicial, vuelve atrás, pasa por *a* (el punto de reposo) y se desplaza hasta *c*, y vuelta a comenzar, hasta que la fuerza de la oscilación se agote.

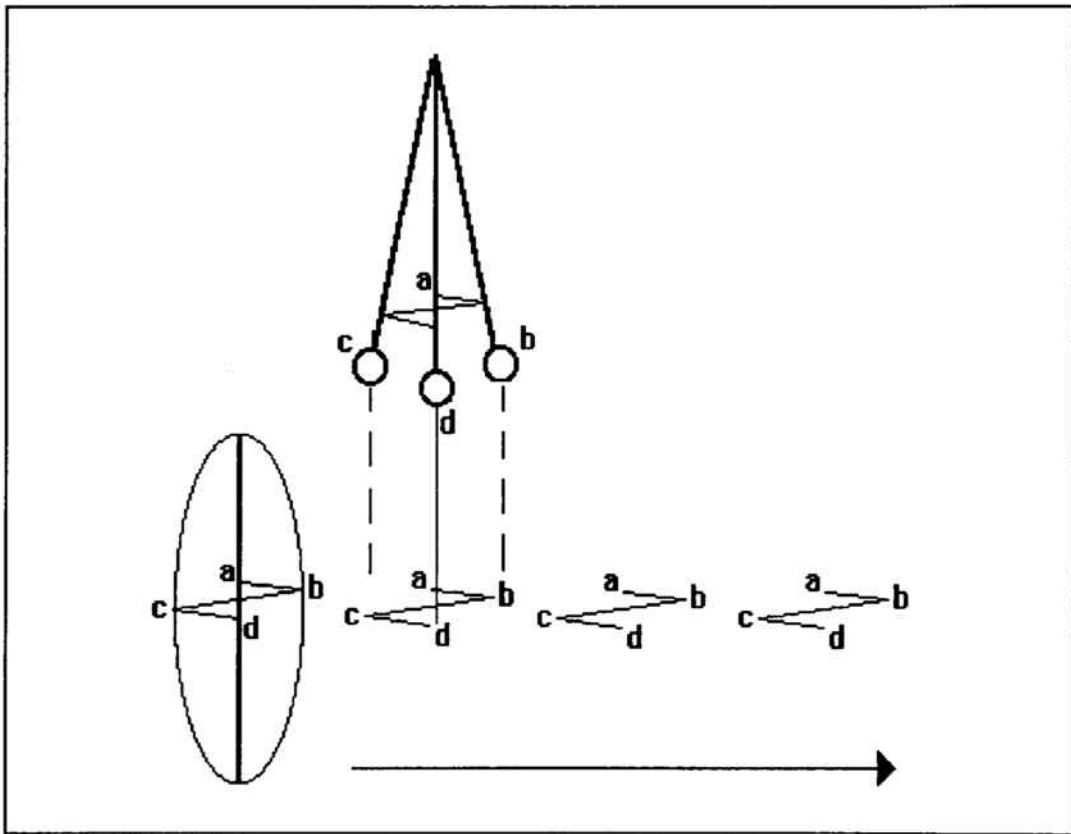


Figura 2 - Esquema de la propagación del sonido.

La vibración original, por tanto, provoca una serie de variaciones de presión en el medio circundante que consiste, como hemos visto, en una serie de movimientos oscilatorios consecutivos a través del espacio: una partícula

empuja a otra, etc. Cuando estas variaciones de presión llegan al oído, nosotros tenemos la sensación de *oír*, pero cuando llegan a otros sensores, como los del tacto, tenemos la sensación de *tocar*. por ejemplo, si ponemos las yemas de los dedos a pocos centímetros de un altavoz, podremos *tocar* el sonido, puesto que lo que notamos en los dedos no es aire, como podría parecer, sino las variaciones de presión a que nos hemos referido. En efecto, el sonido es un

fenómeno táctil, de *presión*, pero que debe ser captado por el oído, cuyos sensores son infinitamente más sensibles y precisos que los sensores del tacto.

El movimiento oscilatorio de las partículas mediante el que se transmite el sonido suele representarse en forma de *onda senoidal* (figura 3): una onda senoidal (llamada así porque es el resultado del seno del ángulo que forma la partícula, en este caso el péndulo, en sus desplazamientos) representa un ciclo completo de una vibración. Sin embargo, cuando hablamos de la *onda sonora* debe entenderse que no es más que una forma de referirse a los movimientos oscilatorios consecutivos de las partículas del medio, y que no es realmente una "onda" en sí, sino un movimiento que representamos visualmente, y de forma esquemática, en forma de onda. No es, por tanto, como las ondas que se producen en el agua

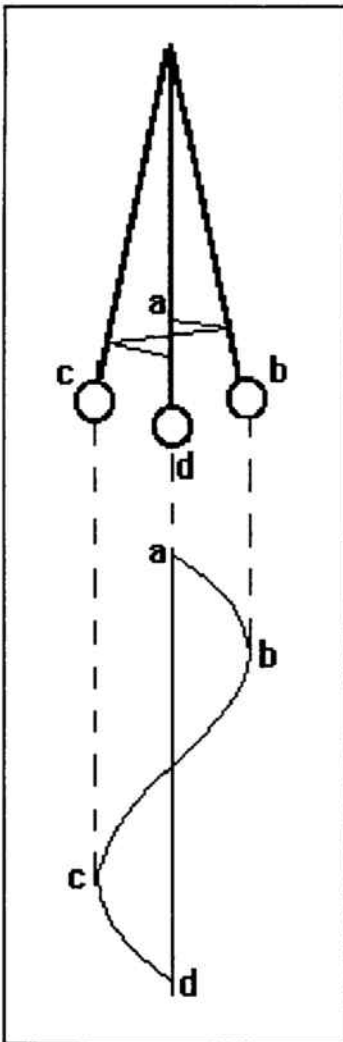


Figura 3 - La onda senoidal como representación de la vibración.

cuando tiramos una piedra, cuyo desplazamiento es espacial, sino que se trata de unas ondas cuyo desplazamiento es temporal (la distancia entre a y d).

1.1.2. Características del sonido

Fijándonos en la representación de la onda periódica simple (figura 4), podemos examinar algunas de las características relevantes del sonido.

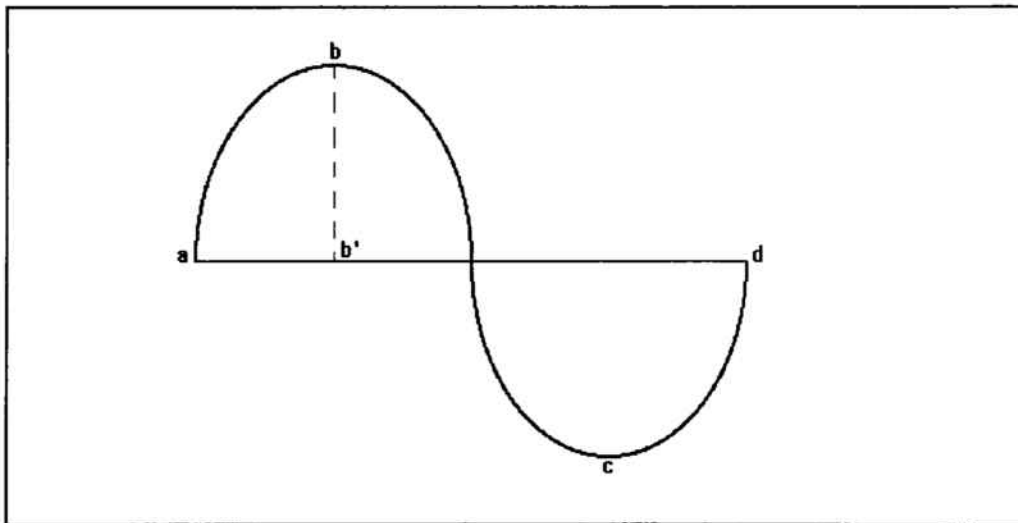


Figura 4 - Onda periódica simple.

Como hemos dicho, la distancia entre a y d es tiempo, medido en segundos: el primer parámetro relevante del sonido es, pues, la *frecuencia* de ciclos completos de la vibración que ocurren en un segundo. Si suponemos que entre a y d ha pasado un segundo, la frecuencia de la vibración representada en la figura 4 sería de 1 ciclo por segundo.

La frecuencia es el parámetro acústico que informa el fenómeno del tono, o de lo que comúnmente se llama "altura musical": a mayor frecuencia, el tono será más agudo; a menor frecuencia, el tono será más grave. La unidad de frecuencia es el ciclo por segundo (c.p.s.) o hertzio (Hz).

La distancia entre b y b' , por su parte, sí es una distancia espacial: corresponde a la amplitud de la vibración, es decir, a la energía empleada en la oscilación. La amplitud de la onda corresponde a la amplitud del desplazamiento del cuerpo vibrante o de las partículas transmisoras de las variaciones de presión, y tiene que ver, por tanto, con la intensidad del sonido: la intensidad de un sonido es directamente proporcional a la amplitud de onda. La unidad de intensidad es el decibelio (dB).

Así, las distintas notas de una guitarra son frecuencias distintas, y la fuerza con que presionemos las cuerdas se corresponderá con una mayor o menor intensidad del sonido.

El tono y la intensidad, sin embargo, no son los fenómenos más relevantes del sonido. Las diferencias de tono e intensidad, por sí solas, sólo nos permitirían una articulación musical del sonido, *y con un solo instrumento*; pero nunca una articulación lingüística. El parámetro acústico más relevante del sonido para la transmisión de información es, en efecto, el timbre.

La representación de la onda sonora mediante una onda simple es una abstracción notable, pues no existen sonidos puros de una sola frecuencia. En realidad, todos los sonidos reales son el producto de multitud de vibraciones simultáneas en el mismo cuerpo vibrante, que provocan otras tantas ondas simples simultáneas, cuya suma es la onda compuesta resultante de la vibración del cuerpo. Así, todos los sonidos (incluso los más puros, como el creado por el diapasón) son el resultado de la adición de varias ondas simples: la onda

sonora siempre es una onda compuesta. En la figura 5 vemos cómo dos ondas simples producidas, por ejemplo, por dos vibraciones simultáneas de un mismo cuerpo, si se suman dan como resultado una onda compuesta.

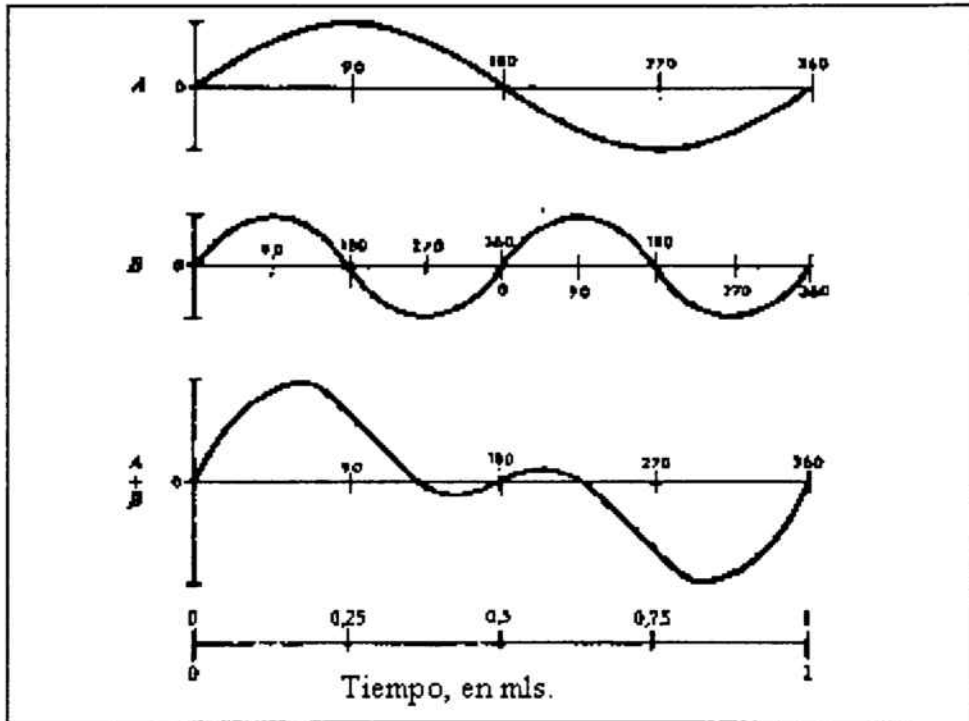


Figura 5 - Suma de dos ondas simples.

En nuestro ejemplo de la cuerda de guitarra, al presionar la cuerda producimos la vibración de la cuerda en toda su longitud, con el resultado de que escuchamos efectivamente la frecuencia para la que estaba afinada; pero, a la vez que vibra todo el cuerpo de la cuerda, vibra también cada una de sus partes: mitades, tercios, cuartos..., produciendo, cada una de ellas simultáneamente, una frecuencia distinta. Nuestra percepción del tono del sonido obedece a la frecuencia de la totalidad de la cuerda: es la *frecuencia*

fundamental (F_0) o tono; la demás frecuencias, llamadas frecuencias secundarias o *armónicos*, no las oímos independientemente, sino como un bloque, y nuestra percepción de ellas, en bloque, no es ya de "tono" o "altura musical", sino de *timbre*, de "calidad de sonido". Así, la diferencia entre una guitarra y un laúd, o un saxofón, o un piano, no es de tono o de intensidad, sino de "calidad", es decir, de timbre: cada instrumento tiene una voz particular, que obedece a una estructura de armónicos distinta.

Si colocamos la misma cuerda en una guitarra o en un laúd, la diferencia tímbrica entre los sonidos resultantes seguirá siendo muy grande, por lo que queda claro que la diferencia no reside en la misma fuente del sonido, sino en el *resonador* que se le asocia, que actúa a la vez como *filtro*.

La fuente de sonido de la guitarra emite una señal de una intensidad despreciable, inaudible, por lo que resulta necesario amplificar la señal para que resulte perceptible. La forma de hacerlo es asociar a la fuente de sonido, en este caso la cuerda, una caja de resonancia que esté "afinada" para vibrar con los sonidos que emite la cuerda, de tal modo que tales sonidos salgan de ella potenciados, amplificados, pues la superficie de la caja es mucho mayor que la de la cuerda.

Un resonador no es otra cosa que un cuerpo que vibra al entrar en contacto con una vibración determinada. El ejemplo clásico es el de los dos diapasones: si golpeamos uno, el otro, afinado en la misma frecuencia, vibrará inmediatamente; si, en cambio, no están afinados en la misma frecuencia, al golpear uno el otro no se inmutará. El otro ejemplo, muy cotidiano, es el de la vibración fantasma que a menudo aparece cuando tocamos una nota en el piano, o cuando suena algún sonido en el tocadiscos, o cuando pasa una motocicleta por la calle: el cristal de la ventana, por ejemplo, está afinado en esa misma

frecuencia. O la famosa anécdota del tenor Enrico Caruso, capaz de romper una copa de cristal con su voz: lo que hacía el tenor era golpear primero suavemente la copa, para escuchar en qué frecuencia estaba "afinada", y después emitir esa misma frecuencia con una gran potencia: la copa vibraba con intensidad hasta romperse, como si la hubieran golpeado con un mazo.

Este es el mismo principio aplicado a los instrumentos musicales, razón por la cual una guitarra debe emplear unas cuerdas de unas cualidades determinadas, o por la que la caja de resonancia debe tener esa forma precisa y ser de los materiales que debe ser: para conseguir la resonancia de la caja y el timbre característico del instrumento.

Pero, a la vez, un resonador es un filtro, pues es imposible que puedan amplificarse exactamente los mismos armónicos, excepto en el caso de dos cuerpos idénticos (como los dos diapasones): así, lo normal es que un resonador sólo amplifique aquellos armónicos producidos por la fuente del sonido que coincidan con los suyos propios. Los demás armónicos no potenciados desaparecen, simplemente porque no han sido amplificados y su intensidad era muy pequeña: son borrados por superposición por los armónicos potenciados; es decir, son filtrados.

Así, el timbre de un sonido es la estructura de los armónicos potenciados a su paso por un resonador. La relevancia lingüística del timbre es obvia: las diferencias entre los sonidos del lenguaje son únicamente de timbre, sin que importe el tono o la intensidad con que son emitidos. Podemos emitir el sonido [a], por ejemplo, variando la intensidad y la frecuencia, pero no dejaremos de percibir el mismo sonido; por otra parte, podemos emitir [a], [u], [i], sin variar la intensidad ni la frecuencia, a pesar de lo cual seguiremos percibiendo sonidos distintos: las diferencias que notamos entre ellos son exclusivamente tímbricas.

Es decir, es como si cada uno de los sonidos del lenguaje que somos capaces de emitir fuera un instrumento musical distinto (cada instrumento siempre tiene el mismo timbre, pues su caja de resonancia es invariable).

Por lo demás, cada persona tiene su propia calidad de voz, su timbre característico, que evita que la confundamos con otra: lo que no impide, sin embargo, que seamos capaces de reproducir y de reconocer el timbre específico de cada sonido del lenguaje, igual para todos los hablantes de la lengua.

1.1.3. La fonación y la articulación del sonido

El ser humano es capaz de emitir multitud de sonidos distintos con su cuerpo, algunos de los cuales son empleados en la comunicación oral y llamados *sonidos del lenguaje*. Nos referimos a la producción de la mayoría de tales sonidos con el término *fonación*.

La fonación consiste en un aprovechamiento brillante del sistema respiratorio en el momento de la espiración, de modo que la columna de aire que se expulsa se convierte en el principal motor en la producción de sonido. Por tanto, el llamado *aparato fonador* no es más que el *aparato respiratorio* en el momento de la espiración. No parece haber ningún órgano diseñado específicamente para la fonación en nuestro cuerpo, a excepción, tal vez, de los pliegues vocales: eso da una idea de lo reciente que es en nuestra propia naturaleza la capacidad de hablar.

Se distinguen tres cavidades en el aparato fonador (figura 6): las cavidades infragloticas (pulmones, bronquios y tráquea), la cavidad glótica (la laringe, donde se encuentran las cuerdas o pliegues vocales) y las cavidades supragloticas (faringe, cavidad bucal y cavidad nasal). Es decir, todo el aparato

fonador está organizado alrededor de la laringe, en cuyo centro, la glotis, se genera el sonido.

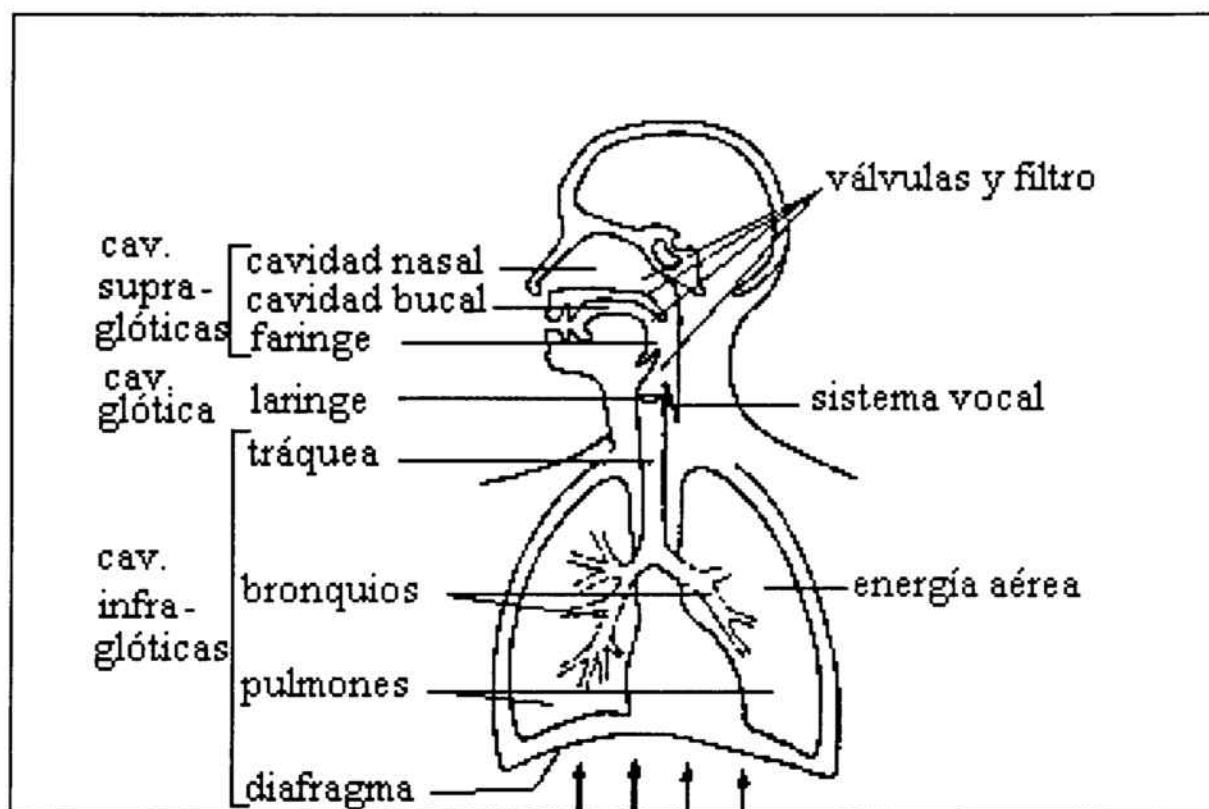


Figura 6 - El aparato fonador.

El músculo diafragma oprime a los pulmones para que expulsen el aire, que es conducido por los bronquios hasta la tráquea, donde se forma una columna de aire que se desplaza con una presión considerable. En la glotis, la columna de aire hace que los pliegues vocales vibren a su paso y se genere el sonido (el mecanismo de esta vibración la examinamos con cierto detalle en el punto 1.3.1.). El sonido creado en la glotis resuena en las cavidades supra-glóticas y sale al exterior.

En este sencillo esquema, repetido hasta la saciedad, conviene puntualizar, sin embargo, algunos aspectos. En primer lugar, hay que decir que en la glotis se crea un sonido que sería irreconocible como voz humana, que consiste en una especie de zumbido, como si se tratara de la boquilla de un instrumento de viento.

Otro aspecto interesante es la relación existente entre el aire y el sonido generado a su paso por la glotis. El aire, como hemos dicho, provoca la vibración de las cuerdas vocales que genera el sonido; pero el sonido y el aire que salen al exterior no son la misma cosa: en realidad, la salida de más aire del necesario puede dificultar una buena resonancia del sonido generado en la glotis, de modo que los profesionales de la voz (profesores, actores, cantantes) suelen aprender a impostarla, es decir, a conseguir una menor pérdida de aire en el momento de la fonación. El "chorro de voz" de un cantante de ópera, así, no es un "chorro de aire" (y por tanto poco tiene que ver con la capacidad pulmonar), sino literalmente un "chorro de sonido", casi sin aire, que ha podido resonar completamente por todas las cavidades. El ejemplo opuesto sería la voz "aterciopelada" de algunos cantantes de música ligera, o la voz "cascada" (como la del actor José Isbert): una voz en la que la salida de aire es muy superior a la necesaria, de modo que el sonido apenas puede resonar por las cavidades supraglóticas. Sobre este aspecto veremos más detalles en el punto 1.3.1.

El sonido generado en la glotis contiene todos sus armónicos, que ya en la primera cavidad supraglótica, la cavidad faríngea, que es el primer resonador, comienzan a filtrarse. En las cavidades supraglóticas el sonido resuena y, por tanto, se filtra, y el resultado es el timbre característico de cada persona, que depende, por tanto, no sólo de la estructura de sus cuerdas vocales (importante, por ejemplo, en el "cambio de voz" de los niños debido al crecimiento, y

también en las diferencias de voz masculina y femenina) sino a las características de tales cavidades.

Así, la morfología semejante de familiares cercanos, como hermanos o padres e hijos, que seguramente se parecen no sólo en el aspecto exterior sino también en las cavidades interiores, hace que su timbre sea muy parecido, porque sus glotis son semejantes, sus faringes, la estructura de sus cavidades bucales, sus cavidades nasales, y en general las dimensiones de la cabeza (toda ella una auténtica caja de resonancia de nuestra voz), la porosidad de los huesos, etc. Todas estas características determinan el timbre de los individuos, y por eso resulta casi imposible encontrar dos voces idénticas, porque eso supondría encontrar dos cuerpos idénticos.

La mayoría de los resonadores que filtran los armónicos del sonido que generamos en la glotis son invariables, fijos, como las cajas de resonancia de los instrumentos, y por eso nuestra voz característica es invariable a lo largo del tiempo (hasta que el envejecimiento cambia las dimensiones, la textura, etc., de nuestros resonadores); sin embargo, también contamos con un resonador móvil, flexible, variable, que nos permite cambiar a placer ciertos rasgos del timbre de los sonidos que generamos: la cavidad bucal.

La boca es, efecto, una caja de resonancia flexible, de modo que podemos cambiar su forma y sus dimensiones para lograr un timbre distinto. Así, la mayoría de los movimientos de la lengua (órgano móvil por excelencia de la cavidad bucal) suponen cambios tímbricos perceptibles (aunque no todos los movimientos son igualmente importantes: Stevens, 1972, demostró que hay movimientos articulatorios mínimos que producen un gran efecto acústico, mientras que otros movimientos apenas producen efectos perceptibles; cfr. Martínez Celdrán, 1994: 112).

Decíamos arriba que cada sonido del lenguaje distinto es como si se tratara de un instrumento distinto, porque consiste en un timbre característico. Así pues, nuestra voz se compone de una serie de rasgos de timbre personales, que caracterizan su personalidad individual e irrepetible, y que se deben a los resonadores fijos de nuestro aparato fonador; y de una serie de rasgos de timbre estándares, comunes a todos los hablantes, que caracterizan los sonidos del lenguaje que empleamos y que se producen en la cavidad bucal. A la producción de estos últimos rasgos tímbricos la denominamos *articulación* del sonido.

De este modo, la tradicional diferencia entre fonética articuladora y fonética acústica se nos aparece en sus justas dimensiones: cada movimiento articulatorio produce un cambio de timbre; los rasgos acústicos con los que caracterizamos los sonidos del lenguaje son rasgos de timbre debidos a movimientos en la cavidad de resonancia, con los cuales cambiamos continuamente el filtrado de algunos armónicos y la potenciación de otros.

En la figura 7 podemos ver el esquema de la posición de los órganos articulatorios en la producción de las vocales, junto a la estructura de armónicos potenciados en esa posición: como puede verse, cada movimiento de la lengua hace que los grupos de armónicos potenciados (llamados *formantes*) sean distintos. En el sonograma, el eje de ordenadas indica la altura en Hz de los formantes, y el eje de abscisas indica el tiempo transcurrido (en milésimas de segundos -ms-). En el espectro, las ordenadas indican la intensidad de los formantes (en dB) y las abscisas su altura en Hz: nótese que los armónicos no potenciados (filtrados) también se hallan presentes en el esquema (a ellos corresponden las depresiones en la curva), aunque no son relevantes para la percepción del timbre. En el caso de las vocales del español, los dos primeros formantes son suficientes para caracterizar su timbre.

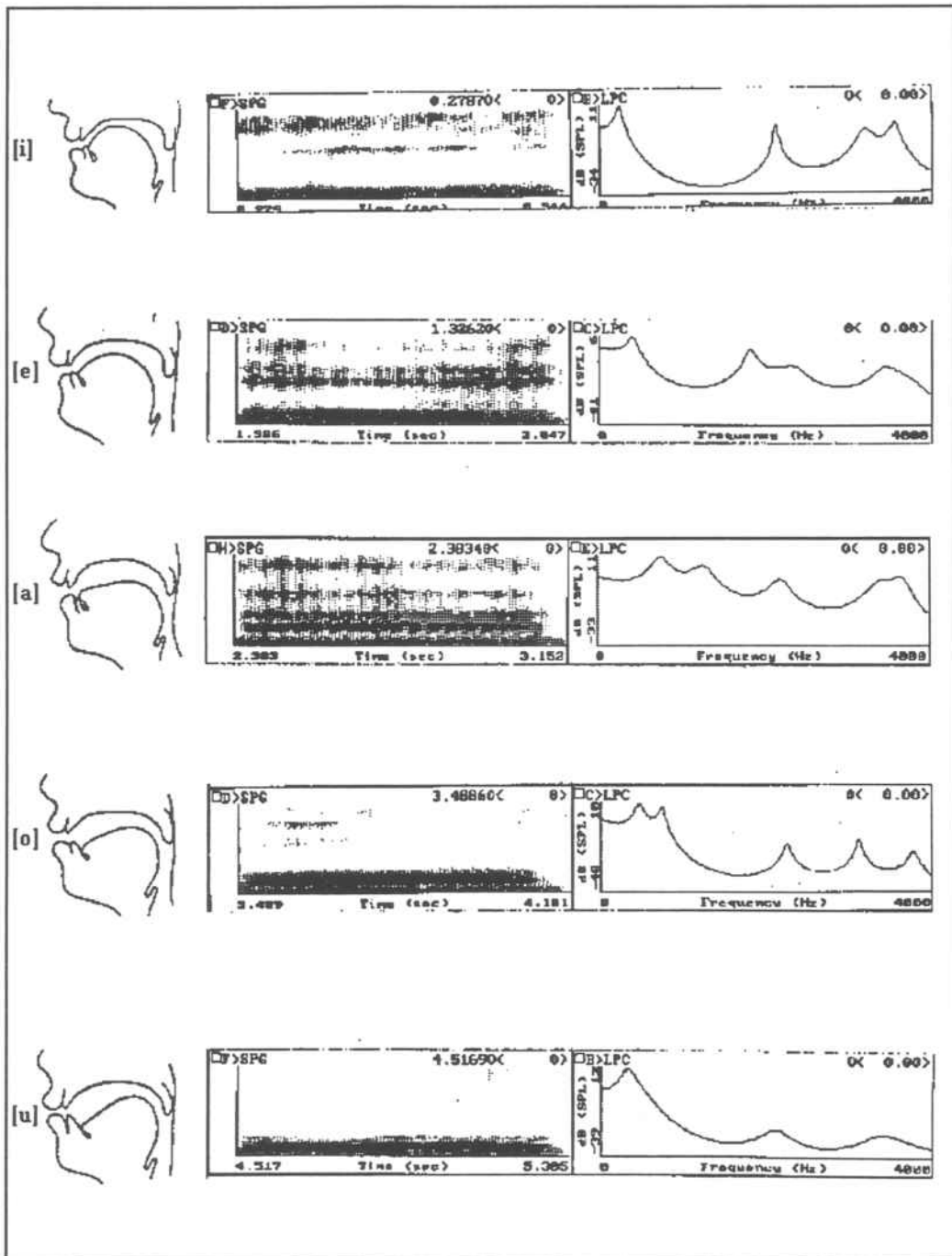


Figura 7 - Articulación de las vocales y estructura tímbrica.

La diferencia entre fonación y articulación, finalmente, se pone de manifiesto en la producción de algunos sonidos consonánticos (fricativos y oclusivos), que, por lo general, se producen directamente en la cavidad bucal, independientemente de que en la glotis se haya generado o no sonido (si es así, se tratará de consonantes sonoras). Así, mientras las vocales, por ejemplo, son producto de una fonación y una articulación posterior que moldea el sonido fonado, las fricativas y oclusivas son producto del rozamiento o choque del aire con los órganos articulatorios, y por tanto producto de una articulación, directamente, y no de una fonación, pues no hay sonido sino un ruido cuya fuente no es la vibración de las membranas vocales, sino la propia columna de aire a su paso por una constricción.

En realidad, también puede darse una articulación glótica, esto es, un ruido (oclusiva o fricativa) producido cuando el aire pasa por la propia glotis, en la zona no vibrante de la misma (y, por tanto, puede ser, simultáneamente, sonoro o sordo, esto es, con vibración de las membranas o sin ella): es el caso de las consonantes faríngeas y glotales (como la aspirada [h], característica de varios dialectos de nuestra lengua).

Así, podemos decir que la fonación es el sonido producido por la vibración de las cuerdas vocales, y que la articulación es el ruido producido por el aire a su paso por una constricción glótica o supraglótica (glotal, faríngea o bucal).

1.2. Cuestiones previas de Fonética y Fonología

1.2.1. Segmentos tímbricos y continuum sonoro

Suele creerse que las lenguas están compuestas por un número limitado de *sonidos* organizados en una serie de unidades funcionales llamadas *fonemas*; cada fonema, unidad funcional y por tanto abstracta, puede tener una o más realizaciones concretas, llamadas *alófonos* o *variantes*, que son sonidos distintos en sí mismos pero iguales en su rendimiento lingüístico, es decir, teóricamente intercambiables, aunque suelen ser complementarios: cada uno aparece en un contexto determinado.

El estudio de los sonidos (esto es, de los alófonos en cuanto sonidos) corresponde a la fonética, que suele ocuparse de su articulación (fonética articuladora) o de los rasgos acústicos que definen su timbre (fonética acústica). La percepción de los sonidos suele considerarse también dentro del ámbito de la fonética (fonética perceptiva), pero ya muy mediatizada por la fonología, esto es, por el rendimiento funcional de tales sonidos y por su identificación como tales por parte de los hablantes. Porque la fonología, como disciplina netamente teórica (frente a la fonética, teórica, experimental e instrumental), se ocupa precisamente de la identidad lingüística de los sonidos, de los rasgos mínimos que los caracterizan en una lengua determinada, de su clasificación, de las relaciones que mantienen, etc.; es decir, de lo que el hablante percibe, identifica y combina conscientemente, al margen de las características físicas de su timbre o fisiológicas de su producción: al margen del sonido en sí.

Considerando, pues, que cada segmento tímbrico tiene una identidad física reconocible, su identidad fonológica no puede ofrecer muchas dudas. Al estudio y clasificación de las unidades fonológicas siguiendo estos parámetros se le llama *fonología segmental* o *fonología de la palabra* (pues el rendimiento

funcional de los sonidos se determina sólo en el contexto de la palabra, de los sonidos que forman una palabra); frente a ella, la *fonología suprasegmental* o *prosodia* se ocuparía del estudio de aquellos fenómenos igualmente fonológicos pero que afectan a más de un segmento a la vez (a más de un sonido), como el acento, el ritmo o la entonación. Completaría el panorama la fonética (articulatoria, acústica o perceptiva), encargada sólo de suministrar los datos que debe manejar e interpretar la fonología, y de someter a pruebas experimentales (generalmente, bajo condiciones controladas) sus predicciones.

Este sería un esquema muy sencillo de la perspectiva tradicional, estructuralista, del componente fónico de la lengua. En esencia, es también la perspectiva que subyace en la mayor parte de los trabajos actuales sobre fonética y fonología, y la concepción que se transmite en los manuales de gramática, de lingüística general, etc.; es la perspectiva que está detrás de buena parte del trabajo dedicado a desarrollar sintetizadores de voz y reconocedores del habla; es la perspectiva en la que se basan todos los métodos de enseñanza de lenguas que se ocupan del componente fónico; y es también la que emplean en su trabajo de reeducación los especialistas en el tratamiento de los trastornos de audición y de habla.

Es decir, los repetidos intentos de fijar nuevas perspectivas en la concepción del componente fónico de la lengua (por ejemplo, los distintos enfoques de fonología no lineal) no han logrado traspasar aún la barrera de los círculos especializados, generalmente porque provienen del ámbito de la lingüística generativa, cuya concepción marcadamente teórica de la disciplina, casi algebraica, el empleo de un aparato formal muy desarrollado y su aparente desprecio por la aplicación inmediata y por los métodos experimentales la han alejado de sus posibles usuarios no especialistas (sobre el cambio de paradigma teórico impulsado por la fonología generativa, v. el apartado 2.3.2.).

La perspectiva tradicional, por lo demás, ha dado buenos resultados, o al

menos resultados aceptables, en los ámbitos de aplicación práctica en los que se han empleado los conocimientos que ha generado. Con todo, dicha perspectiva contiene algunos errores de enfoque obvios desde el punto de vista fonético, y que nos interesan particularmente porque atañen al carácter segmental de los sonidos.

Como decíamos al principio de este apartado, suele creerse que las lenguas están compuestas por un número limitado de sonidos. Cada sonido, como sabemos, es un segmento de la cadena hablada caracterizado por unos rasgos tímbricos determinados. Incluso el concepto de *cadena hablada* implica el encadenamiento de unidades (segmentos) bien definidas. Sin embargo, la realidad acústica es otra.

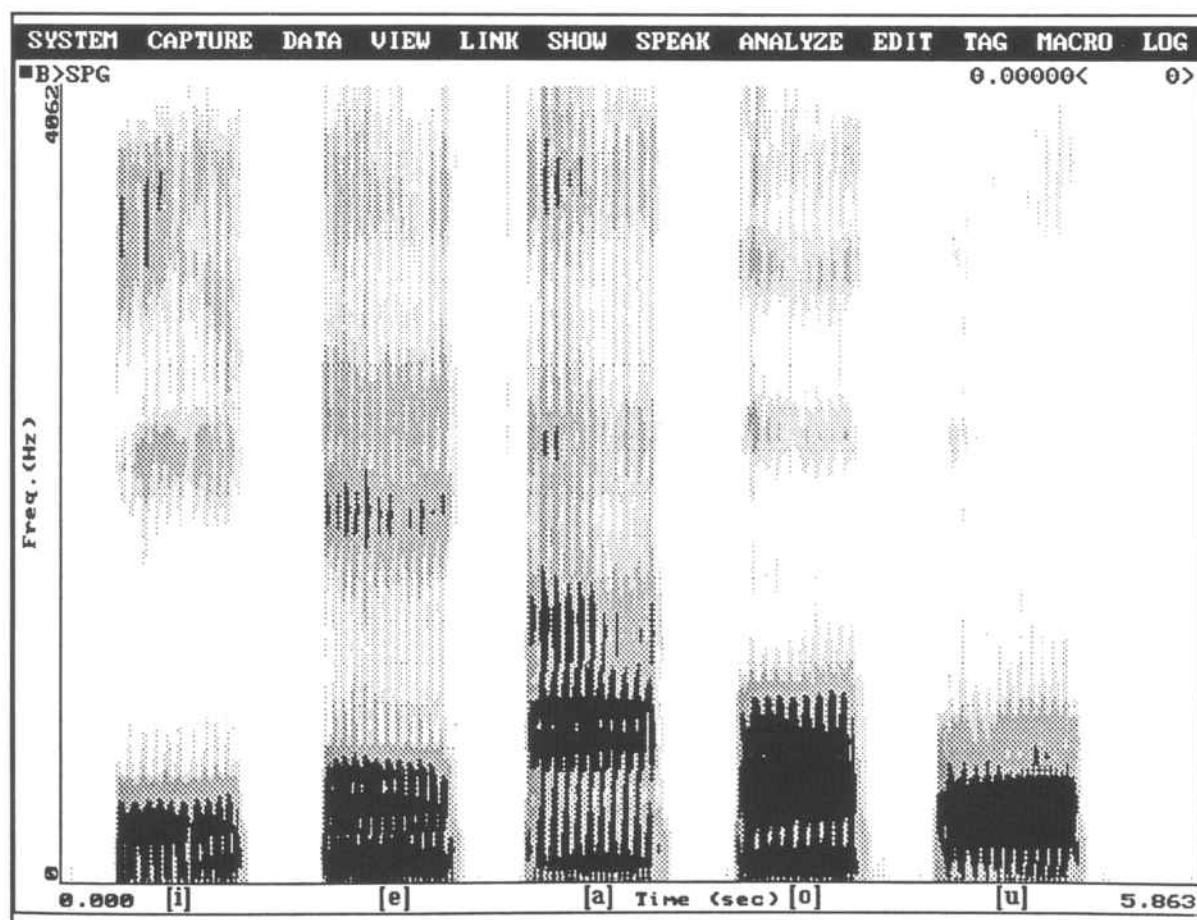


Figura 8 - Sonograma de las cinco vocales del castellano.

Veámos en el apartado anterior (figura 7) el sonograma de las cinco vocales del castellano, que reproducimos aisladamente en la figura 8. En ellos pueden verse con toda claridad el timbre (la estructura de los formantes) de cada una de ellas.

Sin embargo, difícilmente encontraremos en la realidad ningún sonido aisladamente. En la figura 9 podemos ver el sonograma de una cadena hablada en la que aparecen las cinco vocales, algunas de ellas repetidas: "los sonidos se solapan".

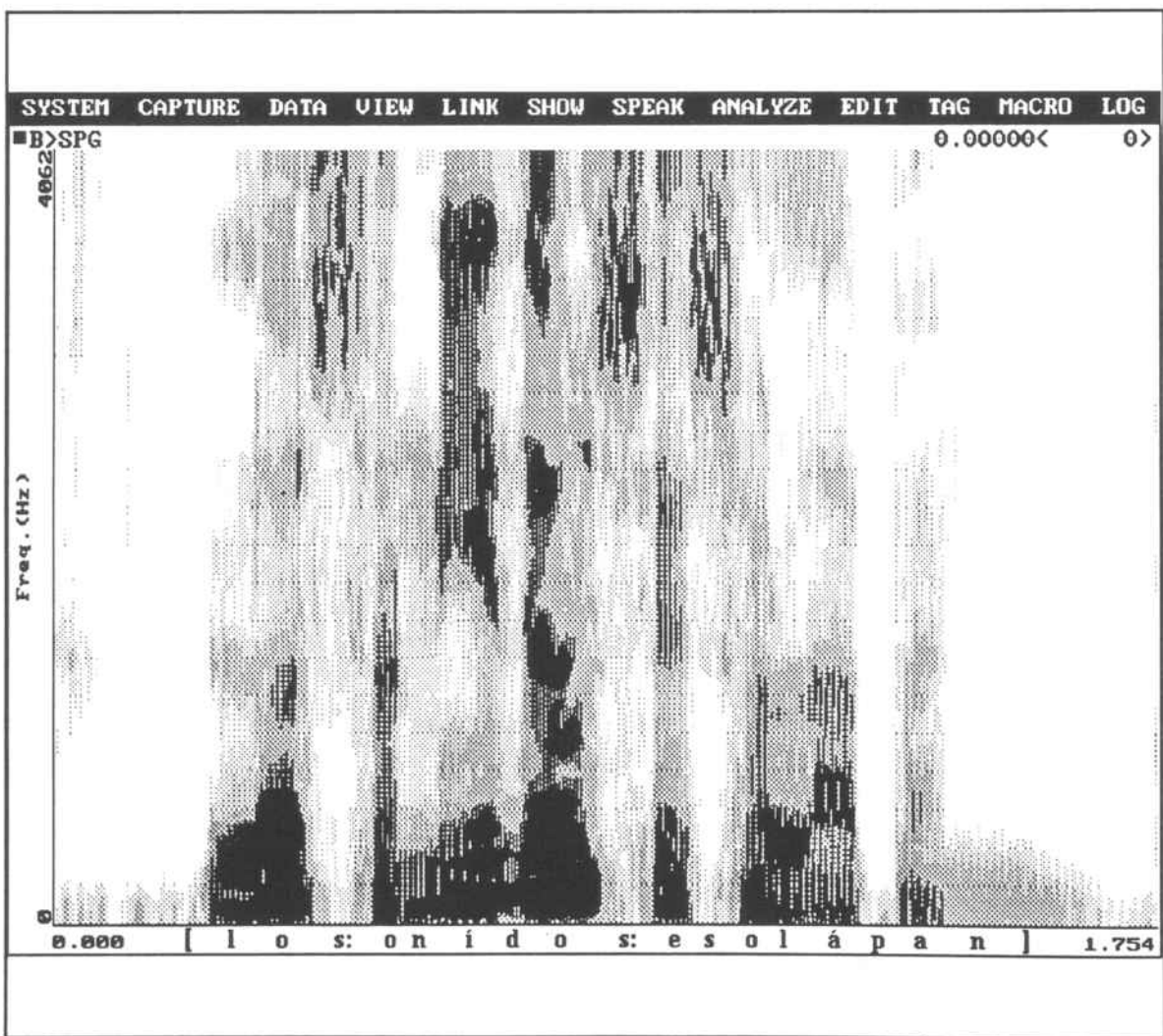


Figura 9 - Sonograma del enunciado: "Los sonidos se solapan".

Como puede comprobarse, ninguna de las vocales se parece a las que veíamos aisladamente, ni a las mismas que aparecen repetidas en la misma cadena. A eso nos referimos cuando hablamos de "variantes contextuales": por ejemplo, las cuatro vocales [o] que aparecen en la figura 9 apenas se parecen entre sí; las dos [a] de "solapan" realmente parecen sonidos distintos; etc. Es decir, en realidad cada segmento tímbrico reconocible en la cadena fónica es en realidad un sonido distinto. Por lo demás, apenas pueden reconocerse los límites precisos entre los distintos segmentos.

Este es uno de los fenómenos que más sorprende a los lingüistas que trabajan en fonética instrumental por primera vez: que no hay límites en la realidad, que la cadena hablada no es una "cadena" de segmentos directamente identificables sino un verdadero *continuum sonoro*, que el timbre de cada segmento no es estable prácticamente en ningún momento y que el continuum sonoro lo es de rasgos tímbricos que atraviesan longitudinalmente a varios supuestos segmentos.

Así, la fonología segmental y la misma fonética tradicional (especialmente la fonética articulatoria), parte de una concepción de la realidad difícilmente comprobable: la misma existencia de segmentos tímbricos estables. Por eso en numerosas lenguas las disputas entre fonólogos sobre cuántas unidades fonológicas debe interpretarse que hay y cuáles son, siguen siendo vigentes indefinidamente (cuántas vocales hay en portugués, las africadas del catalán son un fonema o dos, las glides del español son fonemas o alófonos vocálicos, etc.).

La afirmación de que una lengua está constituida por un número limitado de sonidos es, por tanto, falsa, pues cada realización de un sonido es físicamente distinta a otra, y muestra que en fonética se trabaja, en la práctica, con un nivel intermedio de abstracción, entre el fonema y el sonido real: el "sonido tipo" aislado, el *fono*, caracterizado con unos valores tímbricos estables, generales para todos los hablantes, abstractos. En este sentido, dice Coseriu: "la fonética

práctica abarca normalmente los planos del «alofón» y del «fono» (sonido concreto y sonido abstracto) -pero inclinándose más bien hacia éste-" (1962: 223).

1.2.2. Carácter lecto-escritor de la fonología: la norma fónica

Para muchos estudiantes de "fonética y fonología", sin embargo, el análisis de un sonograma sigue siendo terreno vetado, demasiado especializado para los objetivos del curso, y el único ejercicio práctico suele reducirse a la transcripción fonética (en contadas ocasiones, también se practica la transcripción fonológica). Y es que subyace en la tradición fonética un marcado carácter lecto-escritor: a cada segmento tímbrico, claramente definido y descrito por la fonética, corresponde un símbolo escrito perteneciente a un alfabeto especial, el alfabeto fonético. Los sonidos son realidades sonoras, auditivas, que pertenecen radicalmente, exclusivamente, a la lengua oral, en efecto, pero su estudio sigue haciéndose por medio de la escritura, con alfabeto especial o sin él: pues, además, raramente se efectúa una genuina transcripción fonética de una grabación, sino que suelen realizarse directamente de un texto escrito. Así, el alfabeto fonético pierde todo su sentido originario y su objetivo inicial se ve claramente pervertido.

Pero es que, como hemos dicho, en la tradición fonética subyace un marcado carácter lecto-escritor. Desde el principio, antes de los trabajos fundacionales de la fonología moderna, los fonetistas habían optado por la descripción de los sonidos correspondientes a "las letras", y por su fijación para lograr "una pronunciación correcta": esto era lo más razonable, porque aún no se manejaban los conceptos generales, de planteamiento, que nos permiten hoy distinguir entre sonido, fonema y grafía.

Fundada la fonología como una disciplina descriptiva, sin embargo, los

métodos siguieron siendo muy parecidos, pero sobre todo los resultados: el número de unidades, de variantes, etc. Esa es la razón por la que sólo contamos con una buena descripción fonética y fonológica, completa y generalmente aceptada, en aquellas lenguas que cumplen las siguientes tres condiciones históricas: en primer lugar, ser lenguas de escritura alfabética, en las que el concepto de "segmento" es fácil de intuir y aceptar; en segundo lugar, contar con una tradición fonética anterior a los trabajos de Trubetzkoy, tradición que ya hubiera fijado y descrito los sonidos de la lengua correspondientes a la letras o grupos de letras; y, en tercer lugar, que existiera una norma gramatical perfectamente sancionada y, por tanto, se mostrara una preocupación sostenida por la pronunciación correcta. Este es el caso de algunas lenguas de nuestro entorno, como el alemán, el francés o el inglés; es el caso, igualmente, del español, tal vez una de las lenguas mejor estudiadas; también, con algunas décadas de retraso, es el caso del catalán; no, por ejemplo, del portugués o del italiano. Pero veamos el caso del español.

En el español peninsular puede hablarse de una norma fonética bien definida, cuya mejor representación sigue siendo la voz de los telediarios, la voz *en off* de los anuncios y el doblaje de las películas. Dicha norma fonética, descrita en los apartados de "ortología" o "pronunciación correcta" de las gramáticas académicas y pseudoacadémicas, cayó en desgracia durante la transición democrática, porque fue identificada con el régimen anterior y su política lingüística de negación de otras lenguas y dialectos. Este fenómeno, inédito en toda Europa, ha propiciado un prestigio creciente de todo lo dialectal en nuestro país (lo dialectal se identifica con lo democrático: todos los líderes políticos hablan un dialecto bien definido) y una tolerancia lingüística tan excepcional que ya empieza incluso a decaer.

Sin embargo, el origen de la norma fonética es muy anterior y ajeno a la dictadura: su fuente primordial es el *Manual de pronunciación española* de

Tomás Navarro Tomás, publicado en 1918, reeditado ininterrumpidamente desde entonces, y que sirve aún de base o punto de partida a toda descripción o investigación fonética del castellano.

Hablando de la "pronunciación correcta española", dice Navarro Tomás en su libro (1944: 8):

Señálase como norma general de buena pronunciación, la que se usa corrientemente en Castilla en la conversación de las personas ilustradas, por ser la que más se aproxima a la escritura

Puesto que la lengua estándar -la norma léxica y la norma gramatical-sancionada por la academia tiene como principal objeto la escritura, también la norma fonética había de regirse por ella, único rasero cultural de la época -en las mismas páginas, se llama sistemáticamente "vulgar" a lo "dialectal"-:

Esta pronunciación, pues, castellana sin vulgarismo y culta sin afectación, estudiada especialmente en el ambiente universitario madrileño, es la que en el presente libro se pretende describir.

El "ambiente universitario madrileño" ha cambiado mucho, desde luego, desde 1918; sin embargo, en la actualidad su descripción sigue considerándose plenamente vigente (y es que la lengua escrita y, en general, la norma lingüística apenas ha cambiado desde entonces).

Con todo, el objetivo de Navarro Tomás era fijar una norma "castellana" y "culta", exclusivamente, que sirviera como modelo de aprendizaje, por ejemplo, para los extranjeros o para los actores. No pretendía describir todo el español, sino que marcó unos límites bien definidos a su estudio, acotando incluso la variedad dialectal y la variedad de registro consideradas. En este

sentido, la modernidad de Navarro Tomás es admirable.

En la actualidad, sin embargo, y gracias a una forma de entender el concepto saussureano de "lengua" como algo que todo lo iguala, el objetivo de las descripciones parece ser, más bien, no hacer distinciones dialectales, diastráticas ni de registro, y tomar lo que fue una propuesta muy ponderada de norma culta por norma lingüística y general. Y esto supone algunos problemas de enfoque.

En 1950 aparece el libro de cabecera de la fonología española, titulado, precisamente, *Fonología española*, de Emilio Alarcos Llorach, donde dice el autor (1950: 143):

...vamos a intentar una descripción fonológica del español actual, ateniéndonos exclusivamente al sistema del lenguaje corriente libre de dialectalismos y vulgarismos, así como de afectaciones literarias y académicas. Se trata del mismo estilo de español estudiado fonéticamente por Navarro Tomás

Obsérvese que cuando Navarro Tomás habla sólo de "vulgarismos", Alarcos distingue, con precisión, entre "dialectalismos" y "vulgarismos"; pero que cuando Navarro Tomás habla de "pronunciación castellana y culta", acotando bien los límites de su estudio, como hemos visto, Alarcos habla de "español actual" y de "lenguaje corriente", generalizando entonces la descripción y eliminando todos los límites.

Teniendo en cuenta que los datos fonéticos interpretados fonológicamente por Alarcos *coinciden* en su mayor parte con la descripción de Navarro Tomás, resulta obvio que la fonología española nace viciada ya desde un principio, a expensas de la lengua escrita, de la variedad diastrática culta y del registro literario; y no como resultado de una investigación fonética original y bien

acotada, con datos exclusivamente de lengua oral que se organizaran después fonológicamente.

Paralelamente, la propia descripción fonética de la lengua, cuyas realizaciones concretas son enormemente cambiantes, encuadradas en todas las variedades lingüísticas (dialectales, de registro, etc.), algunas de las cuales encuentran en las diferencias fónicas, precisamente, su personalidad, sólo se preocupa del nivel más abstracto, del más general, del más parecido a la lengua escrita, muy cerca siempre de lo prescriptivo. El riesgo es, obviamente, describir una lengua que nadie o muy pocos hablan: sólo en algún estudio muy concreto de dialectología encontraremos, pues, la descripción de algunos de los sonidos más extendidos de nuestra lengua (como las fricativas palatales [j] y [ç], la aspirada glotal [h], la africada sonora [dç], etc.) o algunos fenómenos generalizados, pero no aceptados aún *oficialmente* como "norma" (la reducción vocálica, el yeísmo, el comportamiento fonológico de la "wau" [ɔ], etc.). En este sentido, la revisión llevada a cabo por E. Martínez Celdrán (1989) del sistema fonológico del español propuesto por Alarcos es un paso decisivo en la normalización de la fonología española como una disciplina puramente *descriptiva*, dependiente sólo de los datos experimentales y alejada de toda tentación *prescriptiva*.

1.2.3. *Fonética y fonología de la "cantidad"*

Según Quilis (1981: 75), "un estímulo acústico cualquiera comprende cuatro elementos físicos constitutivos: cantidad, intensidad, frecuencia del armónico fundamental y estructura formántica de las ondas sonoras". Es decir, a las tres características del sonido que vimos en el apartado anterior (intensidad, tono y timbre) añade el parámetro de "cantidad" o "duración" del sonido. La mayoría de autores, en efecto, cuentan la cantidad como un parámetro acústico

del sonido (medido en milésimas de segundos -ms-), y como un elemento de primera magnitud para su percepción; igualmente, la duración es un parámetro fundamental para caracterizar fenómenos suprasegmentales y contrastivos como el acento o la entonación, e incluso en numerosas lenguas es un rasgo distintivo en el nivel segmental.

Evidentemente, la "cantidad de señal sonora" es un parámetro en sí mismo acústico, es decir, fonético, aunque la información que ofrece, como tal, es escasa: nos referimos, por ejemplo, a la duración total de la señal emitida, o a la duración de un rasgo tímbrico determinado. En este último caso, al referirnos a la duración de un rasgo de timbre determinado, debemos tener en cuenta que se trata de un parámetro, digámoslo así, *de segundo grado*, pues lo que nos interesa realmente es medir un rasgo tímbrico, y la duración no es más que una característica de ese rasgo. No es, por tanto, una característica del sonido que esté al mismo nivel que el tono, la intensidad y el timbre, sino que está siempre referida a estas, como un instrumento para caracterizarlas a su vez.: es una característica de las características del sonido, es un parámetro para medir los parámetros del sonido.

La cantidad, sin embargo, a menudo constituye un rasgo fundamental en la caracterización de un fenómeno fonológico, como sucede con el acento. En tales ocasiones, la tendencia es a tratar la duración como un parámetro acústico más, lo cual hemos visto que no es en absoluto correcto: es más, la duración debería considerarse, directamente, como un rasgo fonológico y dependiente sólo del nivel fonológico.

Como ya vimos arriba, no puede hablarse de segmentos de timbre claramente definidos, dentro del continuum sonoro, sino de una serie de rasgos de timbre solapados y cambiantes que nosotros podemos concebir como segmentos sólo si hacemos un notable esfuerzo de abstracción (una abstracción intermedia entre el casi indistinto sonido real y el fonema, como decíamos).

Entonces, la duración de un sonido estable (en realidad, de un segmento, aunque para medir un segmento habitualmente medimos sólo un rasgo tímbrico mínimamente estable y considerado suficientemente definitorio), una vez más, aparece como un parámetro *de segundo grado*, pues es una característica física, en efecto, pero de una unidad (un segmento aislado) en cierto modo abstracta: es decir, no sólo no se trata de un rasgo acústico comparable a los demás, sino que la propia medición física sólo puede hacerse después del proceso de abstracción y referida sólo al resultado de dicho proceso de abstracción.

Cuando más adelante encontremos definiciones de acento y entonación en las que el parámetro de cantidad juegue un papel importante, deberemos tener en cuenta este hecho: como veremos, el carácter *de segundo grado* que el propio fenómeno de acento tiene ya de por sí (pues se refiere al contraste entre los sonidos vocálicos definidos del continuum sonoro) es coherente con el de la duración, que entonces puede ser un parámetro adecuado para definir el fenómeno; con el fenómeno de la entonación, sin embargo, seguramente habrá que afinar más.

1.3. Aspectos fonéticos de la entonación

1.3.1. Fisiología de la producción tonal

La laringe es una estructura cartilaginosa situada al final de la tráquea y conectada a un complicado sistema de músculos y ligamentos. Los elementos más importantes de la laringe son el cartílago tiroides, el cartílago cricoides, los dos cartílagos aritenoides y las cuerdas o pliegues vocales (v. figura 10).

El cartílago cricoides está conectado al final de la tráquea y es a su borde frontal, que acaba en un estrechamiento, a lo que llamamos "nuez" o "bocado de Adán". Los aritenoides son dos cartílagos más pequeños, conectados al cartílago cricoides por unos ligamentos flexibles que les permiten cierta movilidad; en el extremo opuesto, por delante, también está conectado al cartílago cricoides el cartílago tiroides.

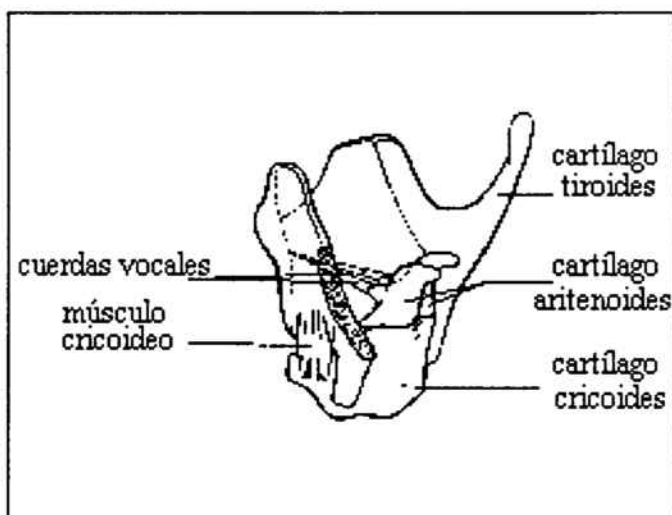


Figura 10 - Esquema de la laringe.

Las cuerdas vocales son dos membranas de tejido muscular que se extienden desde el cartílago tiroides a los dos cartílagos aritenoides, a cada uno de los cuales está enganchada una membrana. A la abertura que hay entre las cuerdas vocales la llamamos glotis, y suele ser más ancha cerca de los aritenoides (por donde pasa la mayor parte del aire durante la respiración) y más estrecha en el extremo opuesto, el frontal, más cerca del tiroides (porción cuyos movimientos son generalmente los responsables de la fonación).

Los músculos más importantes conectados a la laringe son los músculos cricotiroides (cuya función es tensar las cuerdas vocales, tirando del cartílago cricoides, que a su vez tirará de los aritenoides), los interaritenoides (que sirven para juntar los cartílagos aritenoides) y los cricoaritenoides (que sirven para separarlos).

Todo este juego de músculos, ligamentos y cartílagos tiene como función principal ajustar la tensión de las cuerdas vocales y aumentar o disminuir la abertura glotal. Las cuerdas vocales, por tanto, no tienen movimiento propio, y su posición se ajusta sólo en función del estiramiento que provoquen los cartílagos aritenoides, a los que están sujetas; estos cartílagos, a su vez, tampoco tienen movimiento propio, sino que dependen, como hemos visto, de los músculos laríngeos.

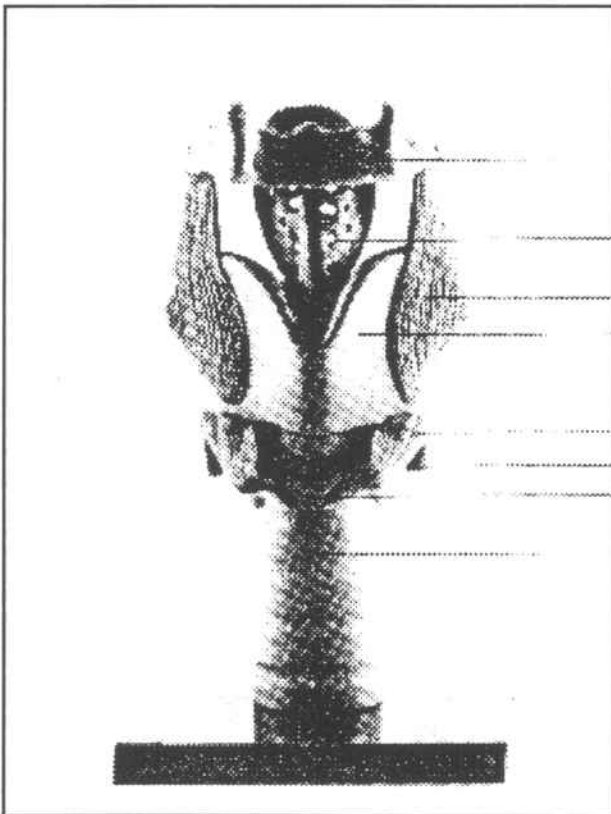


Figura 11 - Vista frontal de la laringe.

Como dijimos en el apartado 1.1.3., la fonación se produce cuando la columna de aire reunida en la tráquea desde los pulmones pasa por la glotis y provoca la vibración de las cuerdas vocales.

En condiciones de respiración normal, la abertura de la glotis es lo bastante ancha como para que el aire pase sin provocar ninguna vibración. Para que las cuerdas vocales vibren, la abertura glotal ha de ser lo

suficientemente estrecha y el flujo de aire, causado por la diferencia de presión

que hay entre los pulmones y las cavidades supraglóticas, ha de ser lo suficientemente grande. La presión infraglótica es provocada, como dijimos, por los músculos respiratorios (fundamentalmente, el diafragma); el estrechamiento de las cuerdas vocales, por su parte, es provocado por la contracción de los músculos interaritenoides, que hace que los cartílagos aritenoides se junten.

La proximidad de las cuerdas vocales hace que estas opongan resistencia

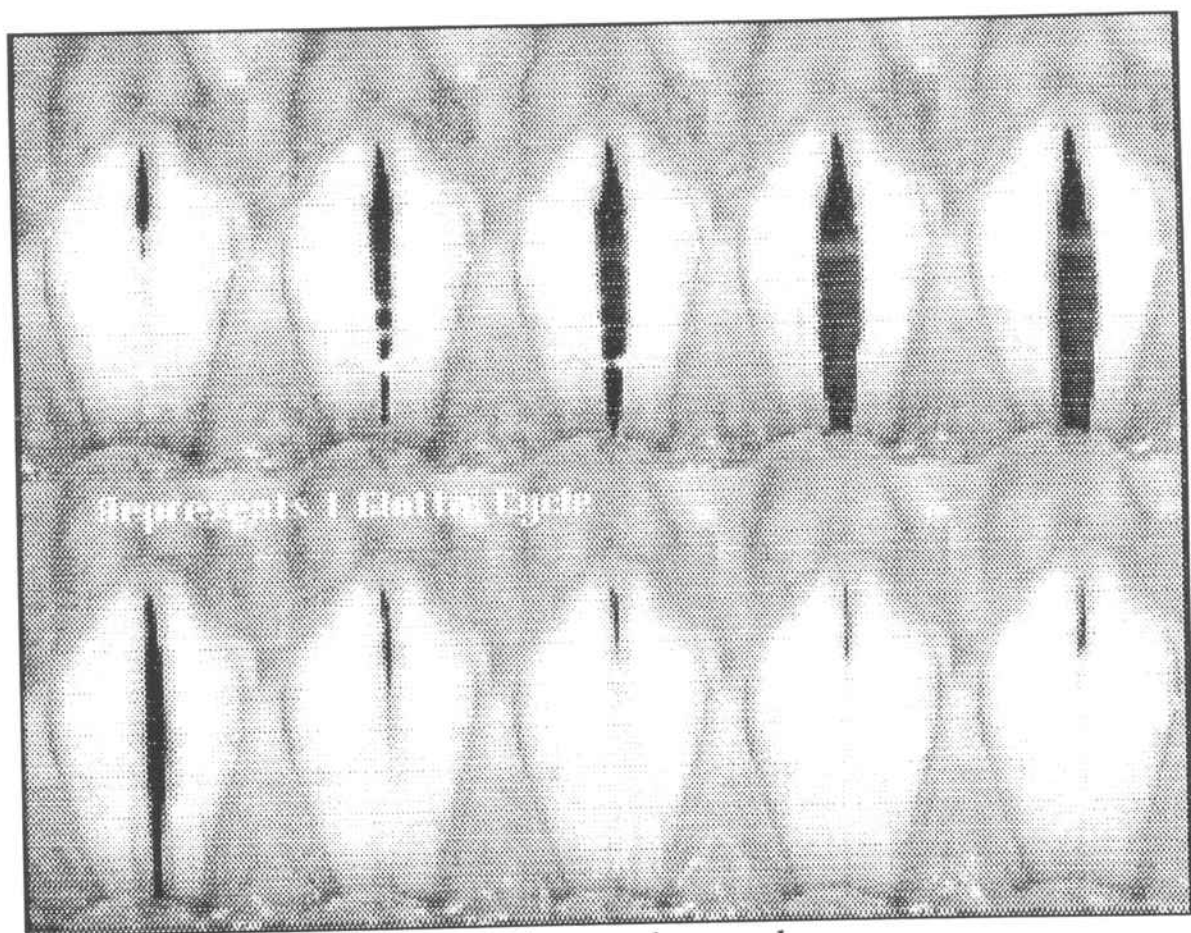


Figura 12 - Ciclo vibratorio de las cuerdas vocales.

al paso del aire, con lo cual la presión infraglótica se ve fuertemente aumentada. Cuando dicha presión es lo suficientemente alta, la corriente de aire aparta a las cuerdas vocales, sobreponiéndose a la fuerza muscular que las mantenía juntas: la abertura glotal, entonces, deja pasar el flujo de aire libremente. Las cuerdas vocales, replegadas sobre sí mismas, tienden a volver a la posición inicial: pero,

en cuanto la abertura glotal se estrecha un poco, la velocidad de la corriente de aire aumenta. Entonces tiene lugar el llamado *efecto Bernoulli*, que provoca un descenso de presión en el espacio que media entre las cuerdas vocales, a causa del cual se produce una succión que cierra bruscamente la glotis y hace que las cuerdas vocales se junten. Una vez se ha cerrado la glotis, cesa el efecto Bernoulli, las cuerdas vocales vuelven a su posición inicial y comienza un nuevo ciclo vibratorio. El sonido consiste en las variaciones de presión producidas por el paso alternativo del aire proveniente de la tráquea, interrumpido por los cierres de la glotis. Así, las cuerdas vocales actúan como verdaderas *válvulas* que regulan y fragmentan la salida del aire, en cuyo proceso se genera el sonido.

En la figura 12 puede verse una serie de fotografías de las cuerdas vocales durante el ciclo vibratorio.

A la explicación de los movimientos vibratorios de las cuerdas vocales se le ha llamado *teoría mioelástica-aerodinámica de la fonación* (Lieberman, 1967), en contraste con la llamada *teoría neurocronámica* (Husson, 1962), según la cual una serie de impulsos nerviosos provocaría directamente la vibración de las cuerdas vocales por medio de los músculos laríngeos. Curiosamente, y a pesar de que hace ya dos décadas que no se concede apenas validez a la teoría neurocronámica, sigue hablándose hoy de "teoría mioelástica", en vez de referirnos simplemente al "mecanismo de la fonación".

La fonación, en fin, puede definirse exactamente como la repetición cíclica de los movimientos de abertura y cierre de la glotis, como efecto de la fuerza aerodinámica de la corriente de aire proveniente de los pulmones (que fuerza su paso separando las cuerdas vocales), y de la fuerza aerostática provocada por el descenso de presión entre las cuerdas vocales cuando estas tienden a recuperar su posición, concretada en el efecto Bernoulli, que hace que se cierre la glotis y se genere el sonido.

Una reproducción curiosa y bastante aproximada de todo este movimiento

vibratorio podemos conseguirla si juntamos dos hojas de papel y soplamos con fuerza por entre ellas: las hojas se separarán y volverán a juntarse una y otra vez, como hacen las cuerdas vocales (con la única diferencia de que las hojas de papel no participan activamente en la vibración, y las cuerdas vocales sí, con su continua tendencia a volver a su posición inicial, que es justamente lo que provoca que actúe sobre ellas el efecto Bernoulli, pues de otro modo se encontrarían demasiado separadas, como en la respiración normal).

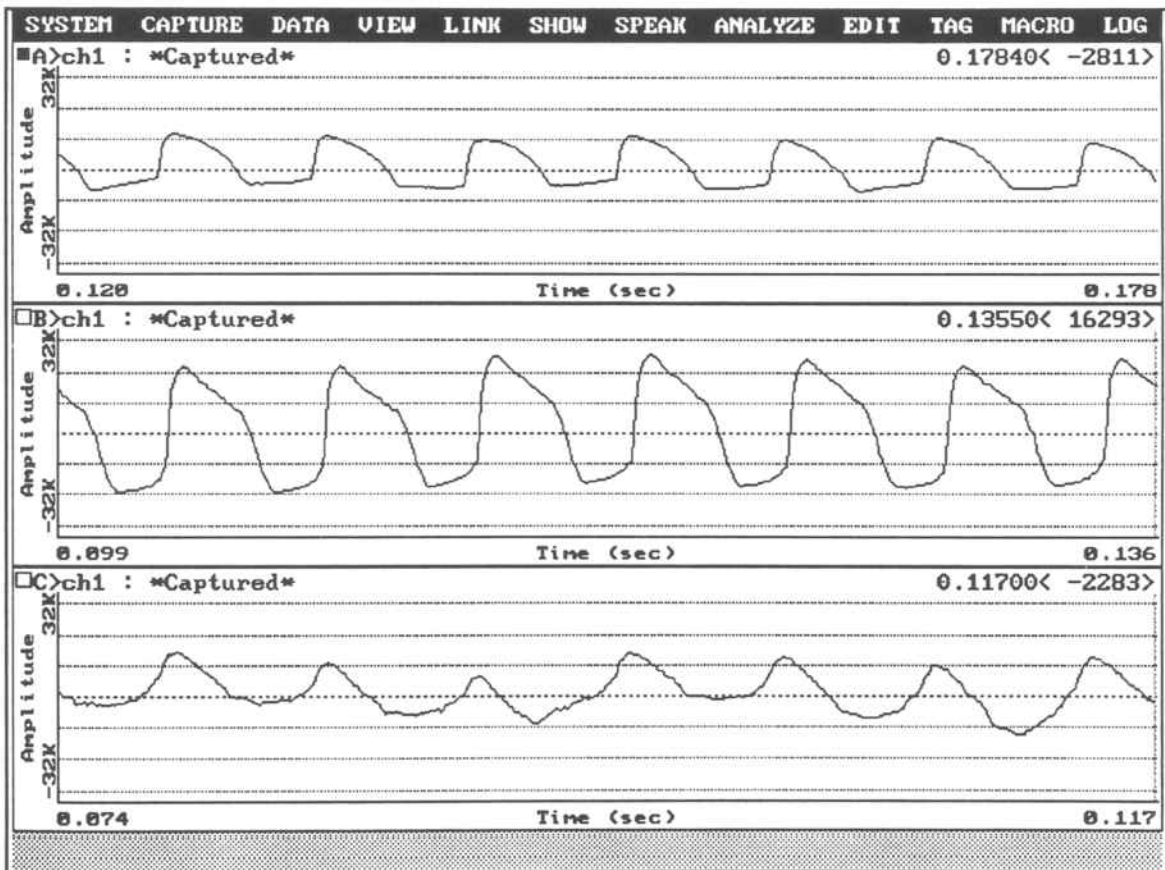


Figura 13 - Laringografías de voz normal (A), voz impostada (B) y voz "cascada" (C).

Las laringografías de la figura 13 muestran cómo los pliegues vocales se juntan y se separan, en tiempo real: cada pico del esquema corresponde a un contacto de las cuerdas vocales, es decir, a un cierre total de la glotis (llamado también *pulso glotal*), y cada depresión en el esquema corresponde a la

separación de las cuerdas, es decir, a la máxima abertura glotal, que es cuando sale el aire. La figura 13a muestra una laringografía de una fonación normal. La figura 13b, en cambio, muestra una fonación impostada: obsérvese cómo el contacto de las cuerdas vocales se ha alargado, para facilitar la buena resonancia del sonido por todas las cavidades supraglóticas, y el tiempo de abertura glotal es menor, de modo que la salida de aire sea muy pequeña, no entorpezca la resonancia del sonido creado, el gasto de aire sea mínimo y pueda prolongarse la fonación. En la figura 13c, finalmente, vemos el esquema de la fonación de una voz "cascada": el pulso glotal es mínimo, y la abertura muy prolongada, con lo que el sonido creado apenas puede resonar limpiamente antes de ser arrastrado por el propio aire, que se agotará inmediatamente, razón por la cual las personas con este tipo de voz se fatigan fácilmente y hablan incluso con dificultad.

La frecuencia fundamental del sonido emitido en la fonación dependerá, en todo caso, del número de vibraciones por segundo de las cuerdas vocales. Según Lieberman (1967), es la presión infraglótica, fundamentalmente, la causante de un aumento de la frecuencia fundamental, fenómeno que evidencia la estrecha relación que mantienen el tono y la intensidad. Sin embargo, hay otros elementos que inciden en los cambios de frecuencia, tanto en lo que se refiere al tono medio del individuo como al propio mecanismo de cambio de frecuencia durante la fonación.

En primer lugar, el grosor de las cuerdas vocales, su masa en la parte vibratoria: a mayor grosor, la vibración será más lenta y la frecuencia fundamental será menor. El tono medio de una persona depende en gran parte de este fenómeno, así como el que los hombres tengan un tono de voz más grave que las mujeres, o el cambio de voz que se produce en los adolescentes, justo en la época en que aumenta la masa de sus pliegues vocales.

Por otra parte, el elemento fundamental para controlar las variaciones

tonales durante la fonación es la tensión de las cuerdas vocales, que podemos controlar, como hemos visto, mediante los músculos laríngeos: una mayor tensión de las cuerdas provocará un movimiento de recuperación de su posición inicial más rápido y por tanto la actuación del efecto Bernoulli, así como un restablecimiento de tal posición tras el cierre glotal también más rápido, con lo que todo el proceso vibratorio se acelerará considerablemente, obteniendo como resultado un aumento de la frecuencia. La precisión con que somos capaces de controlar la frecuencia durante la fonación es, curiosamente, muy grande, teniendo en cuenta que nosotros no controlamos realmente las cuerdas vocales, sino un complejo sistema de, digámoslo así, "palancas" (músculos, cartílagos y ligamentos).

Otros elementos son también importantes en el control del tono, tales como el acoplamiento entre la frecuencia vibratoria de las cuerdas vocales y las frecuencias de resonancia infraglótica (tráquea y pulmones), que puede hacer cambiar el efecto tonal de las vibraciones glóticas.

1.3.2. Audición y percepción del tono

La fonación es un fenómeno fundamentalmente tonal, pues de las cuatro características del sonido (tono, intensidad, timbre y cantidad), la cantidad ya vimos que no es un parámetro comparable a los demás, sino que depende de la semi-abstracción, basada en el timbre, en que consiste detectar un segmento o un rasgo (tímbrico) estable; por su parte, la intensidad apenas es un parámetro informativo, ya que no hay ningún fenómeno lingüístico que dependa directamente de él (ni siquiera el acento, como veremos más adelante), seguramente porque la intensidad es un fenómeno coarticulatorio inevitable, que depende de la energía del chorro de aire espirado y cuya única finalidad es que la señal sea audible: por eso es lógico que las variaciones de intensidad no sean

informativas, porque han de graduarse según el nivel de ruido circundante, es decir, porque no dependen del mensaje en sí mismo, sino de factores externos (ruido ambiente, pero también pérdida de audición del hablante -fenómeno que, subjetivamente, es igual a la existencia de ruido ambiente-, u otros factores como el aumento de la tensión muscular general causado por un estado de excitación, etc.).

Los únicos parámetros de la voz directamente informativos en la lengua son los parámetros tonales: la frecuencia fundamental (el tono, que informa los fenómenos de acento y entonación) y la estructura de frecuencias secundarias (armónicos) reforzadas por los resonadores (el timbre, que informa las diferencias entre los distintos sonidos del lenguaje). Pero además es que la fonación es un fenómeno fundamentalmente tonal porque es en la vibración de las cuerdas vocales donde se produce el sonido, vibración cuya frecuencia es, precisamente, la frecuencia fundamental del sonido.

La audición, por tanto, es la sensación provocada por las variaciones de presión captadas por el oído, centrada en los fenómenos tonales que hemos examinado: la intensidad, una vez más, es captada sólo colateralmente (sólo si hay suficiente intensidad podrá haber audición, si hay demasiada intensidad la audición se verá distorsionada, etc.); por su parte, la percepción de la cantidad no pertenece al ámbito estricto de la audición en sí, que es una *sensación*, sino de la *percepción*, que actúa en un nivel más elaborado del funcionamiento cerebral, un nivel mediatizado ya por unidades captadas como tales por la conciencia.

Hay que distinguir, por tanto, entre *audición* propiamente dicha (la sensación transmitida tal cual por los sensores auditivos hasta el cerebro) y *percepción auditiva* (la interpretación que el cerebro hace de tal sensación, mediatizada ya por el aprendizaje cultural). Como siempre que se trata del tema de "sensación y percepción", sin embargo, ambos conceptos llegan a confundirse

en algunos momentos: también en el aspecto auditivo, especialmente cuando el condicionamiento cultural de lo que se espera *percibir* es tan fuerte que anula la propia *audición* de sonidos que ocurren realmente y que ha sido captados por el sentido: por ejemplo, los hablantes de una lengua están literalmente "sordos" a diferencias fonéticas que, si fueran hablantes de otra lengua, indudablemente percibirían. La *percepción*, por tanto, está fuertemente mediatizada por el funcionamiento lingüístico de la lengua del oyente.

Cuando se habla de "percepción del tono", sin embargo, no está claro si los distintos autores están tratando de describir el umbral diferencial de audición de la frecuencia (las variaciones mínimas de frecuencias que puede detectar el oído humano) o si intervienen otros factores de carácter lingüístico (las variaciones mínimas de frecuencia captadas por el oído que se consideran cambios significativos en una lengua).

Sobre la audición propiamente dicha del tono sabemos algunas cosas: el umbral auditivo del tono se establece entre 16 hz y 20.000 hz, aproximadamente, en una persona sana; sabemos también que para detectar la frecuencia fundamental de un sonido no es imprescindible que esté presente dicha frecuencia directamente en el *input* recibido (como ocurre con la voz telefónica, que filtra las bajas frecuencias de la voz y por tanto puede eliminar totalmente la frecuencia fundamental en las voces más graves), sino que el oído es capaz de inferirla automáticamente hallando el máximo común divisor de los armónicos disponibles (se cree que con sólo dos armónicos contiguos puede procesarse la F_0 , siempre que no se hallen solapados y que la diferencia de intensidad entre ellos sea de 20 ó 25 dB -'T Hart, Collier & Cohen 1990: 26-); en cuanto a la detección de cambios de frecuencia, parece ser que pueden detectarse cambios en una F_0 de a partir de 40 hz (presente o virtual), y que a partir de 4.000 hz apenas puede juzgarse ya con precisión (*ibid.*); sabemos también que aunque las vibraciones de las cuerdas vocales no son exactamente

periódicas, sino que hay algunas fluctuaciones incontroladas, la percepción del tono actúa como si la señal fuera estrictamente periódica, supliendo así las pequeñas diferencias ocurridas en la fonación (en la figura 14 puede verse la laringografía de una fonación en la que se pronunciaba la vocal [a] manteniendo el tono durante 3,2 segundos: nótese las diferencias que hay a lo largo de la señal, debidas a pequeñas fluctuaciones en la vibración).

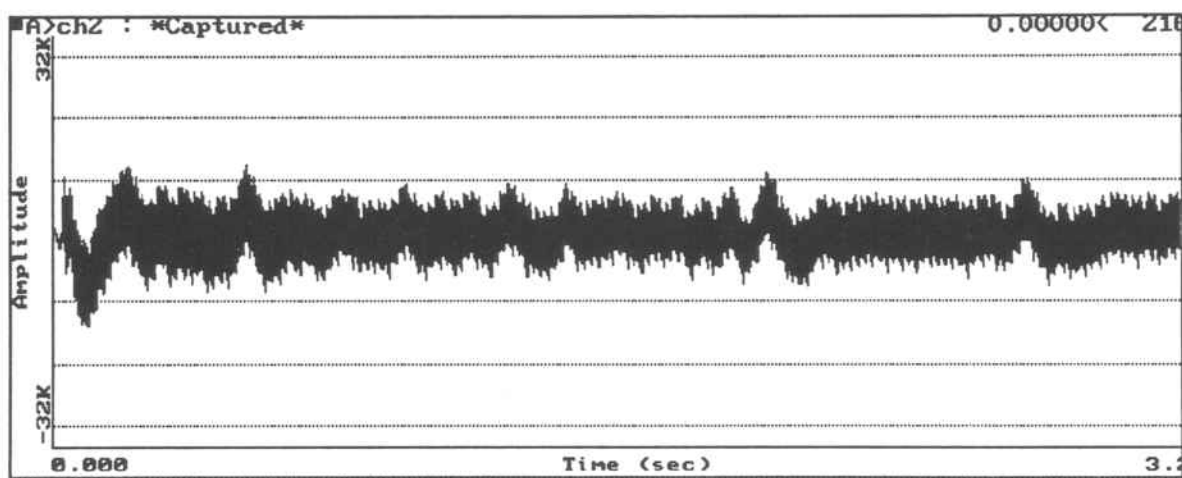


Figura 14 - Fluctuaciones en la vibración de las cuerdas vocales (laringografía de una fonación de 3,2 segundos).

Sin embargo, cuando comienza a discutirse sobre el umbral de percepción diferencial del tono empiezan las discrepancias: en 'T Hart, Collier & Cohen (1990: 25-36) hay una completa revisión de la bibliografía al respecto, en la que pueden observarse las diferencias de concepto que han animado a los distintos autores, como resultado de lo cual las discrepancias, en términos absolutos, son muy notables.

Sin embargo, parece bastante obvio que hay algunos factores que determinan tal percepción, los más importantes de los cuales son: *a)* el campo frecuencial en el que ocurra la variación de tono (la percepción es más precisa en los niveles tonales más bajos: a partir de un nivel de 1.000 hz, la precisión decae invariablemente); y *b)* la duración del estímulo tonal (con estímulos de 10

segundos, las diferencias percibidas son de 1 hz, con estímulos mucho más cortos, de 100 milisegundos, las diferencias percibidas nunca son menores a 150 hz -íbid.: 30-).

En realidad, ambos factores parecen tener bastante que ver con las diferencias intrínsecas que encontramos entre lenguas diferentes: desde los campos frecuenciales más comunes (que sirven a Tomatis, 1991, por ejemplo, para caracterizar las distintas lenguas) hasta la duración relativa de los sonidos en la cadena hablada, que depende del patrón rítmico de la lengua: así, podríamos aventurar que la percepción de las diferencias tonales está en cierto modo determinado por la lengua hablada por el oyente, lo que justificaría algunas discrepancias entre los distintos autores.

En cualquier caso, parece cierto que podemos percibir diferencias de entre 1 y 3 hz en los niveles frecuenciales en los que suele ocurrir la F_0 (entre 80 y 200 hz en los hombres, y entre 180 y 400 hz en las mujeres), lo cual es suficiente para nuestros objetivos. Seguramente, sin embargo, y tratándose de *percepción*, en algunos puntos de la cadena hablada especialmente informativos se percibirán diferencias mucho menores con una mayor precisión y en otros puntos menos informativos diferencias mayores serán imperceptibles: de este fenómeno se desprende la convicción de que una descripción completa de los fenómenos entonativos sólo puede llevarse a cabo con garantías de éxito teniendo en cuenta su carácter lingüístico y jerarquizado, y no sólo midiendo indistintamente las variaciones fonéticas del tono.

1.3.3. *Extracción y análisis de la F_0*

El tono o F_0 de un sonido, como sabemos, es el número de ciclos por segundo de la vibración que lo constituye, y su unidad de medida es, precisamente, el ciclo por segundo (c.p.s.) o hertzio (hz). Ya vimos que en la

fonación la F_0 se corresponde con la frecuencia de vibración de las cuerdas vocales o, lo que es lo mismo, con la frecuencia de cierres glotales.

En principio, la medición de la F_0 debería ser muy sencilla, teniendo en cuenta que toda la fonación es en sí un fenómeno tonal; sin embargo, la fluctuación de las vibraciones vocales y la propia fluctuación tonal durante la fonación convierte su medición en un problema no baladí: no basta, como podría suponerse ingenuamente, con medir el tiempo de fonación y "contar" el número de cierres glotales habidos, pues a lo largo de la fonación las variaciones de F_0 han sido continuas. Por tanto, había que encontrar algún modo de medir la F_0 en cada punto de la cadena hablada, en cada momento de la fonación, para poder seguir dichas variaciones tonales.

Durante algún tiempo, los fonetistas se consideraban a sí mismos algo así como "músicos del habla", y confiaban en su oído como en el instrumento más preciso. Así, la "melodía del habla" era comparada a la "melodía de la música" y se medía con sus mismas unidades (semitonos), empleando incluso instrumentos musicales para tratar de capturar y emular sus variaciones tonales. Sin embargo, había una barrera infranqueable para los que mantenían una concepción "musical" de las variaciones de tono en el habla: el sistema musical es un modelo arbitrario y cultural, constituido por unidades discretas absolutas y estacionarias (la escala musical) entre las cuales no es posible encontrar ni concebir ninguna otra diferencia (en música, la única unidad menor al semitono es la "coma", cuyo valor, sin embargo, no es constante, y varía según el sistema de afinación de la escala empleado por el ejecutante: fundamentalmente, entre los sistemas de Pitágoras, de W. Holder, de Zarlino y el sistema temperado de Ramos de Pareja; cfr. Zamacois, 1954: II, 146 y ss.).

Incluso con otros sistemas musicales no pertenecientes a la tradición occidental, con escalas más nutridas de unidades discretas (como el sistema hindú, que divide la octava no en doce sino en veintidós unidades), el posible

modelo con el que comparar la melodía del habla sigue siendo un sistema arbitrario de unidades discretas absolutas y estacionarias.

Una de las características de la melodía del habla es, precisamente, que apenas hay momentos de tono estacionario, y que, si los hay, no corresponden con las unidades discretas en las que segmentamos el discurso: sonidos y sílabas. Por tanto, difícilmente podremos analizar las variaciones de F_0 en el discurso hablado tomando como modelo el esquema (cultural) de las canciones, en las que una nota musical corresponde a una sílaba.

Por lo demás, los valores absolutos que componen los distintos sistemas musicales han sido fijados arbitrariamente por la tradición, y no dejan de ser estadios tonales demasiado definidos e inamovibles como para servir de modelo a las sutiles variaciones de F_0 que ocurren en el habla, que podrían constituir, desde luego, un sistema definido de estadios tonales, pero que difícilmente iba a coincidir con el resultado de una tradición como la musical, extremadamente encerrada en sí misma, metacultural y casi algebraica. Con todo, hay varios autores que afirman, como veremos, que puede describirse la entonación de una lengua con un sistema de dos, tres o cuatro tonos distintos (lo cual constituye, curiosamente, una escala de tonos muy parecida a la empleada por el canto gregoriano).

El modo de medir la F_0 , en fin, ha de ser acústico y no tomar como referencia más que los propios valores acústicos del sonido. Si tomamos la definición de F_0 , por ejemplo, veremos que la manera más rápida de medirla es, precisamente, medir la duración de un ciclo, cuyo valor inverso será, justamente, el valor de la F_0 . Observando el oscilograma de la figura 15, a la distancia entre los picos de cada ciclo (en general, entre dos puntos claramente reconocibles de cada ciclo), que es una distancia temporal, la llamamos T_0 : la F_0 la obtendremos de la sencilla fórmula $F_0 = 1 / T_0$.

En la figura, el oscilograma mostrado corresponde a la vocal [a]. La distancia entre los picos sucesivos de cada ciclo (entre a y $a' = T_0$) es de 0,0085 segundos (8,5 ms.); la F_0 del sonido representado, entonces, es de 117,6 Hz.

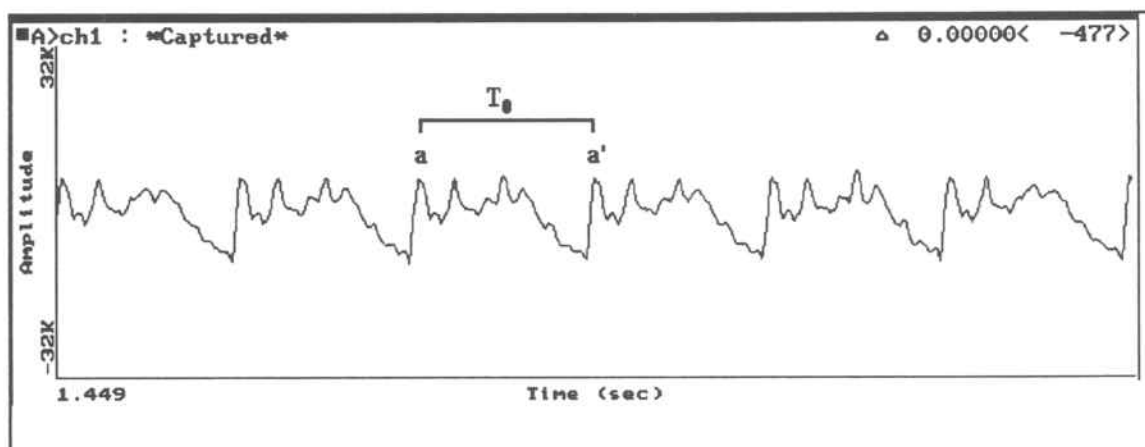


Figura 15 - Oscilograma de la vocal [a].

El oscilograma, con todo, no es más que una forma de representar visualmente el sonido, en la que aparecen sólo los parámetros de tiempo (eje de las abscisas) y de amplitud de la onda (eje de las ordenadas).

El sonido, en realidad, debería ser representado en tres dimensiones, pues los parámetros a partir de los cuales se definen sus características son tres: tiempo (en función del cual se define la F_0 , como hemos visto), amplitud de la onda (que define la intensidad del sonido) y altura frecuencial (en función de la cual definimos la estructura resultante del filtrado de armónicos: los formantes). Este último parámetro, el de altura frecuencial, es obviamente un producto de la abstracción, que consiste en la descomposición de la onda compleja en todas las ondas simples que la componen, como puede verse en la figura 16, que es una representación tridimensional del sonido. A este tipo de representación tridimensional se le llama *objeto sonoro*.

Entre los puntos a y b se representa el transcurso del tiempo; entre los puntos b y c , se representan las frecuencias de cada uno de los armónicos, como resultado de la descomposición de la onda compuesta en las ondas simples que la componen; entre los puntos b y d , finalmente, se representa la amplitud de las ondas.

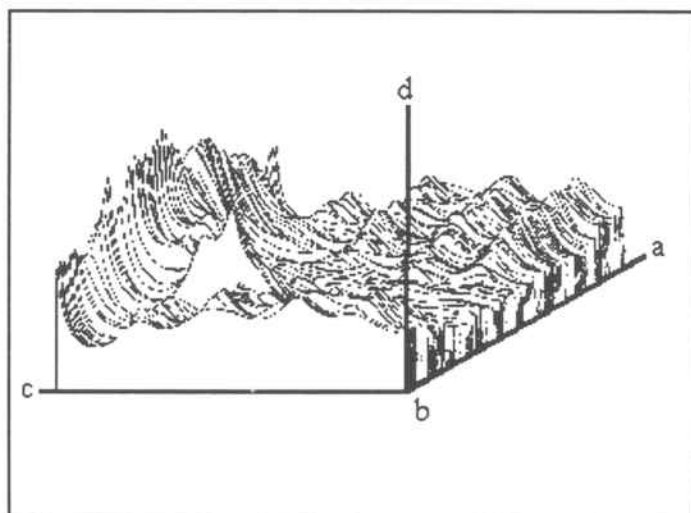


Figura 16 - El objeto sonoro.

Así, el oscilograma, que es la representación más fidedigna de la onda compuesta, es el resultado de *mirar* el objeto sonoro desde la perspectiva "frontal": $a-b-d$.

Desde la perspectiva "lateral" $d-b-c$, en cambio, obtenemos un *espectro* (figura 17), en el que visualizamos los

sucesivos armónicos, su altura frecuencial y su valor relativo dentro de la estructura formántica. En el espectro puede medirse muy fácilmente el valor frecuencial del primer armónico (que es la F_0); sin embargo, el espectro no muestra el valor temporal, sino que es una imagen estática de un momento del sonido, con lo que no pueden seguirse las variaciones de F_0 a lo largo de la cadena hablada.

Desde una perspectiva "aérea" del objeto sonoro, en cambio, sí pueden seguirse las variaciones del primer armónico o F_0 : la perspectiva $a-b-c$ muestra el conjunto de armónicos en su evolución temporal, su altura frecuencial e incluso su amplitud relativa, gracias a la convención de representar la intensidad del armónico mediante la intensidad del color: a mayor negror, mayor intensidad, con lo que el resultado permite visualizar completamente, y de un vistazo, la estructura de armónicos potenciados y de armónicos filtrados. A esta

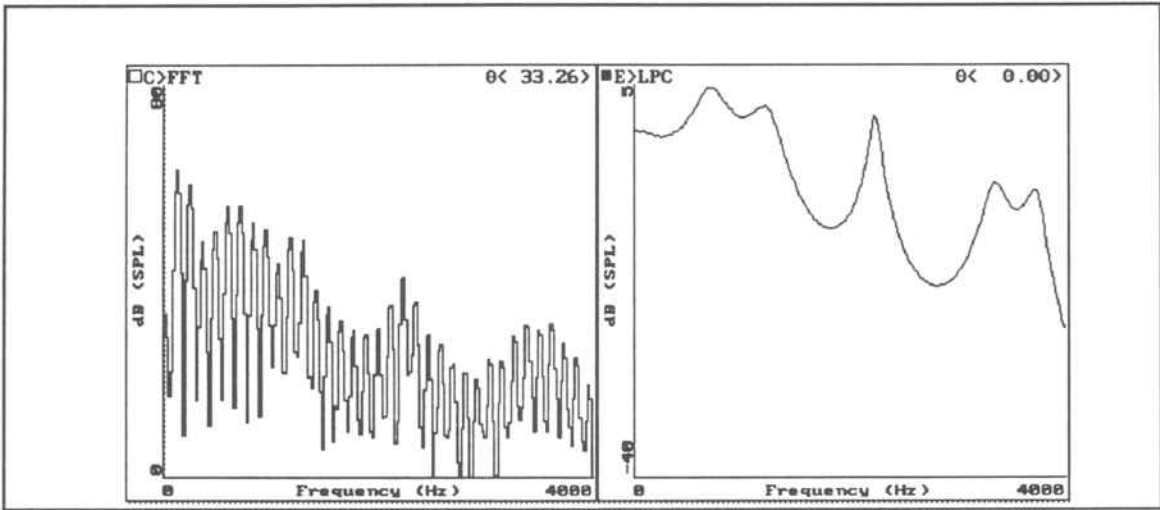


Figura 17 - Espectro de [a]: a la izquierda, el espectro según la Transformada de Fourier (FFT); a la derecha, según el Cálculo de Predicción Lineal (LPC).

representación del objeto sonoro, tal vez la más completa, se le llama *espectrograma* o, más comúnmente, *sonograma*, y viene a ser como una sucesión de espectros.

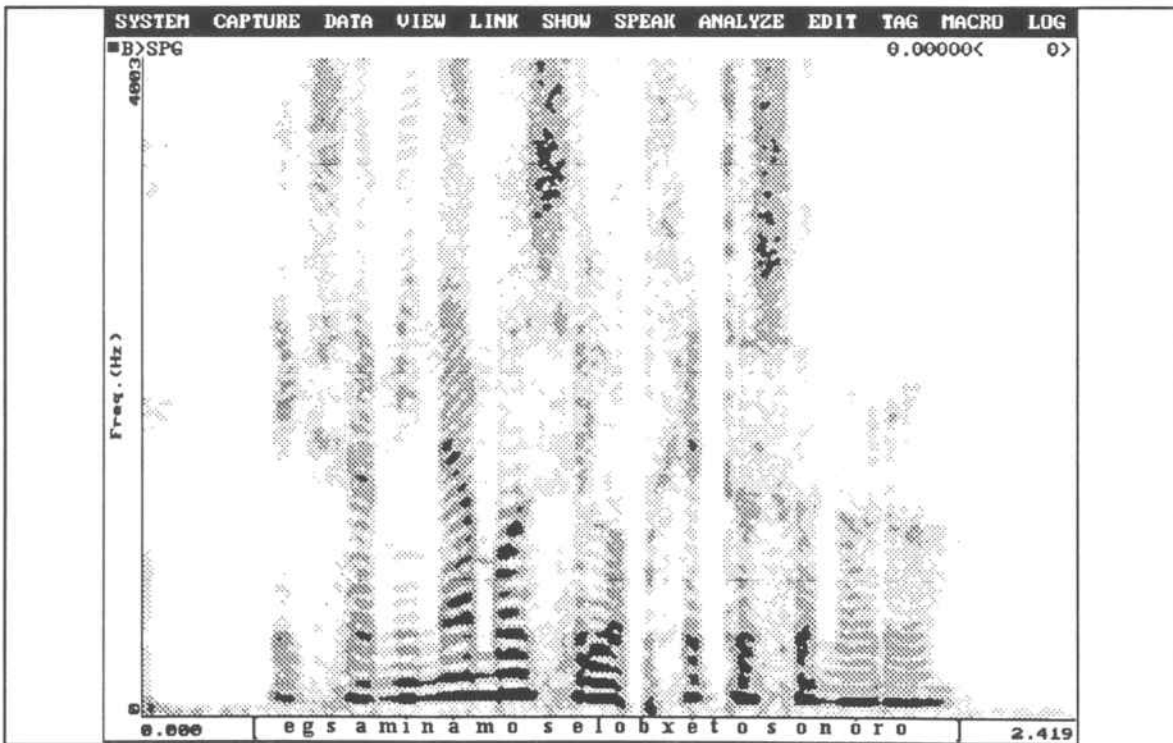


Figura 18 - Sonograma del enunciado: "Examinamos el objeto sonoro" (filtro de banda estrecha).

El sonograma de la figura 18 muestra la cadena "examinamos el objeto sonoro", obtenido con un filtro de banda estrecha (una muestra cada 29 hz), y en él pueden observarse claramente los armónicos y su evolución. El primer armónico, sin embargo, apenas puede seguirse con la vista, con lo que la medición del tono así puede resultar complicada.

Puede medirse, a cambio, un armónico superior, ya que sabemos que los armónicos son todos múltiplos de la F_0 : si medimos el quinto armónico, por ejemplo, la F_0 se obtendrá de dividir su valor entre 5.

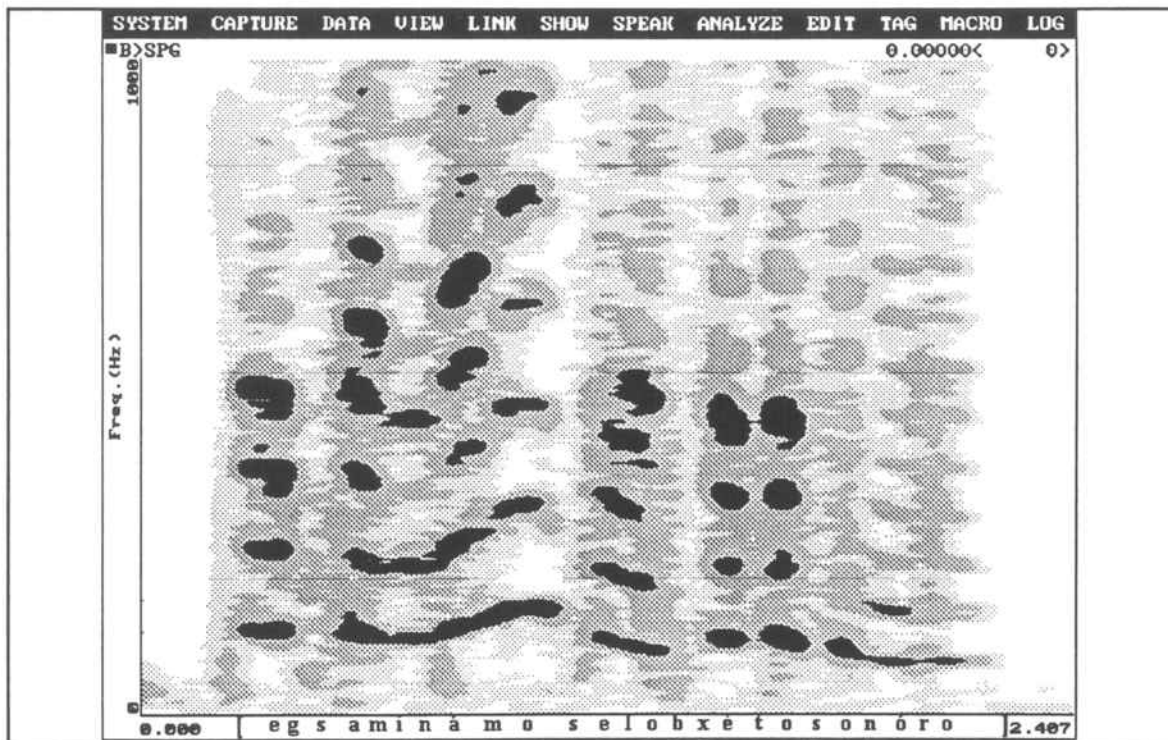


Figura 19 - Ampliación de los primeros 1000 Hz. del sonograma anterior.

Puede optarse también por ampliar la zona de bajas frecuencias, para "acercarse" a la visión del primer armónico; en la figura 19 se muestra la ampliación de los primeros 1000 hz del sonograma anterior: la evolución del primer armónico es más visible, aunque las variaciones de tono apenas se notan aún.

También puede calcularse la F_0 de un sonido mediante un sonograma

obtenido con un filtro de banda ancha. En la figura 20, la misma cadena hablada de las figuras 18 y 19 se muestra ahora en un sonograma de banda ancha (una muestra cada 146 hz). Obviamente, la representación de la estructura de los formantes es ahora mucho mejor, aunque se ha perdido la definición de los armónicos individualmente.

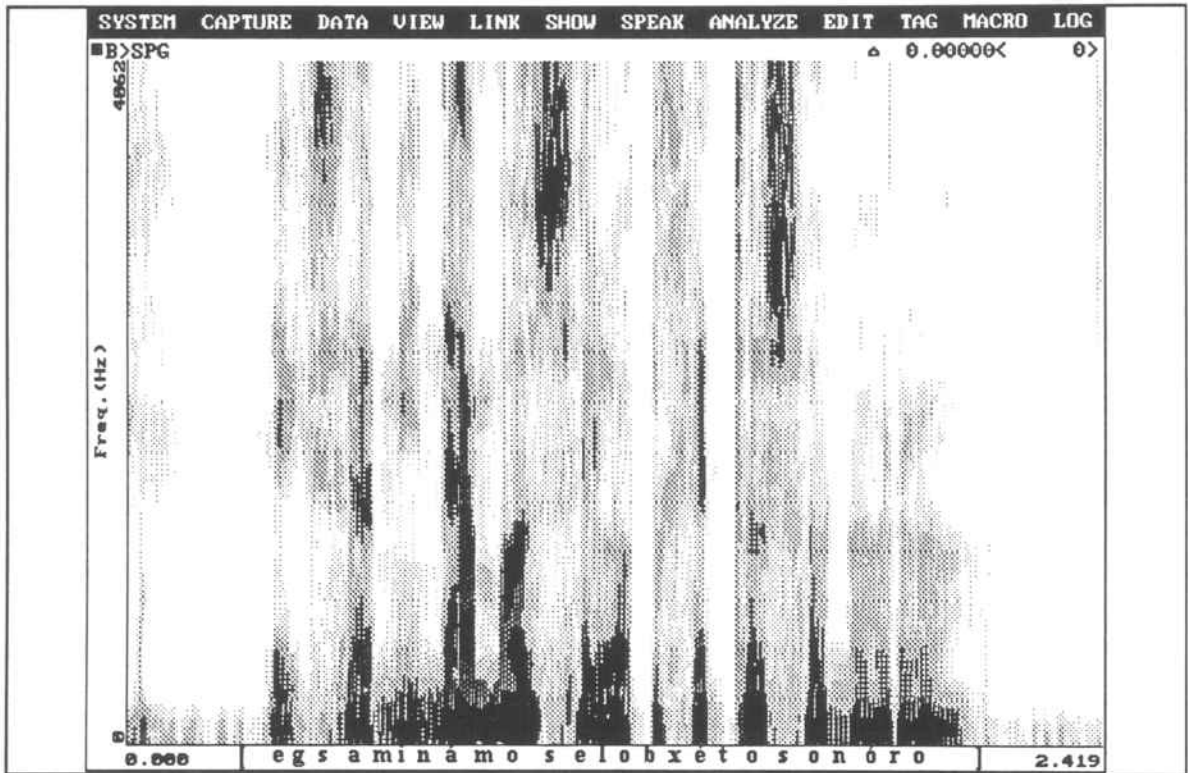


Figura 20 - Sonograma del enunciado: "Examinamos el objeto sonoro" (filtro de banda ancha).

Obsérvese que ahora aparecen en el sonograma unas líneas verticales paralelas, a veces muy definidas (como en las vocales) y a veces muy confusas (en los ruidos, como la [x] de "objeto"): tales líneas paralelas corresponden a los pulsos glotales (los cierres y aberturas de la glotis); así, cada línea corresponde a la salida de una "bola de aire" y, entre línea y línea (entre abertura glotal y abertura glotal), no hay nada, porque en ese momento estaban cerradas las cuerdas vocales.

Cada línea vertical del sonograma, por tanto, equivale al pico de un ciclo

en el oscilograma: si medimos la distancia entre tales líneas (entre pulsos glotales) estamos midiendo la distancia T_0 , con lo podremos obtener también el valor de la F_0 en ese punto. Aunque, obviamente, extraer de este modo la evolución del tono a lo largo de la cadena puede ser muy trabajoso.

Como vemos, todos los modos de extraer el tono de una emisión de voz por medios acústicos se reducen al mismo principio: medir la distancia T_0 e inferir el valor de F_0 , ya sea en el oscilograma, en el espectro o en las distintas modalidades de sonograma.

También podemos extraer el valor de la F_0 con medios articulatorios, en los que el principio sigue siendo el mismo: medir las vibraciones de las cuerdas vocales.

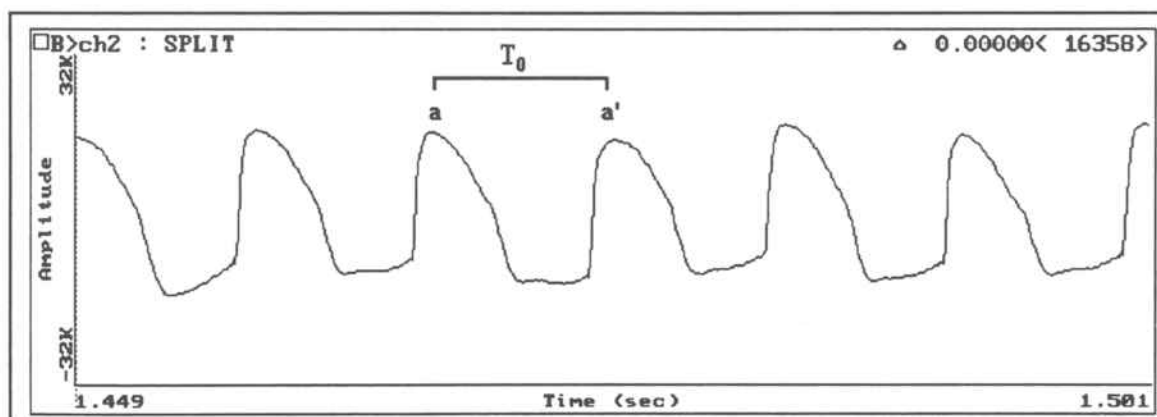


Figura 21 - Laringografía de la vocal [a] (correspondiente al oscilograma de la figura 15).

Podemos hacerlo con el laringógrafo, que nos ofrece una representación de la vibración de las cuerdas vocales en forma de onda triangular (la *Lx waveform*, que no debe confundirse con la *Sp waveform*, que mide la salida de aire y la presión infraglotica: cfr. Fourcin, 1981; Abberton, Howard & Fourcin, 1989) en la que el pico corresponde al máximo cierre de la glotis (el momento de máximo contacto entre las cuerdas vocales, y por tanto de máxima

conductibilidad, detectada por unos sensores colocados en el cuello del informante) y la depresión a la máxima abertura (el momento de más salida de aire).

En la figura 21, la laringografía muestra la misma [a] cuyo oscilograma se mostraba en la figura 15: la distancia entre a y a' es la misma que en el oscilograma, y por tanto el valor de F_0 .

La extracción de tono mediante laringografía es mucho más fiable que por medios acústicos, cuya fuente no es el propio informante, sino un micrófono, con lo que eso supone de pérdida de información, de distorsión en las intensidades, etc. Actualmente, hay disponible en el mercado un laringógrafo conectado a una tarjeta de sonido controlada por ordenador que permite extraer, en tiempo real, la evolución del tono mediante el análisis de los contactos vocales (figura 22).

Con todo, los analizadores acústicos permiten la extracción y el análisis del tono de una señal grabada, lo cual constituye una ventaja indudable en lo que se refiere al análisis de la entonación en el habla espontánea.

Tales analizadores automáticos de tono muestran una representación temporal de la curva tonal, en tiempo real, con lo que todos los métodos reseñados de extracción manual del tono

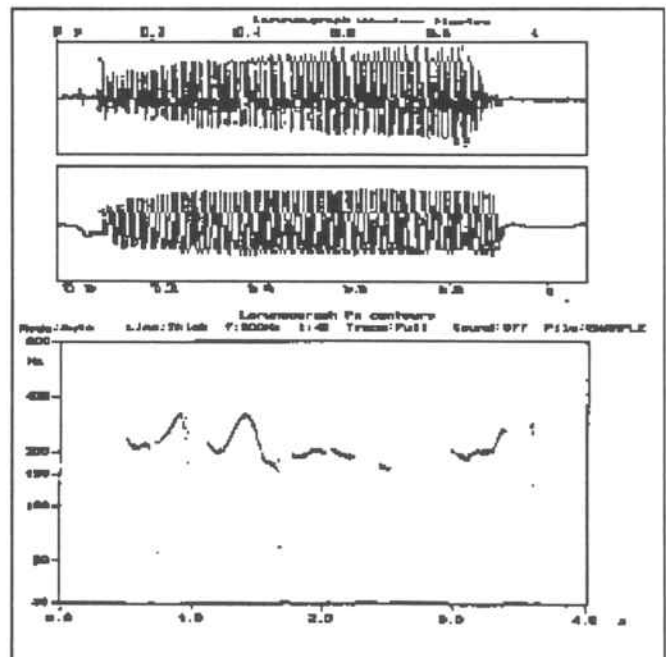


Figura 22 - Pantalla del programa de extracción de F_0 por laringografía PCLX.

pertencen al pasado, es decir, al propio funcionamiento automático de estos instrumentos computerizados.

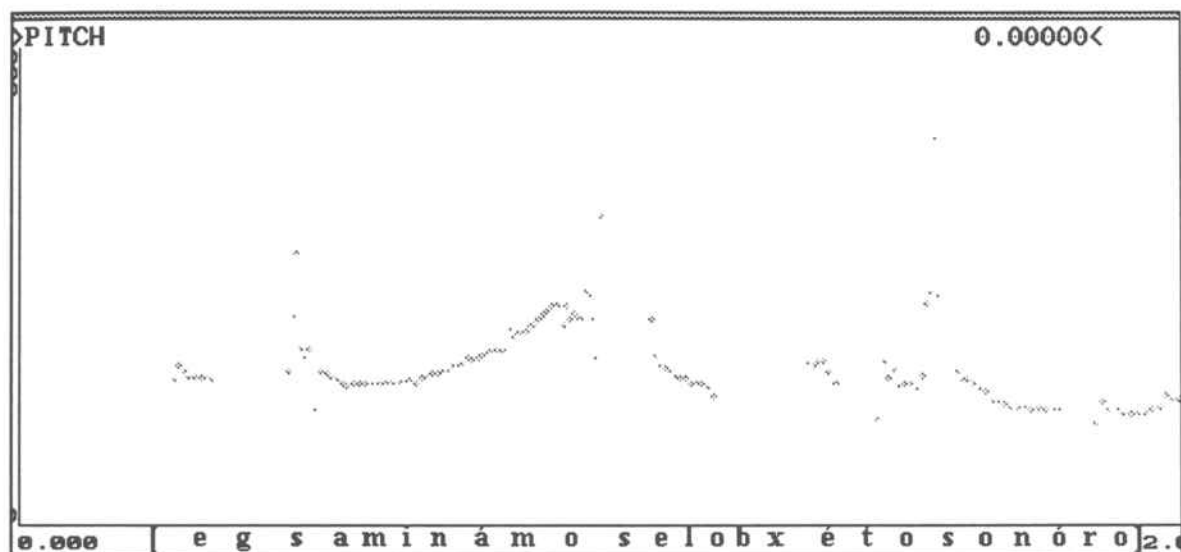


Figura 23 - F₀ del enunciado: "Examinamos el objeto sonoro", correspondiente a las figuras 18, 19 y 20.

En la figura 23 puede verse la curva tonal de la misma señal representada en los sonogramas de las figuras 18, 19 y 20, extraída por uno de los analizadores más populares de la actualidad: el *Visi-Pitch* de la casa *Kay*. La visualización, como puede comprobarse, no tiene ya nada que ver con la confusa línea del sonograma: ahora es mucho más clara, sin ningún elemento extraño, y con los valores en datos absolutos y en tiempo real.

1.4. El concepto de entonación

1.4.1. Definición y parámetros de la entonación

El continuum sonoro en que consiste el habla se halla constituido, básicamente, y como vimos en apartados anteriores, por una frecuencia fundamental (F_0) y por una serie de frecuencias secundarias o armónicos potenciados a su paso por los resonadores supraglóticos: tales frecuencias (fundamental y secundarias) pueden tener en su conjunto una mayor o menor intensidad, pero la fonación es, como sabemos, un fenómeno de carácter fundamentalmente tonal. Por tanto, el continuum sonoro se halla constituido, básicamente, por un tono y un timbre que fluctúan y cambian a lo largo de toda la emisión y que forman, en el caso del timbre, una serie de segmentos más o menos diferenciados (la cadena de sonidos), y, en el caso del tono, una suerte de melodía interrumpida sólo por los segmentos sordos.

Las variaciones de timbre a lo largo de la emisión de voz es estudiada por la fonología segmental; las variaciones de tono, por la fonología suprasegmental.

Sin embargo, y como puede observarse claramente, la misma terminología empleada para referirnos a los ámbitos de estudio parece querer indicar que la base de todos los fenómenos fónicos es el segmento, el sonido independiente, que sería la unidad fónica indiscutible (ya vimos cómo esta concepción tiene un origen abiertamente lecto-escritor y, aún más, alfabético); ni siquiera la evidencia fonética de que los sonidos no son más que un haz de rasgos tímbricos fluctuantes cuyo límite no siempre está muy claro ha sido capaz de eliminar de las mentes de los propios especialistas el mito de que la emisión de voz la constituye una cadena de sonidos definidos.

Así pues, si bien suele aceptarse que la fonología suprasegmental, y en

concreto la entonación, estudia las variaciones de tono a lo largo de la emisión de voz, suele entenderse también que tales variaciones se "superponen" a la cadena de sonidos (de ahí "*suprasegmental*"): por tanto, su estudio se ha considerado tradicionalmente secundario, pues se trataría de un fenómeno difícilmente segmentable, fuera de la doble articulación del lenguaje y que seguramente ni siquiera es relevante en cuanto a información lingüística; en el que no merece la pena, pues, detenerse más de lo debido.

Además de la entonación, suele hablarse de otros dos fenómenos suprasegmentales: el acento y el ritmo. El acento es un fenómeno lingüístico que depende en cierto modo del nivel segmental, pues consiste en poner de relieve unos segmentos frente a otros, según una definición generalmente aceptada (aunque ya veremos más adelante, cuando tratemos el tema con detenimiento, que es una definición que conviene revisar). El ritmo, por su parte, es un fenómeno que a menudo ni siquiera se considera aparte del acento, pues consiste precisamente en la recurrencia temporal de los segmentos acentuados. Así, los otros dos fenómenos suprasegmentales (para muchos autores, uno solo) tienen un anclaje muy definido en el ámbito segmental, aunque sean fenómenos tonales.

Pues el parámetro principal del acento, en efecto, es el tono: una inflexión de tono convierte el segmento en acentuado, pero también una mayor duración, y sólo en contadas ocasiones la intensidad del sonido es determinante.

Si los parámetros físicos del fenómeno suprasegmental acento son, por tanto, la F_0 , la duración y la intensidad, no puede extrañar que los parámetros del otro fenómeno suprasegmental, la entonación, puedan ser los mismos. Así, nos encontramos, en principio, con dos definiciones distintas y posibles de entonación: para una, la entonación sería las "variaciones de la F_0 " a lo largo de la emisión de voz que aportan información lingüística y, para otra, las "variaciones de F_0 , intensidad y duración".

Ya vimos que hay una relación en ocasiones muy estrecha entre la F_0 y la intensidad, que es un parámetro originado directamente por la presión infraglótica, cuyo aumento también puede incidir en un aumento de la F_0 . Por otra parte, la tradicional explicación del acento como fenómeno de intensidad (hoy sabemos que la intensidad tiene un papel mínimo en la percepción de los segmentos acentuados, como hemos dicho -Solé, 1984-) ha ayudado también a concebir la entonación como un fenómeno relacionado con la intensidad.

Pero, sobre todo, la concepción de la entonación como el principal transmisor de la expresividad, como el fenómeno de la emoción, que subyace en buena parte de los autores que lo han tratado, ha contribuido definitivamente a suponer la intensidad como un parámetro físico fundamental de la entonación: como dice Quilis (1981: 394) citando a León (1972), "la intensidad del patrón melódico tiene también un valor simbólico directo con la intensidad del sentimiento expresado".

En cuanto a la duración como parámetro entonativo, de nuevo el máximo argumento es que figura como parámetro muy relevante en el fenómeno del acento; por lo demás, parece ser que hay algunos efectos causados por ella, por ejemplo en el habla cuchicheada, donde no hay F_0 .

En efecto, en el habla cuchicheada no hay fonación, no hay producción tonal, por lo que, en un principio, el aire saldría sordo y rozando por las cavidades supraglóticas como ocurre en la articulación de las consonantes fricativas. Sin embargo, el habla cuchicheada es inteligible, no se reduce a una mera sucesión de fricaciones sordas y podemos distinguir claramente en ella las vocales de las consonantes. La razón de ello es que se produce un efecto por el que el aire roza en la glotis, donde se produce un ruido que en la cavidad bucal se moldea como si fuera el sonido producido en la fonación normal. Por tanto, puede decirse que el timbre es independiente de la propia fonación, pues incluso sin ella la cavidad bucal actúa de filtro y genera el timbre.

Así, podemos observar formantes en las vocales cuchicheadas, que presentan los mismos rasgos generales de timbre que las caracterizan en el habla normal, aunque con modificaciones notables, como un incremento en la altura de los primeros formantes, la potenciación del tercer y cuarto formante, etc. Entonces, el mismo efecto acústico que permite reconocer los formantes vocálicos permitiría inferir de ellos (como vimos en el apartado 1.3.2.) una F_0 virtual: la sensación entonativa, por tanto, puede ser difusa, pero cierta, e incluso en las lenguas tonales puede reconocerse un tono virtual en las vocales cuchicheadas, aunque sean fonéticamente sordas.

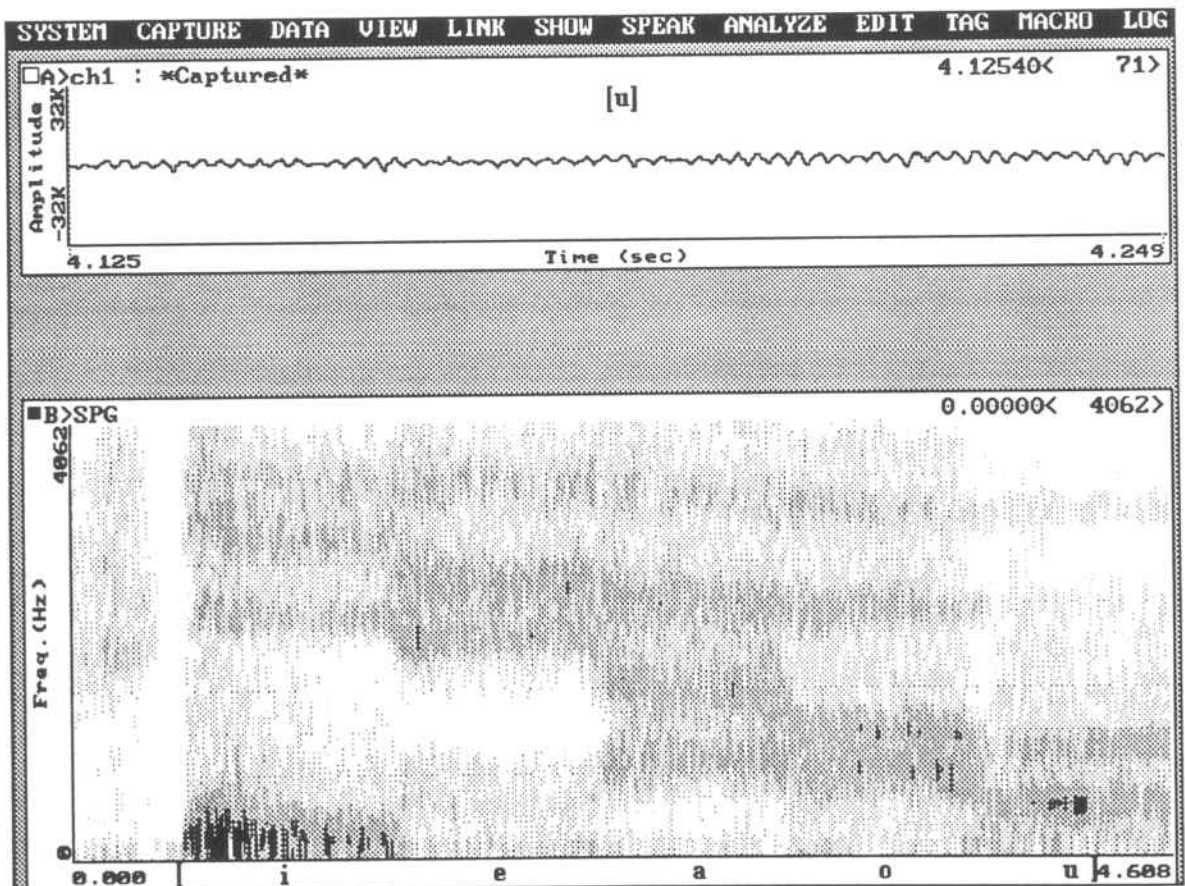


Figura 24 - Oscilograma de la [u] y sonograma de las cinco vocales, cuchicheadas.

Es decir, las explicaciones según las cuales en el habla cuchicheada, al no haber tono, cumplen la función entonativa la intensidad y la duración resultan

poco convincentes porque: el timbre de las vocales es un fenómeno tonal (pues consiste en el reforzamiento de bloques de armónicos o formantes); la duración es un parámetro que depende de la identificación de los segmentos tímbricos; si hay percepción de segmentos más largos y más breves es que hay algún tipo de percepción tonal (en este caso, tímbrica); las dificultades que el oído pueda tener para reconstruir la F_0 virtual de una emisión de voz sorda pero con algún rasgo tímbrico coincide con los resultados experimentales que muestran entre un 30% y un 40% de error en la percepción de entonaciones marcadas como la interrogativa (Quilis, 1981: 411).

En la figura 24 pueden verse claramente los formantes de las cinco vocales cuchicheadas, y el oscilograma correspondiente a la [u]. En él, algo parecido a un ciclo periódico permite establecer la " F_0 virtual" del ruido: 223Hz. En el sonograma, los picos regulares del ciclo presentan también un dibujo muy semejante al que presentarían los pulsos glotales: en resumen, información suficiente como para detectar el tono.

Muy a menudo se presenta también el timbre como un parámetro no determinante pero sí de algún modo influyente en la entonación: cuando se habla del tono intrínseco de las vocales (Hombert, 1978; para el español, Mateo, 1988) y de las consonantes (Di Cristo, 1982; Gili Gaya, 1924). El timbre es un fenómeno tonal, pero condicionado en última instancia no por las vibraciones de las cuerdas vocales sino por la estructura que adoptan los resonadores; así, el tono intrínseco de los sonidos (que es un efecto del juego de los formantes que los caracterizan) ha de estar supeditado al tono general de la emisión vocal, cuya fuente de producción es distinta: difícilmente, por tanto, podría influir en él.

La diferencia entre los dos "tonos" puede verse claramente en un sonograma: la altura de los formantes condiciona el llamado tono intrínseco del sonido, pero es sólo la distancia (temporal) entre los pulsos glotales la que

marca la F_0 del mismo. Es decir, el tono intrínseco es mensurable sólo en condiciones de aislamiento tímbrico.

En resumen, definir la entonación como un fenómeno producido por otros parámetros además de la F_0 supone un cierto grado de confusión sobre la naturaleza de tales parámetros: como hemos visto, el timbre es un parámetro tonal, pero independiente de la altura del primer armónico: las cuerdas vocales producen el sonido y todos los armónicos, pero su frecuencia de vibración tiene ya poco que ver con la estructura de armónicos potenciados a su paso por los resonadores; la duración, por otra parte, depende directamente del nivel segmental, y no comparte el mismo rango de parámetro físico que el tono, la intensidad y el timbre: no es un parámetro comparable, porque actúa a un nivel en el que hay un punto de abstracción; la intensidad, finalmente, ni siquiera informa al fenómeno del acento (como se creía), con lo que se hace patente su naturaleza de parámetro "acompañador", imprescindible para que la voz sea audible (dependiente por tanto del contexto), pero no relevante lingüísticamente (pues consideramos que la "tensión emocional" no es un producto lingüísticamente codificado sino motivado fisiológicamente).

Obviando las relaciones muy estrechas que sin duda ha de tener la entonación con los otros fenómenos suprasegmentales (acento y ritmo), igualmente producidos por la F_0 , y que trataremos más adelante, podríamos definir, entonces, la entonación como *las variaciones de F_0 que cumplen una función lingüística a lo largo de la emisión de voz.*

Sin embargo, no todos los autores estarían de acuerdo con esta definición, pues siempre ha habido ciertos reparos en considerar la entonación como un fenómeno claramente "lingüístico".

Dos son las posturas que se adoptan para cuestionar el carácter lingüístico de la entonación y que Martinet (1960), por ejemplo, resume de este modo: por un lado, señala que la entonación no entra en el juego de la doble articulación,

y que, al tratarse de un fenómeno no segmentable, no hay posibilidad de analizar el significante que pudiera constituir; por otro lado, al afirmar que la entonación cumple "funciones mal diferenciadas" insiste en su carácter fuertemente motivado, cuya función principal sería la función expresiva, es decir, de transmisión no codificada de emociones.

A la primera afirmación han salido al paso numerosos autores, insistiendo en el carácter segmentable de la entonación, en la existencia de "fonemas suprasegmentales", etc.: tratamos de ellos en el apartado 2.3.1. Sin embargo, el segundo tema, el carácter motivado y emocional de la entonación (la doctrina de cuyo máximo defensor, D.W. Bolinger, exponemos en 2.2.2.) rara vez ha sido combatido, y la posición más generalizada es la de aceptarlo como un hecho que, de cualquier modo, no niega su carácter igualmente lingüístico (v. 5.2.).

Así, actualmente suele considerarse que la entonación es, en efecto, un fenómeno de carácter lingüístico que actúa, sin embargo, en diversos niveles, algunos de los cuales no son lingüísticos sino expresivos, etc. Según el resumen que hace Quilis (op.cit.: 376 y ss.; v. 2.5.2.), la entonación actuaría en un nivel claramente lingüístico (en el que cumpliría fundamentalmente una función distintiva), en un nivel sociolingüístico (en el que cumpliría una función informativa sobre el hablante e identificativa) y en un nivel expresivo (cuya función principal sería la transmisión del estado emocional del hablante).

1.4.2. F_0 , melodía y entonación

Una vez establecido que la entonación tiene que ver únicamente con las variaciones de la F_0 (y no con la duración, la intensidad o el timbre), conviene distinguir entre ambos conceptos, pues "entonación" y " F_0 " no son términos homologables.

La F_0 es un parámetro acústico producido por las vibraciones de las

cuerdas vocales, que está en la base también del parámetro de timbre. Ya hemos insistido suficientemente en el carácter tonal de la fonación e incluso de la propia audición.

Por su parte, la entonación es un fenómeno lingüístico, producto de una abstracción teórica muy importante: "la sucesión de variaciones relevantes de la F_0 en una emisión de voz".

La F_0 informa al fenómeno de la entonación, pero también a otros fenómenos, como el acento. La relación que se establece entre la F_0 y la entonación, pues, equivale a la relación que hay entre los fenómenos físicos (articulatorios, acústicos o auditivos) y los fenómenos perceptivos, para explicar los cuales media la construcción de un modelo teórico.

Así, si la F_0 se mide en Hz., la entonación no puede medirse propiamente, porque consiste en una serie de modelos abstractos o contornos, con unos significados más o menos definidos, que cumplen unas funciones lingüísticas determinadas teóricamente y que intervienen en los intercambios comunicativos como unidades operativas.

Sin embargo, no siempre se ha sabido diferenciar lo fonético de lo fonológico. Lieberman (1967), por ejemplo, pretende dar una explicación completamente fisiológica de los fenómenos entonativos, según una concepción que no considera a la entonación un fenómeno lingüístico, sino puramente fisiológico; sin embargo, hace entrar en juego conceptos claramente fonológicos, como "grupo espiratorio" (al que hace coincidir, además, con el "sintagma gramatical") o como "prominencia" (un concepto semejante al de "acento"), a pesar de que pretende darles una explicación únicamente fisiológica (pero usando otros conceptos como el de "duración", "sílabas", "segmento", etc., abstracciones fonológicas todos ellos, como hemos visto repetidamente): en el capítulo siguiente examinaremos con más detalle su doctrina.

En realidad, los fenómenos lingüísticos ni siquiera existen como tales *per*

se, sino en función de nuestra organización teórica de los fenómenos físicos: así, podemos hablar de acento y ritmo, o sólo de acento, según nuestra manera de interpretar los datos; podemos hablar, también, de acento y entonación, o sólo de prosodia, si interpretamos que ambos fenómenos actúan a la vez. La realidad será la misma, al margen de nuestra concepción teórica. Pues bien: la F_0 es la realidad cambiante, y la entonación es la interpretación que podemos hacer de tales cambios.

La sucesión de tonos a lo largo de la emisión de voz constituye una *melodía*, como dijimos, interrumpida sólo por los segmentos sordos (generados sin vibración glotal): una melodía constituida por tonos más o menos estables y por tonos que se deslizan, que fluctúan, etc. Esa melodía, sin embargo, y consecuentemente con lo que hemos dicho, tampoco es la entonación.

La "melodía" es la mera sucesión de tonos; la "entonación", una vez más, la interpretación lingüística de la melodía. En la melodía, por tanto, hay multitud de variaciones irrelevantes lingüísticamente (variaciones que han dado en llamarse *micromelódicas*); en la entonación, sólo los elementos melódicos relevantes. La melodía, en fin, es también realidad; la entonación, un conjunto de modelos abstractos: la relación que mantienen *melodía* y *entonación*, entonces, equivale exactamente a la que mantienen *sonido* y *fonema*.

Pero tampoco F_0 y melodía son términos perfectamente homologables, a pesar de referirse a fenómenos *físicos*: la unidad de medición de la F_0 , por ejemplo, es el Hz, un valor absoluto; la unidad de análisis melódica, sin embargo, no puede ser un valor absoluto, sino relativo, pues la melodía sólo puede definirse midiendo las sucesivas variaciones de tono. La F_0 es el valor en sí, y la melodía la sucesión relativa de valores.

La unidad relativa más utilizada desde antiguo en los análisis melódicos, como vimos, es el *semitono*, una unidad logarítmica que designa la distancia (D) entre dos valores frecuenciales (f_1 y f_2) según la siguiente fórmula ('T Hart,

Collier & Cohen, 1990: 24):

$$D = 12 \log_2 f_1 / f_2 = 12 / \log_{10} 2 \cdot \log_{10} f_1 / f_2$$

En el ejemplo de 'T Hart, Collier & Cohen (op.cit.), se plantea qué diferencias puede haber entre dos distancias frecuenciales, una de 100 a 150 hz y otra de 180 a 270 hz: en un caso, la distancia es de 50 hz y en otro de 90 hz. En términos absolutos, pues, se trataría de dos distancias muy diferentes. En términos relativos, sin embargo, la distancia es, en los dos casos, de 7.02 semitonos. Se trata, entonces, de la misma melodía, pero emitida seguramente por un hombre o por una mujer. Así, los modelos entonativos deben basarse en melodías establecidas en términos relativos, y no en valores frecuenciales absolutos, pues las diferencias entre individuos son siempre muy grandes (diferencias que se deben a la estructura de los órganos fonadores: grosor de las cuerdas vocales, flexibilidad, etc.).

Otra posibilidad, equivalente al empleo del semitono como unidad melódica, es referirse a la distancia D como un porcentaje de variación de f_2 con respecto a f_1 . En el mismo ejemplo, la distancia en los dos casos es de un 50% de variación (tomando siempre la f_1 como el valor 100 sobre el que calcular el porcentaje de variación de la f_2).

En cualquier caso, la medida de la distancia entre dos frecuencias consecutivas en la melodía debe establecerse según un sistema logarítmico: a frecuencias más bajas, una diferencia menor equivale a una diferencia mayor a frecuencias más altas. Siguiendo con el ejemplo anterior, la distancia entre una f_1 de 20 Hz y una f_2 de 30 Hz seguiría siendo la misma, a pesar de tratarse sólo de 10 Hz: una D cifrada en un 50 % de variación relativa, o en 7.02 semitonos.

La ventaja del sistema de porcentajes consecutivos es que ofrece la misma información logarítmica que el sistema por semitonos, pero con valores lineales,

más fáciles de asumir por un no profano, y más fáciles de calcular. Debe notarse, pues, que un 50 % de distancia entre dos frecuencias no equivale a la mitad de una octava (6 semitonos), sino a algo más (7.02 semitonos), mientras que un 100 % de variación entre dos frecuencias equivale exactamente a una octava (12 semitonos).

1.4.3. Conclusión

A lo largo de este capítulo hemos visto cómo la emisión de voz constituye un verdadero *continuum* sonoro de rasgos tímbricos que fluctúan y cambian, cuya segmentación en sonidos independientes está fuertemente condicionada por una perspectiva lecto-escritora: más adelante este será un tema recurrente, por cuanto el mismo concepto de fenómeno *suprasegmental* está basado en dicha perspectiva. Así, algunos autores hablarán de reglas de generación (y *superposición*) del acento; otros, de la posibilidad, de la necesidad o de la imposibilidad de segmentar la entonación; y otros, aún, de la entonación como fenómeno no lingüístico, considerando el discurso oral idéntico al escrito y prescindiendo, por tanto, de su carácter *previo*, inseparable del habla. Desde esta perspectiva, se observa una cierta incompreensión del carácter continuo de la emisión de voz, y del carácter *artificial* de la lengua escrita.

También hemos visto cómo la fonación y la audición son fenómenos fundamentalmente tonales, y cómo la intensidad y la duración son parámetros no comparables, por distintas razones, al tono y al timbre. En primer lugar, el timbre consiste en la estructura de frecuencias secundarias (o armónicos) potenciadas a su paso por los resonadores, frecuencias que dependen directamente de la frecuencia fundamental (F_0) o tono: se trata, por tanto, de un fenómeno claramente tonal.

En segundo lugar, la intensidad es un parámetro coarticulatorio,

dependiente de la presión infraglótica del aire (que a su vez puede estar condicionada, por ejemplo, por la tensión emocional del individuo), y del que no depende directamente ninguna información estrictamente lingüística, pues su función principal es la de hacer audible la emisión de voz.

En tercer lugar, la duración es un parámetro referido directamente a los segmentos tímbricos, por lo que se trata de un parámetro, como dijimos, de *segundo grado*, no comparable a los otros tres, pues los segmentos tímbricos no existen independientemente *per se*, sino en función de nuestra propia segmentación, es decir, tras un proceso intermedio de abstracción entre la realidad acústica misma y el concepto *fonema*.

Hemos examinado el mecanismo fisiológico de la producción del tono, su percepción y el modo de extraerlo de una señal acústica, visualizarlo y analizarlo.

Finalmente, nos hemos detenido en la definición del fenómeno de la entonación, distinguiéndolo de los demás fenómenos suprasegmentales, y examinando los parámetros que realmente pueden tomar parte en él. Según nuestra perspectiva, la entonación está informada únicamente por el parámetro de tono, y consiste en las variaciones de F_0 relevantes lingüísticamente.

Hemos distinguido, además, los conceptos de F_0 , de melodía y de entonación, asignando sólo a este último carácter fonológico, y relacionando los otros dos en el sentido de que la F_0 es el parámetro físico (el valor absoluto en Hz) de cada punto *tonalmente* estable de la emisión de voz (no necesariamente un segmento tímbrico, como veremos más adelante) y la melodía es la sucesión de tonos entendida como la *relación* de los sucesivos valores absolutos, es decir, el esquema de variaciones relativas de tono a lo largo de la emisión de voz, sean relevantes o no lingüísticamente. La entonación, por tanto, se reduciría únicamente a los rasgos melódicos relevantes, que constituirán modelos significativos.

Hasta ahora, sin embargo, nuestra introducción al tema ha estado dirigida al establecimiento de los fenómenos desde un punto de vista marcadamente fonético y muy general. No hemos entrado a fondo, en realidad, en ninguno de los tópicos que constituyen la interpretación lingüística de la entonación. Así, el siguiente capítulo estará dedicado a recorrer las ideas y los autores más relevantes en el estudio de la entonación de los últimos años, en los que ha habido un verdadero *boom* en el interés por el tema, y en los que su estudio ha dado un vuelco fundamental, gracias a la diversificación de los enfoques teóricos, por un lado, y a la generalización, por otro, del método experimental, propiciada por unos avances técnicos impensables hace sólo unos años.

Sólo después de examinar el *estado de la cuestión* y plantear una discusión de conjunto emprenderemos nuestra propia descripción del fenómeno, en los capítulos siguientes.

2. PRINCIPALES ENFOQUES Y MÉTODOS DE ANÁLISIS DE LA ENTONACIÓN

2. PRINCIPALES ENFOQUES Y MÉTODOS DE ANÁLISIS DE LA ENTONACIÓN

2.1. Introducción

2.1.1. Acentos, niveles y configuraciones

El análisis de la entonación ha seguido, tradicionalmente, dos caminos bien diferenciados: el análisis *holístico* de la curva entonativa por las configuraciones o contornos significativos que la constituyen (llamado *análisis de configuraciones*) y el análisis *segmental* por niveles tonales, acentos y junturas en que se divide (llamado *análisis de niveles*).

En un principio, ambos métodos tienen una vida paralela e independiente y son menos incompatibles de lo que a primera vista pudiera parecer: son numerosos los autores modernos, por ejemplo, que los conjugan en sus trabajos, siguiendo la propuesta conciliadora de Daneš (1960) de analizar la entonación por "configuraciones de niveles". Dicha propuesta tiene sentido sólo como un intento de superar la controversia creada entre los dos métodos, controversia que, sin embargo, sólo apareció después de que el norteamericano Bolinger (1951) arremetiera decididamente contra el análisis de niveles; hasta ese momento, todos los autores suponían implícitamente que ambos métodos eran compatibles e igualmente eficaces, y que el suyo era el más adecuado para el dialecto del inglés que había escogido estudiar: así, se decía que en el inglés norteamericano había una serie de niveles tonales bien establecidos, cosa que no ocurría en el inglés británico (Ladd, 1980: 10).

A partir del citado trabajo de Bolinger, se inicia una polémica, a veces bastante dura, entre los seguidores de ambos métodos, sin duda propiciada por cierto componente extralingüístico (al parecer típico en las relaciones anglo-

norteamericanas) que se desprende de que el análisis de niveles sea un método norteamericano y el análisis de configuraciones un método británico: no es de extrañar, pues, que sean autores ajenos a tales planteamientos "nacionalistas", y generalmente europeos, los que proponen salidas conciliadoras y los que las adoptan.

Las diferencias entre los dos métodos, sin embargo, son ciertas y afectan a cuestiones de fondo, aunque en realidad no son muy numerosas: es decir, se puede prescindir de ellas con relativa facilidad, especialmente si no se entra en planteamientos muy detallados y lo único que interesa es mostrar un catálogo de "configuraciones de niveles".

El análisis de niveles, por una parte, constituye un intento muy serio de segmentar la entonación en unidades discretas y oponibles entre sí, semejantes a los fonemas; el análisis de configuraciones, por otra, considera la entonación como un fenómeno suprasegmental no divisible en unidades discretas, que sólo constituye unidades mayores (las configuraciones) que no se oponen entre sí sino que contrastan gradualmente. Es decir, en el primer método subyace la afirmación del carácter plenamente fonológico y lingüístico de la entonación, mientras que en el segundo el tema se deja en suspenso, se obvia, y más bien parece abonado a considerar la entonación como un fenómeno motivado, activo sólo en niveles expresivos no lingüísticamente codificados.

Crear que ambos métodos son directamente compatibles, por tanto, supone no entender bien el marco teórico en el que se desenvuelven: en Quilis (1981: 366), por ejemplo, se argumenta dicha supuesta compatibilidad aduciendo que el análisis de configuraciones incluye sucesiones de niveles tonales y junturas igual que el análisis de niveles, pues una melodía no es más que una sucesión de niveles tonales más altos, más bajos o iguales al anterior; es decir, se ignoran los planteamientos fonológicos de base que implica cada uno de los dos métodos y que afectan a la propia concepción del fenómeno entonativo y

se queda en la mera superficie, tomando como argumento teórico de igualación la evidencia física de que una melodía es una sucesión de valores relativos, sin darse cuenta de que es precisamente esa evidencia la que quieren describir ambos métodos desde puntos de vista divergentes.

El análisis de niveles, así, considera que la entonación está formada por una serie de "fonemas entonativos": los niveles tonales (de tres a seis niveles, según el autor que los describa), los acentos (primario, secundario, terciario y débil según la gradación más extendida) y las junturas (marcas de principio y final de frase, la más importante de las cuales suele ser la pausa). La sintaxis de tales "fonemas", que afectan a las sílabas, da lugar a los "morfemas entonativos", que serían ya unidades significativas y que afectarían a una cláusula o sintagma gramatical. Sin embargo, los análisis con este método rara vez han sobrepasado el nivel "fonológico" puramente descriptivo, de modo que no contamos realmente con "lexicones entonativos" generados por él: el tema del significado entonativo, por lo demás, nunca se ha tocado desde esta perspectiva de un modo sistemático.

El análisis de configuraciones, en cambio, se ha detenido mucho más en el significado de las configuraciones, a las que se considera unidades completas inanalizables y significativas: la unidad mínima de la entonación, por tanto, es directamente la configuración. En realidad, la configuración suele dividirse en "cabeza" (del principio de la frase al acento principal o "acento de frase") y "núcleo", que es la parte significativa (desde el "acento de frase" al final); pero ambas partes son inseparables y afectan siempre a una oración completa (y no a una cláusula, como en el análisis de niveles).

Una de las críticas más poderosas que ha recibido el análisis de configuraciones, así, ha sido el carácter no segmental, no claramente fonológico de su estudio, que contrasta con la función distintiva, claramente lingüística, que desempeña la entonación en algunos contextos: para los autores norteamericanos,

por tanto, acostumbrados a los análisis estructuralistas de Bloomfield, los británicos no han sabido ver con claridad esa realidad lingüística. Sin embargo, desde el marco teórico en el que se desenvuelve la fonología británica esta crítica no siempre ha sido comprendida: ya desde los trabajos de Firth (1948), por ejemplo, el nivel prosódico está bien integrado teóricamente en el sistema fonológico, que se compone, según su división, de "fonemas" y "prosodias", entre las que se cuentan el acento, algunas variaciones de timbre (en fenómenos como la armonía vocálica) y la entonación. Es un ejemplo más de la incomprensión mutua que reina entre ambas tradiciones lingüísticas y que, en fonología, podría concretarse en las figuras contrapuestas de Blommfield y Firth.

La mayor crítica recibida por el análisis de niveles, por su parte, no proviene del campo de la lingüística británica (que siempre se ha mostrado un tanto displicente al respecto) sino de la propia lingüística norteamericana: es Bolinger quien señala, en el artículo citado, que los niveles tonales no corresponden a ninguna realidad sensible determinada (cada nivel tonal no se corresponde, por ejemplo, con una gama de frecuencias, sino que su determinación es relativa, siempre distinta), que su establecimiento es totalmente variable (para unos autores hay tres, para otros cuatro, cinco o incluso seis niveles tonales) y que no constituyen realmente un sistema de unidades oponibles como los fonemas, sino de unidades graduales contrastantes (con lo cual buena parte de su potencialidad como unidades discretas de análisis desaparece, así como su homologación al resto del sistema fonológico: con lo cual, también, desaparece una diferencia fundamental con el análisis de configuraciones y es posible, entonces, hallar los puntos en común).

Pero además del carácter segmental o no del método de análisis y todo lo que en sí conlleva, hay otra diferencia muy importante entre los dos métodos, la segunda diferencia de fondo que conviene tener en cuenta y que fue puesta de relieve agudamente por Ladd (1980): el distinto tratamiento que recibe en

ambos métodos el "acento de frase", es decir, la cuestión de las relaciones que mantienen acento y entonación.

Para los seguidores tradicionales de ambos métodos, el acento y la entonación se consideran fenómenos lingüísticos claramente diferenciados, que no deben confundirse y que actúan independientemente el uno del otro. Con todo, al tratarse de fenómenos suprasegmentales siempre puede haber alguna coincidencia: así, en el análisis de niveles los acentos son unidades co-ocurrentes, paralelas, pero que no forman parte de la estructura melódica de la cláusula, y el llamado "acento de frase" (*sentence stress*) no es más que un mero portador de la principal inflexión tonal de la frase; es decir, el acento de frase y la inflexión tonal coinciden normalmente, pero podrían no coincidir porque son fenómenos distintos.

Para el análisis de configuraciones, en cambio, y aunque mantiene la misma postura general de distinguir ambos fenómenos, el acento de frase es el "núcleo" de la oración entonativa, y en él ocurre la inflexión tonal porque debe ocurrir en él: no es un fenómeno casual de coincidencia sino de necesidad; es decir, el acento de frase o núcleo del contorno es un fenómeno entonativo, y no acentual.

Los autores norteamericanos, entonces, entienden que los británicos no distinguen entre acento y entonación, y nuevamente se aviva la controversia: en realidad, la única diferencia es que tradición británica considera el acento de frase como una parte de la entonación y la tradición norteamericana lo considera teóricamente ajeno al fenómeno entonativo, aunque admita que suelen coincidir el acento primario y la inflexión tonal; es decir, en los dos métodos se distingue entre acento y entonación, aunque el límite entre ambos fenómenos se establece en un lugar distinto. De nuevo se trata de una especie de malentendido, digámoslo así, terminológico, aunque se trate de una diferencia teórica real.

Con el resto de los fenómenos acentuales ambas tradiciones están

básicamente de acuerdo: además del acento principal (o primario, o acento de frase, o núcleo) hay otros acentos no nucleares en la frase, que contrastan gradualmente (secundario, terciario y débil en el análisis de niveles y sólo secundario y débil en el análisis de configuraciones).

La gradación de acentos en cuatro niveles de intensidad, sin embargo, característica del análisis de niveles (Chomsky & Halle -1968- habían postulado incluso una graduación indefinida en el contraste acentual), ha sido puesta en entredicho en numerosas ocasiones: Halliday (1967), por ejemplo, se pregunta irónicamente quién sería capaz de oír alguna diferencia entre un acento "terciario" y un acento "débil", o si podrían ponerse de acuerdo dos autores distintos en una misma transcripción.

Por otra parte, la evidencia instrumental de que el parámetro fundamental en la acentuación de sílabas no es la intensidad, como creían todos los autores de ambas tradiciones, intrujo un nuevo elemento de discordia: por un lado, se daba una parte de razón al método británico, pues en él ya se contemplaba el carácter entonativo del núcleo acentual, que ahora aparece realmente como únicamente una inflexión tonal; por otro, sin embargo, obligaba a replantear el tema de los acentos secundarios, pues si igualmente se trata de acentos con una inflexión tonal hay algo que obviamente falla en la descripción tradicional.

El resultado ha sido la aparición de dos nuevos métodos de análisis o, más bien, de dos nuevas perspectivas de análisis al abrigo de los dos métodos tradicionales. La primera se ha integrado completamente en el análisis de configuraciones de la tradición británica, y consiste en distinguir el acento melódico (el núcleo entonativo) del acento no melódico (sea debido este a la variación tonal, a la duración o a la intensidad, pero sin una función entonativa relevante), y en considerar dicho acento melódico como el elemento más informativo y fundamental del contorno.

Es la perspectiva, básicamente, de Bolinger, cuya ideas generales sobre

la entonación no han encontrado mucho eco entre los distintos autores, pero cuyo ataque al análisis de niveles, por una parte, y su propuesta de análisis basado en el acento melódico ha sido fundamental en el desarrollo moderno del análisis de configuraciones y en la conformación de la tradición británica actual (se trata, con mucho, del autor más influyente en dicha tradición, a pesar de tratarse de una excepción teórica y, además, de un autor norteamericano). Bolinger distingue, pues, entre *accent* y *stress*, entre acento melódico, con una función entonativa fundamental, y acento dinámico (aunque sea tonal): como Bolinger, la mayoría de los autores británicos (y europeos) actuales. La perspectiva de Bolinger la examinaremos con cierto detenimiento más adelante.

La otra perspectiva aparecida como consecuencia de la evidencia de que el acento es un fenómeno tonal ha sido el análisis basado en los esquemas rítmicos o, más bien, métricos: en concreto, la fonología métrica de Liberman & Prince (1977).

Ya Halliday (1967) había propuesto un análisis basado en el esquema rítmico de la frase y atribuía, por ejemplo, la prominencia del acento principal no a la inflexión tonal sino a un efecto rítmico; antes que Halliday, Householder (1957) hablaba del efecto que produce un determinado esquema rítmico, que hace oír "cosas que no están ahí". Lo que hacen Liberman & Prince es desarrollar esta idea y llevarla a sus últimas consecuencias, hasta afirmar que, en realidad, el principal parámetro para la percepción del acento es la estructura métrica (o rítmica), y no el tono; aunque, como el acento principal suele ir asociado a la inflexión tonal de la frase, en los experimentos en los que no se tenga en cuenta el factor rítmico puede parecer (equivocadamente) que la inflexión tonal sea determinante. Es decir, según esta perspectiva no sólo permanece intacta la tradicional separación entre acento y entonación, sino que se puede incluso negar el carácter tonal del acento: una jugada maestra de apego a las tradiciones, combinada con la renovación más radical. En su momento,

también examinaremos con más detalle esta perspectiva.

En resumen, la cuestión del acento de frase o núcleo como parte de la entonación o como fenómeno aparte constituye, de entrada, una diferencia notable entre los dos métodos de análisis tradicionales; con el establecimiento experimental del carácter fundamentalmente tonal del acento, además, se deshace el prejuicio del "accento de intensidad", que tan bien iba para distinguir claramente ambos fenómenos, y surgen dos nuevas perspectivas: el análisis basado en el "accento melódico" (*accent*) de Bolinger (bien integrado en el análisis de configuraciones y que admite y aprovecha el carácter tonal del acento), y el análisis métrico de Liberman & Prince (que constituye un cambio de rumbo en el análisis de niveles y que no admite el carácter tonal, sino el carácter rítmico del acento).

En la actualidad, por tanto, las diferencias entre los dos métodos de análisis de la entonación, que parecían limadas con el tiempo y el concurso de diversos autores empeñados en compatibilizarlos (Daneš -1960-, Vanderslice & Ladefoged -1972-, Stockwell -1972-, Quilis -1975-, etc.) nuevamente se han profundizado, esta vez hasta límites insospechados, pues incluso parten de descripciones físicas diferentes, y no hay acuerdo ni siquiera en el parámetro físico que informa a un fenómeno de la importancia del acento.

2.1.2. *Panorámica previa*

Aunque en los manuales suele simplificarse mucho diciendo que hay sólo dos métodos de análisis de la entonación, el análisis de configuraciones y el análisis de niveles (y que es justamente lo que hemos hecho nosotros hasta este momento), en realidad las perspectivas e incluso los métodos de análisis están mucho más diversificados, y el panorama es mucho más rico.

Sí puede hablarse de dos tradiciones mayores: la tradición británica y la

tradición norteamericana, que básicamente coinciden con los dos métodos citados, pero en la que se encuentran inmersos otros autores de relevancia que no siguen exactamente el mismo método, o que siguen otros caminos.

Así, en la tradición británica, fijada por Jones (1909), Palmer (1922) y Amstrong & Ward (1926), incluimos también a Bolinger, que, como hemos dicho, es un autor norteamericano y que no es exactamente un abanderado del análisis de configuraciones tradicional, como a veces se ha dicho: más bien, merecería un capítulo aparte, pues su planteamiento teórico original difiere en mucho del de los autores británicos (que, por otra parte, han sabido enriquecerse muy bien con él: por ejemplo, en el exponente más reciente y completo de la tradición, el manual de entonación de Cruttenden, 1986). Otros autores británicos, como Halliday, también representan un punto de ruptura en la tradición, por sus planteamientos originales, y fueron muy influyentes durante algún tiempo.

En la tradición norteamericana, por otra parte, el panorama está mucho más diversificado, pues el análisis de niveles es un método claramente estructuralista, inspirado en Bloomfield y fijado por Wells (1945), Pike (1945) y, sobre todo, Trager & Smith (1951); sin embargo, la irrupción de la lingüística generativa supuso un cambio de planteamiento radical, aunque siempre dentro de unos límites asumibles por el método. Chomsky & Halle (1968), por ejemplo, adoptan en lo tocante a la entonación una perspectiva que sigue muy de cerca estos trabajos estructuralistas (como hicieron con el binarismo de Jakobson, al que también siguen muy de cerca) aunque sus planteamientos teóricos eran muy novedosos. Sin embargo, ya Lieberman (1967) se había enfrentado al análisis de la entonación desde un planteamiento generativista, en el que se negaba el carácter fonológico de la entonación en virtud de su carácter universal, innato y dependiente de la estructura sintáctica. Desde el mismo generativismo, pero con planteamientos radicalmente distintos, Liberman & Prince (1977) suponen

un cambio de perspectiva fundamental, que luego ha seguido Pierrehumbert (1987), imprimiendo un nuevo impulso. El panorama de la tradición norteamericana, por tanto, aunque mantiene una continuidad interesante, resulta también muy diverso.

Pero hay otras tradiciones que también nos interesan: en concreto, lo que ha dado en llamarse *escuela holandesa*, cuyas aportaciones figuran entre las más destacadas de los últimos años y que no pueden considerarse exactamente en la línea del análisis de configuraciones, como a veces se ha dicho: es el mejor exponente del método instrumental y perceptivo en el análisis de la entonación, llevado a cabo de un modo sistemático por primera vez en la historia, y sus planteamientos teóricos son originales y novedosos. Su estandarización de contornos, por lo demás, constituye un avance considerable en la fijación de una metodología experimental eficaz con la entonación. Merecen, por tanto, un capítulo aparte, pues constituyen, además, el único ejemplo en el que la interpretación fonológica de la entonación está claramente condicionada por la descripción fonética y perceptiva, y no al revés.

Finalmente, y como es obvio, nos interesa describir nuestro entorno más inmediato, la llamada *escuela española*, cuya aportación central y más original es la doctrina de Navarro Tomás (1918, 1944), retomada una y otra vez por los distintos autores.

En realidad, los trabajos de Navarro Tomás importan, más que por la doctrina propiamente dicha, por el notable corpus entonativo que se reúne en ellos, corpus que ha servido de base para todas las investigaciones posteriores en español, e incluso en otras lenguas peninsulares (como el catalán: cfr. Salcioli, 1988). La selección y el orden de los datos que ofrece, por lo demás, implican una cierta doctrina, como veremos, y en eso consiste, básicamente, la enorme influencia que ha tenido y sigue teniendo este autor sobre los especialistas españoles.

Al margen de Navarro Tomás, cuyo análisis puede considerarse un tipo de análisis de configuraciones, el panorama de la tradición española resulta un tanto disperso: Stockwell, Bowen & Silva-Fuenzalida (1956), Bowen & Stockwell (1960), J. H. Matluck (1965), etc., introducen el análisis de niveles en español, rápidamente aceptado por algunos autores españoles (incluso por la Real Academia: 1973), que A. Quilis (1981, 1993) intenta compaginar con el análisis de configuraciones. Antes, R. E. Hoyos Andrade (1970), y tras un análisis espectrográfico de los patrones (por niveles) propuestos por Matluck, reivindica una vuelta al método por configuraciones.

Los autores más recientes, en efecto, protagonizan una vuelta al análisis de configuraciones: V. Salcioli (1988) realiza un trabajo sistemático de análisis instrumental, siguiendo los esquemas entonativos de Navarro Tomás, para la lengua catalana; M. Cabrera-Abreu (1991), desde otra perspectiva, realiza un intento de adaptación del análisis de configuraciones, según la tradición británica, a la entonación española (y, en este sentido, su trabajo pertenece más bien a la tradición británica que a la española); finalmente, J. M. Garrido (1991) se aplica a estandarizar los contornos entonativos del español, siguiendo también la tipología fraseológica de Navarro Tomás, con el objetivo de implementar la síntesis de voz y el reconocimiento automático del habla en español.

Este es el panorama que examinaremos a lo largo del presente capítulo, deteniéndonos sólo en las aportaciones más relevantes, con la intención de presentar sistematizados, al final, las ideas y los elementos de análisis que han empleado los distintos autores en su estudio de la entonación, como punto de partida y como guía para nuestra propia descripción del fenómeno.

2.2. La tradición británica

2.2.1. El análisis por contornos tradicional

La tradición británica de análisis de la entonación puede decirse que comienza con la publicación en 1909 del libro de Daniel Jones *Intonation Curves*. Jones grabó un nutrido corpus de frases en inglés y francés, y se dedicó a escucharlas y transcribirlas detalladamente sobre un pentagrama, en forma de curvas. Tales curvas se correspondían de un modo impresionista con la evolución de la F_0 a lo largo de la frase, suponiendo que la entonación era como una melodía similar a las melodías musicales; sin embargo, y como vimos, las representaciones que nos suministran actualmente todos los visualizadores instrumentales están fuertemente influenciadas por este tipo de representación y nos ofrecen justamente eso, una curva continua que muestra con toda exactitud la evolución de la F_0 a lo largo de la frase. El trabajo de Jones, pues, supone un hito en la concepción de la entonación como una sucesión de valores que constituyen una *curva*.

Publicado casi sin comentarios, no intenta ninguna interpretación lingüística de las entonaciones que describe ni glosa los significados: se trata de un trabajo puramente fonético (posteriormente, Jones asumirá el análisis de Armstrong & Ward y en otros trabajos lo aplicará a su propio corpus). Teniendo en cuenta, además, que se usa un magnetófono para escuchar repetidamente las entonaciones que van a transcribirse (lo cual supone ayudarse de un instrumento tecnológico, como en la moderna fonética instrumental), *Intonation Curves* se nos aparece, pues, como una obra de una gran modernidad.

Con todo, la importancia del libro estriba en que describió un corpus sobre el que posteriormente se podrían llevar a cabo sucesivas interpretaciones lingüísticas: en el apartado 1.2.2., en el capítulo anterior, decíamos que sólo

cuentan con una descripción fonológica completa las lenguas que habían sido descritas fonéticamente con anterioridad. Así, si los trabajos de Sweet (1878, 1892) sirven de base y corpus previo para el desarrollo de la fonología segmental del inglés, el trabajo de Jones es la descripción fonética previa necesaria para todo el desarrollo posterior de la tradición británica de análisis de la entonación (y, en un principio, también de la norteamericana). En España, por ejemplo, veremos que esta función histórica la desempeñan los trabajos de Navarro Tomás.

Después del trabajo de Jones, hay otros dos libros fundamentales en la fijación del análisis de configuraciones británico: uno es el libro de Palmer *English intonation, with systematic exercises* (1922); y el otro es *Handbook of English intonation* (1926), de Armstrong & Ward, cuyo análisis aplicó el propio Jones en los años treinta, en la 3ª edición de *An outline of English phonetics* (1918; 1932³).

Nótese que en ambos casos se trata de manuales para enseñar la entonación inglesa correcta (la *Received Pronunciation*), dirigidos a un público culto, y que por tanto no se trata de trabajos de investigación científica. Este es un hecho de enorme trascendencia en el análisis de configuraciones, pues, en contraposición al análisis de niveles, nunca se planteó sistemáticamente más descripción de la entonación que la estrictamente utilitaria, ni más sistema de análisis y representación que el pedagógicamente eficaz.

Es más: en realidad, el análisis de configuraciones no es un sistema de análisis de la entonación, sino un mero sistema de representación, en el que es fácil seguir las entonaciones por parte del lector aprendiente y es fácil reproducirlas intuitivamente. Esta es la característica fundamental del análisis de configuraciones: su aplicación didáctica inmediata e intuitiva. Muy lejos, pues, del sistema de análisis norteamericano por niveles, cuya finalidad es establecer el carácter fonológico y segmental de las unidades entonativas, pero no servir

para la enseñanza del inglés correcto y de sus entonaciones típicas.

En el manual de Armstrong & Ward (1926), por ejemplo, se establece la descripción sólo de dichas entonaciones típicas, de "las formas más simples de la entonación usadas en la conversación", fiel a su carácter didáctico. Así, las demás variedades entonativas no se consideran esenciales, y por tanto no se contemplan.

Para Armstrong & Ward el acento es un fenómeno de intensidad, y recae sobre las palabras más importantes de la frase; sin embargo, se postula una relación muy estrecha entre acento y entonación, y se distingue entre "prominencia" (*prominence*), que es el núcleo de la frase entonativa, y "acento" (*stress*), siguiendo la definición de Jones (1918), para quien la "prominencia" es un efecto perceptivo que tiene que ver con la entonación. Entre los trabajos de Jones, pues, y Armstrong & Ward, las influencias son múltiples y en las dos direcciones.

En principio, Armstrong & Ward consideran la oración como la unidad entonativa fundamental, especialmente en la conversación, y en lo que se refiere a su rendimiento semántico; sin embargo, también contemplan que las oraciones largas puedan incluir más de un grupo de entonación, llamados entonces "grupos de sentido" (*sense-group*). Así, la unidad mínima de entonación sería el *sense-group*, que puede ocupar una oración completa o una parte coherente de la oración. La división en *sense-group*, sin embargo, no es fija y depende del hablante, del estilo que emplee, etc.

Cada *sense-group* puede tener una de las dos formas o contornos en que se divide la entonación: el contorno característico de las oraciones afirmativas e imperativas, *TUNE I*, y el contorno característico de las oraciones interrogativas absolutas (preguntas que requieren respuesta de sí o no) y de las oraciones de incertidumbre, el *TUNE II*.

El *TUNE I* comienza en un tono medio y continúa así, con alguna variación ascendente en las palabras tónicas, hasta el final del grupo, en el que el tono desciende (*fall*) rápidamente. Por su parte, el *TUNE II* comienza en un tono medio o alto, que desciende gradualmente a lo largo del grupo y al final asciende bruscamente (*rise*), o bien se mantiene (*level*).

En oraciones largas, con varios *sense-group*, el *TUNE II* se usa en los grupos no terminales, excepto en algunos casos, como en las enumeraciones, en los que los grupos no terminales usan el contorno *TUNE I*.

El trabajo de Armstrong & Ward fijó, por tanto, la división tradicional entre las dos entonaciones supuestamente universales: la entonación afirmativa, de final descendente, y la entonación interrogativa, de final ascendente. Todos los autores, en adelante, se referirán a la división entre *TUNE I* y *TUNE II* como el paradigma de las dos entonaciones, sigan o no el análisis de configuraciones. En este sentido, la influencia del trabajo traspasa con mucho los límites de la tradición británica, pues en él se establece una de las distinciones recurrentes en todos los autores que han estudiado la entonación, así como una de las afirmaciones básicas más generalizadas y cuya aceptación previa se supone.

Otro de los aspectos del trabajo más influyentes, aunque ya en el ámbito de la escuela inglesa, es el tipo de transcripción, que se ha erigido en el más importante de la tradición, y al que se han incorporado elementos de otros autores, fundamentalmente de Palmer. La transcripción empleada por Armstrong & Ward es distinta de la que usaba Jones: en lugar de una línea inserta en un pentagrama, los autores optan por marcar sobre la línea de escritura fonética una serie de puntos y rayas, cuya altura relativa intenta mantenerse de un modo intuitivo y sin seguir ninguna pauta definida como el pentagrama.

Cada punto o raya corresponde a una sílaba: los puntos, a las sílabas átonas; las rayas, a las sílabas tónicas. Con las rayas se representa la inflexión

tonal de la sílaba tónica, con una subida, una bajada o el mantenimiento del tono. Esta transcripción, mediante puntos y rayas sobre cada sílaba, es, en esencia, la que se ha mantenido a lo largo de toda la tradición británica, y difiere mucho de la transcripción de Jones, que se ha mantenido, sin embargo, en las demás tradiciones que usan el análisis de configuraciones, como la tradición española.

El manual de Palmer (1922) puede considerarse la obra más influyente y la que más ha ayudado a crear lo que conocemos como tradición británica. Cada oración, según Palmer, constituye un "grupo de entonación" (*tone-group*), caracterizado por un "modelo de entonación" (*tone-pattern*). Así, si para Armstrong & Ward una oración podía contener varios *sense-group*, para Palmer la unidad mínima es, directamente, la oración.

Todo *tone-group* se divide en tres partes: la "cabeza" (*head*), que va desde el principio de la oración hasta el *nucleus* (para Armstrong & Ward, *prominence*); el "núcleo" (*nucleus*), que es la parte más importante y definitoria del grupo, porque en él ocurre la inflexión tonal (el *tone-pattern*); y la "coda" (*tail*), que va del núcleo al final de la oración.

La división de Palmer ha sido aceptada y enriquecida posteriormente por numerosos autores, que, o bien añaden una *pre-head* (lo que Cruttenden veremos que llama "anacrusis", tomando el término de Chao, 1932), o bien eliminan la separación entre el núcleo y la coda. Curiosamente, los seguidores más directos de Palmer, Schubiger (1958) y Kingdon (1958), son los que añaden más partes a su división del contorno, y hablan de *pre-head*, *head* y *body* donde Palmer hablaba sólo de *head*.

En la figura 25 (tomada de Ladd, 1980: 16) puede observarse cómo todos los autores (incluso los autores no pertenecientes a la tradición británica) están de acuerdo al menos en separar el núcleo, como parte más informativa, del resto

del contorno: el propio Ladd, que opta por una división más detallada del mismo (*pre-head, head, nucleus, tail*), llama la atención sobre el indudable interés que tiene distinguir, en todo caso, *pre-head* y *head* de *nucleus* y *tail*, a la manera de Chao (que los distingue como *head* y *body*) y decide llamarlos, genéricamente, *head* (*pre-head* + *head*) y *nucleus* (*nucleus* + *tail*).

	De	Tooqve	vill	stud	ied	at	Prince	ton
Kingdom (1958)	Prehead	Head	Body				Nucleus	Tail
Shubiger (1958)								
O'Connor & Arnold (1961)	Prehead	Head				Nucleus	Tail	
Crystal (1969)								
Ladd (1980)	Head				Nucleus			
Chao (1932)	Anacrusis	Main Head				Nucleus	Tail	
	Head				Body			
Palmer (1922)	Head				Nucleus	Tail		
Hockett (1955/58)	Remainder Pendant				Intonation Head			
Pilte (1945)	Precontour				Primary Contour			

Figura 25 - División del contorno, según diversos autores.

En el *nucleus* (que en la tradición norteamericana se llamará *sentence-stress*) está la parte más informativa del grupo de entonación: el *tone-pattern*. Volviendo, pues, a Palmer, vemos que este describe seis *tone-pattern* fundamentales (v. figura 26).

El modelo 1) es el característico de las frases asertivas, de las declaraciones, de las órdenes, de las preguntas retóricas y de las preguntas con partícula interrogativa (obsérvese, pues, que para Palmer las preguntas relativas pertenecen a la misma categoría que las declarativas: en la mayoría de los autores posteriores, sin embargo, se insiste en hacer una distinción que sólo puede justificarse tomando como parámetro fundamental el *significado semántico* de la oración, pero no la forma de la entonación, ni el *significado entonativo*; volveremos a este tema más adelante).



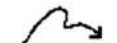
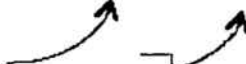


1	
2	
3	
4	
5	
6	

Figura 26 - Los seis *tone-pattern* de Palmer (1926).

El modelo 2) es el propio de las oraciones asertivas, así como de las oraciones de una sola palabra.

El modelo 3) es el de las afirmaciones y los comentarios.

El modelo 4), que tiene dos formas posibles, es el característico de las preguntas absolutas (sí / no) y de las preguntas eco.

El modelo 5) es el de las concesiones y los contrastes.

El modelo 6), finalmente, y aunque claramente descrito por Palmer y empleado a lo largo de su manual en numerosas ocasiones, no parece corresponder a ningún significado oracional definido, y más bien depende del contexto en el que se encuentre.

En resumen, Palmer ofrece una suerte de "lexicón entonativo" cuyo

significado y cuya función se supone que depende en parte del contexto en el que la entonación ocurre, pero que es lo suficientemente estable como para poder identificar y repetir los modelos de entonación.

En realidad, Palmer únicamente se aplica a deducir algunos de los usos más frecuentes que ha detectado en su análisis de la entonación en la conversación, como buen fonetista de principios de siglo, y no le interesan las generalizaciones lingüísticas que puedan hacerse de un corpus que ofrece únicamente para ayudar a mejorar la pronunciación y la dicción del inglés *R.P.* De ahí que a veces se pierda en consideraciones semánticas y pragmáticas muy particulares, analizando alguna frase, o que sea tan dado a relatar toda la casuística que conoce de cada entonación.

En lo sucesivo, sin embargo, los diversos autores van a intentar hacer generalizaciones lingüísticas de sus observaciones, pero siguiendo con el mismo método de detallar cada contexto y cada uso particular de cada entonación. Así, los tratados de entonación aún hoy son lo más parecido a los tratados de fonética novecentistas, en los que se mezclan afirmaciones y reglas genuinamente lingüísticas, de cierto empaque teórico, con observaciones y anotaciones detalladas de cada variación encontrada, sin ninguna trascendencia teórica, y opiniones sobre significados y usos posibles, sin más patente científica que la propia intuición y la experiencia del entonólogo.

El estilo disperso y detallado, y a menudo divagatorio, de los trabajos que tratan de la entonación, pues, tiene su origen en una mala interpretación de las intenciones y los métodos de autores no formados en la moderna lingüística sino en el positivismo lingüístico del siglo XIX: Palmer elabora un manual de entonación con una mentalidad puramente didáctica, y lo adereza de agudas observaciones sobre tal o cual uso concreto y, por inercia, los sucesivos autores adoptan sus elementos pedagógicos como postulados teóricos y su desenfado metodológico (que procura informar y amenizar al lector-estudiante) como la

metodología a seguir.

Jassem (1952), Schubiger (1958) y Kingdon (1958), puestos a menudo como ejemplo de un análisis de configuraciones tradicional, siguen en lo esencial a Palmer, aunque ya desde una perspectiva estrictamente científica. La transcripción empleada por O'Connor & Arnold (1961), Crystal (1969) y, como veremos, Cruttenden (1986), consiste en una sabia mezcla de la transcripción propuesta por Armstrong & Ward y la propuesta por Palmer: puntos en las sílabas átonas, rayas (inflexiones tonales) en las sílabas tónicas y los *tone-pattern* (llamados *nuclear tones*) en el núcleo del contorno.

En el apartado anterior vimos cómo los fenómenos del acento y la entonación tienen en la tradición británica una definición característica: el acento se considera de intensidad, pero el núcleo del contorno es un fenómeno tonal, entonativo, un fenómeno no propiamente de acento sino de prominencia, y en la inflexión tonal que en él sucede está la parte informativa del contorno.

Para Halliday (1967), sin embargo, y como ya comentamos, la prominencia silábica no es un efecto únicamente tonal sino también rítmico: cada "grupo de entonación" (*tone-group*) está dividido, según su perspectiva, en varios "pies rítmicos", cuya primera sílaba es siempre una sílaba acentuada, una "saliente" (*salient*), y cuyas otras sílabas son débiles; una o dos sílabas "salientes" de cada grupo de entonación son sílabas "tónicas" (término que, por tanto, equivale al *sentence-stress*, al *prominence* o al *nucleus* de otros autores).

Así, la jerarquía acentual, según Halliday, sería de tres grados: sílabas "salientes tónicas", sílabas "salientes no tónicas" y sílabas débiles. Las tónicas lo serían por efecto de una inflexión tonal, o bien de una mayor duración; las sílabas salientes únicamente por efecto de la estructura rítmica del grupo: cada pie rítmico comienza por una saliente; las sílabas débiles, finalmente, estarían determinadas por contener vocales reducidas (más adelante nos referiremos al

fenómeno de la reducción vocálica con cierto detenimiento) y por su duración relativa dentro del pie rítmico, según el número de sílabas que lo compongan.

Otros aspecto de la entonación tratado por Halliday es la relación que mantiene o puede mantener con la sintaxis: en general, la tradición británica se limita a interpretar los usos concretos de los contornos, y a menudo interviene en la descripción algún elemento gramatical. Sin embargo, no es un tema planteado específicamente por ningún autor hasta que Halliday lo convierte en el tema central de su investigación.

En la mayoría de los autores existe una cierta tendencia a considerar la entonación como un fenómeno no estrictamente lingüístico, sino más bien motivado, meramente expresivo y no arbitrario. Esta consideración no fonológica de la entonación supone, lógicamente, su dependencia de los niveles de significación propiamente lingüísticos: fundamentalmente, del nivel léxico-gramatical y semántico. Obviamente, no comienza a hablarse de nivel pragmático hasta finales de los años setenta, y sólo trabajos como el de Ladd (1980) le conceden cierta relevancia; sin embargo, en todos los autores subyace una mentalidad claramente pragmática, sobre todo en los abundantes comentarios sobre usos particulares de los contornos entonativos y sobre sus significados en cada contexto a los que nos referíamos arriba.

Para Halliday, la relación que existe entre la gramática y la entonación es una relación de interdependencia: considera la entonación como un sustituto ocasional de ciertas estructuras gramaticales, como por ejemplo la estructura interrogativa, que puede estar marcada por el orden de las palabras, por la presencia de una partícula interrogativa o por una entonación de pregunta.

Así, sólo los constituyentes rítmicos más los constituyentes sintácticos del enunciado marcarían su esquema acentual y entonativo. La entonación, desde este punto de vista, funcionaría a un nivel paralelo y complementario del nivel

gramatical y no tendría ninguna relevancia como fenómeno lingüístico aislado: no tendría, por ejemplo, ningún tipo de autonomía semántica.

Sin embargo, hemos visto que Palmer, y como él otros autores, se aplica a describir el significado característico de cada inflexión tonal característica; es lo que hemos llamado "lexicón entonativo", que puede considerarse muy definitorio de lo que es y lo que pretende el análisis de configuraciones: esto es, un sistema de representación y descripción de la entonación enfocado prioritariamente a su enseñanza. Así, importa describir formalmente el contorno, tanto como describir su uso más corriente, sus significados posibles, los casos especiales en los que se encuentran, etc.

El trabajo de Ladd (1980), muy bien documentado y totalmente integrado en la tradición de la que se siente partícipe, se aplica, precisamente, a intentar discernir de qué tipo es el significado que aporta la entonación, despreciando displicentemente a los autores que le niegan autonomía semántica. Después de hacer un recorrido por los diversos autores y de haber transitado sistemáticamente, tal vez por primera vez, el aspecto específicamente semántico de los contornos entonativos, la conclusión a que llega podría considerarse como un resumen del tratamiento que ha tenido el significado de la entonación en la tradición británica: según Ladd, "el significado entonativo es ... *léxico, gradual, fonestésico, y relacional*" (1980:205) (v. 4.2.2.2.).

Léxico, por cuanto cada contorno contiene un significado unitario, como si se tratase de una palabra: el significado de los contornos entonativos, por tanto, es semejante al significado conceptual, aunque sin referentes concretos en la realidad.

Gradual, porque no es posible hacer oposiciones discretas ni binarias entre contornos, ni entre sus significados.

Fonestésico, esto es, metafórico, aproximativo, porque no corresponde a

la conceptualización de realidades definidas, como hemos dicho, sino a la conceptualización de actitudes con respecto al propio enunciado y a los interlocutores en la conversación.

Y, finalmente, *relacional*, no *segmental*, porque implica a varios factores interrelacionados, factores y unidades suprasedgmentales no definibles en términos segmentales.

2.2.2. *La perspectiva de D. L. Bolinger*

Un caso especial dentro de la tradición británica de análisis de la entonación lo constituye D. L. Bolinger, autor norteamericano que debe incluirse en dicha tradición por varias razones: en primer lugar, fue el primero en realizar una crítica fundamentada al sistema de análisis por niveles propio de la tradición norteamericana, como vimos; en segundo lugar, aboga decididamente por el análisis de configuraciones desde los años 50, y ya en los 60 aparece como un claro inspirador de los autores británicos (un ejemplo relevante es Crystal, 1969); en tercer lugar, en Bolinger las principales tendencias que se apuntan en los autores británicos muy a menudo se llevan a sus últimas consecuencias, como por ejemplo el carácter no arbitrario de la entonación, su relación con la sintaxis, con el gesto, con la emoción, así como la propia metodología empleada, más "novecentista" que en ningún otro autor; finalmente, Bolinger representa un momento importante de inflexión en la tradición británica, pues con él se inicia una línea de investigación no dedicada ya a la mera enseñanza del inglés correcto (objetivo de los autores británicos anteriores) sino a la profundización del conocimiento científico, cuyos productos, por tanto, no son manuales utilitarios de dicción sino trabajos científicos para especialistas, al estilo norteamericano: después de Bolinger, los autores británicos se irán decantando, progresivamente, no hacia la enseñanza de la pronunciación correcta en su

aspecto entonativo sino hacia la pura investigación científica.

Las opiniones de Bolinger sobre la entonación, sin embargo, no han tenido una acogida demasiado favorable entre los especialistas, a pesar de ser un autor enormemente influyente en otros aspectos: básicamente, en su crítica al análisis de niveles (que ya examinamos someramente) y en su concepto de *accent*, que Ladd (1980), por ejemplo, considera una aportación de primer orden al análisis de la entonación.

La obra de Bolinger es muy extensa, y a menudo sus opiniones están formuladas de modos muy distintos en distintos artículos, pues abarcan desde mediados de los 50 a final de los 80: treinta años publicando y aportando nuevas ideas. En 1986 publicó *Intonation and its parts*, un compendio de su teoría de la entonación, en el que quiso reunir de un modo sistemático sus opiniones sobre el análisis formal y "lingüístico" de la entonación. En 1989, *Intonation and its uses*, subtítulo *Melody in grammar and discourse*, viene a ser la continuación natural del libro anterior, en el que no le interesa ya poner orden a sus ideas sino ofrecer un vasto panorama de datos y comentarios sobre usos particulares, en la mejor tradición (como apuntábamos en el apartado anterior) de la fonética "novecentista": según Ladd (1990), el segundo libro de Bolinger es "un gran apéndice de datos, un compendio de soporte para las afirmaciones teóricas de *Intonation and its parts*".

En *Intonation and its parts*, Bolinger comienza distinguiendo entre los tres usos lingüísticos que puede tener el tono: su uso en las lenguas tonales para distinguir palabras (como en chino-mandarín), su uso en la entonación (en las lenguas no tonales) y su uso en el *accent* (término que podría traducirse por "acento melódico", como propone I. Mascaró en su traducción del manual de Cruttenden, 1986, y que en inglés debe distinguirse de *stress*, "acento dinámico", que suele quedar para referirse al "acento de intensidad": nótese que todos los autores mencionados mantienen este uso del término *stress*, con el que queda

salvaguardada su identidad como acento "de intensidad", según la tradición, a despecho de las evidencias experimentales según las cuales todo "accento" es producto de una inflexión tonal; entrevemos aquí, una vez más, la pugna por distinguir a toda costa los fenómenos del acento y la entonación).

El *accent* es un mecanismo que consiste básicamente en una inflexión tonal (aunque también pueden intervenir otros factores, como la duración, el énfasis articulatorio de las consonantes de la sílaba, o incluso la intensidad) con la cual se pone de relevancia la palabra subjetivamente más importante para el hablante o la parte más informativa de la frase (el rema). "El *accent* es la entonación al servicio del énfasis (...) hace que ciertas sílabas se superpongan en varios grados a otras, revelando a nuestro oyente lo importante que es para nosotros las palabras que contienen, y revelando también (...) lo importante que es el mensaje que contienen" (1989: 3).

Para Bolinger también existe una clara jerarquía acentual, que sin embargo no coincide con las ideas que hemos visto hasta ahora: según su perspectiva (1964), en inglés hay sílabas "largas" (*long syllable*), en las que puede recaer un acento de palabra (*stress*) porque contienen una "vocal plena" (*full vowel*), y sílabas "breves" (*short syllable*), sobre las que no puede recaer acento alguno porque contienen una vocal inestable, a menudo una "vocal reducida" (*reduced vowel*).

Las sílabas largas pueden estar acentuadas o no (*stressed syllable* vs. *unstressed syllable*), según el esquema rítmico de la frase. El *stress* es un fenómeno, generalmente, de intensidad. Finalmente, las sílabas *stressed* pueden contener o no un *accent*, y dar lugar, por tanto, a una *accented syllable* o bien a una *unaccented syllable*: el *accent*, como hemos dicho, es un fenómeno estrictamente tonal (similar al *nucleus*, etc., de otros autores), y tiene que ver ya no con el "accento" (*stress*) sino con la entonación.

Del mismo modo que Palmer define una serie de *tone-patterns* y otros autores hablan de *nuclear tones*, Bolinger da una lista de las inflexiones tonales posibles de un *accent* en inglés: los "perfiles" (*profiles*), a los que considera "las unidades MORFOLÓGICAS mínimas de la entonación", cuya sintaxis da lugar a los contornos, que son las "unidades SINTÁCTICAS" de la entonación (1986: 140 -el subrayado es del propio Bolinger-).

El *profile*, según lo define, es la inflexión tonal, el movimiento entonativo límite que puede ocurrir en una palabra monosílaba. Hay tres perfiles simples (A, B, y C) y cuatro perfiles compuestos básicos (CA, AC, CAC y CB). En la figura 27 podemos ver la forma (intuitiva, según es norma en el análisis de configuraciones) de los perfiles en la frase *It's John* (op. cit.: 141).

Perfil A	It's J _ó hn
Perfil B	It's J ó h ⁿ
Perfil C	It's J ó h ⁿ
Perfil CA	It's J _ó h n
Perfil AC	It's J _ó h ⁿ
Perfil CAC	It's J _ó h ⁿ o

Figura 27 - Los *profiles* de Bolinger.

Estos siete perfiles básicos pueden combinarse indistintamente para formar distintos contornos. Bolinger distingue, en general, contornos que acaban en A, contornos que acaban en B, contornos que acaban en C y contornos que acaban en AC (v. figura 28).

El contorno es, pues, "la forma de una entonación *completa*. Los contornos son a los perfiles más o menos como las oraciones a las palabras y tienen el mismo tipo de límites indeterminados que la oración" (op. cit.: 254).

Contornos que terminan en A	B + A A + A C + A	
Contornos que terminan en B	B + B A + B C + B	
Contornos que terminan en C	C + C A + C B + C	CB + C CA + C
Contornos que terminan en AC	B + AC A + AC C + AC	AC + AC CA + AC

Figura 28 - Tipos de contorno según Bolinger.

Un contorno puede tener un perfil o más, aunque lo más común es encontrar uno o dos perfiles por contorno. Generalmente, las marcas de incompletitud de contorno suelen ser: el ritmo (el límite rítmico coincide con el límite del contorno), la ausencia de pausa (que, generalmente, marca final de contorno), el gesto (de los

ojos, los labios, etc., que indican que el enunciado aún no ha acabado), la cohesión morfosintáctica, y los clichés del hablante (personales, como las muletillas, que sirven para unir perfiles, o frases hechas estereotipadas en la lengua, como la *tag-question*).

Los contornos, en fin, son sumas de perfiles, pero no puede hablarse de contornos "holísticos", ni de "frases entonativas", como si fueran nuevas unidades originales: sólo los perfiles son significativos. Así, no puede hablarse de "entonación de pregunta", sino de "perfil que marca la incompletitud del enunciado".

Los significados que se asocian a los perfiles provienen, según Bolinger, "de las metáforas asociadas a lo ascendente y lo descendente, y que comparten las expresiones faciales y los gestos corporales" (op. cit.: 341).

Así, el perfil A está relacionado con la terminación, con la marca de final, de que la frase ha terminado y puede tomar el turno otro interlocutor. Además,

y en contraste con el perfil B, A es usado, dentro del propio enunciado, para introducir un hecho independiente. La forma descendente, según Bolinger, es un indicativo modal de la finalidad, pero también de la seguridad: por eso es la usada en las frases asertivas o imperativas.

El perfil B, por su parte, es esencialmente ascendente, y por tanto es el opuesto natural de A: la subida de la entonación indica incompletitud, que lo dicho "queda en el aire"; también indica ansia, sorpresa, excitación, y en general una "alta emotividad", como la propia "curiosidad que conlleva la pregunta". El gesto análogo más común en la expresión facial es el de enarcar las cejas. El perfil B no sólo sirve para formar preguntas, sino también para marcar la subordinación y, en general, como soporte a "numerosas aplicaciones morfológicas y sintácticas".

El perfil C, finalmente, se caracteriza porque la sílaba acentuada (*accented syllable*) se mantiene abajo, sin ningún salto "desde" ella. Así, frente a A y B, en los que la sílaba acentuada es de un tono más alto, "C se opone a ambos en que retiene cualquier movimiento ascendente en el punto del *accent*". Las asociaciones metafóricas tienen que ver, esta vez, con la restricción, con la moderación, el control, la cortesía, etc.: "en una pregunta con la que se aborda a alguien, o dando una disculpa, o contradiciendo suavemente a alguien (...) para tranquilizar a un adulto o confortar a un niño (...) para minimizar la importancia de algo...". El gesto facial de acompañamiento más común es una leve inclinación de la cabeza, como una reverencia, "que se adecúa a la inclinación (o reverencia) entonativa de C" (todas las citas, 1986: 337 y ss.).

Sin embargo, "hacer un catálogo completo de contornos y sus significados podría ser imposible, primero porque no hay límite en el *número* de perfiles que un contorno puede contener, segundo por la variación entonativa dentro de los contornos, que altera gradualmente algunos aspectos de su significado, y tercero..." porque en cada contexto (verbal o situacional) el significado del

contorno puede cambiar (op. cit.: 274).

Para Bolinger, la entonación tiene que ver con la oposición "arriba y abajo" (*up and down*), muy relacionada con los gestos de los músculos faciales y de los hombros, y con significados metafóricos asociados a tales movimientos, que no están codificados en ninguna lengua en concreto sino que pertenecen al bagaje humano universal, es decir, que están motivados genéticamente.

Sin embargo, también apunta que "la simple oposición arriba / abajo está enormemente enriquecida por las modificaciones que pueden operarse en ella" (op. cit.: 223), es decir, que la entonación no es una mera sucesión de valores en dos niveles (alto/bajo), como podría suponerse (y que es la posición que adoptarán algunos autores norteamericanos, como Pierrehumbert, que proponen sólo dos niveles), sino que entran en juego otros elementos:

Glide vs. jump: un salto en la entonación no es igual a un deslizamiento del tono, sino que significan cosas distintas, y por tanto deben considerarse "movimientos entonativos" distintos. "En el mundo gradual de la entonación, *cada cosa* que pueda ser detectada es potencialmente significativa" (op. cit.: 225).

Monotone vs. rise or fall: frente al valor de la subida y la bajada de la entonación, el mantenimiento del tono es también un elemento significativo, siempre en términos binarios, es decir, la suspensión del tono frente al movimiento tonal, que a su vez puede ser ascendente o descendente.

Intensity of segment / Length of segment: la intensidad del segmento acentuado y su duración también pueden ser elementos significativos.

Relative height: un perfil entonativo será el mismo si mantiene su estructura, sea a la altura que sea. En este sentido, Bolinger distingue tres "registros" tonales: la voz de falsete (*falsetto*), que contempla los valores de tono más altos posibles de la voz; la voz normal (*modal*), que contempla los valores

medios; y voz profunda o "cascada" (*fry or creak*: "frita o crujiente", según la traducción literal), que contempla los valores tonales más bajos posibles.

Frente al cambio de "registro", que consiste en un cambio en el timbre de la voz para forzar cambios tonales, Bolinger también hace uso del término *key*, con el que se refiere a los cambios tonales (alto, medio o bajo) que se operan dentro de un registro. Sin embargo, no acierta a distinguir perfectamente los dos términos y opta por usar *key* sólo "para un uso informal refiriéndonos a un tono alto que corresponda a lo que haya sido afinado alto por alguna razón (informativa u otra), pero no en un sentido técnico" (op. cit.: 244): este es un buen ejemplo, por tanto, de cómo Bolinger ha descuidado los aspectos teóricos de su descripción, pues es capaz de introducir en su manual-compendio un concepto ajeno a su propia teoría, no acertar a definirlo bien y aceptarlo finalmente a medias.

Sequence: el orden de los factores sí altera el producto, por lo que la sucesión de valores tonales dentro de un perfil y de perfiles dentro de un contorno depende de su orden y no sólo de su presencia; no es lo mismo, sino todo lo contrario, un movimiento tonal de arriba abajo que de abajo arriba.

Combination: Cada modelo, cada perfil, puede combinarse consigo mismo y con cualquier otro modelo, y cada combinación puede observar variaciones de intensidad, de duración, de altura, etc., que serán significativas.

Por otra parte, para Bolinger la entonación no constituye un fenómeno estrictamente lingüístico y arbitrario, del mismo modo que sí lo es la gramática o la fonología. Por el contrario, considera que "la entonación forma parte de un complejo gestual cuya función primitiva, que aún pervive, es el señalamiento de la emoción" (op. cit.: 195).

Según esta perspectiva, la entonación reside en el hemisferio derecho del cerebro, donde se localiza todo el "procesamiento prosódico y afectivo". En

realidad, el propio emplazamiento del *accent* obedece a razones subjetivas, de realzar lo más importante emocionalmente, de modo que una palabra se remarca y se entiende que está remarcada sólo de un modo intuitivo.

Esto quiere decir que la entonación constituye un fenómeno no particular de cada lengua determinada, sino universal, y que los significados metafóricos que se entienden no pueden ser formalizados del mismo modo que formalizamos los significados conceptuales: todo es intuitivo y aproximativo, como en el mundo del gesto. Ciertamente, reconoce un cierto nivel de independencia de algunas entonaciones particulares, y admite que pueda haber un cierto grado de arbitrariedad en ellas, aunque los parecidos entre las lenguas, insiste, sobrepasan con mucho a las diferencias.

Ya hemos visto que el análisis de configuraciones es un método en el que es fácil despreciar el carácter lingüístico de la entonación, que es la razón de ser, por ejemplo, del análisis de niveles. En un momento, Bolinger deja claro por qué es así: "el hecho de que las configuraciones entonativas estén emparejadas con las configuraciones de la expresión facial y los gestos del cuerpo, de que ambas operen muchas veces paralelamente, y de que sus semejanzas correspondan a semejanzas de función, apunta al enfoque de configuraciones como el más cercano al éxito" (op. cit.: 337).

Según Ladd, el enfoque de Bolinger se caracteriza "por insistir en que todos los significados abstractos básicos que descubrimos en la entonación deben ser (a) directamente vinculados a los usos prelingüísticos del tono en la expresión de la emoción, y consecuentemente (b) universal", "directamente basado en un código emocional universal" (1990: 808).

Esta perspectiva, en realidad, está muy cercana a deducciones que pueden hacerse de otros autores, como Halliday, que sin embargo difícilmente podría aceptar una formulación así. En este sentido, Bolinger constituye una excepción en un ámbito en el que cada vez hay más convencimiento sobre el rendimiento

estrictamente lingüístico de la entonación.

Como anota Ladd (1990), las aportaciones ideológicas de Bolinger han sido insistentemente despreciadas en el mundo académico, a pesar de la influencia que ha ejercido en dos generaciones de entonólogos británicos: una influencia constante, aunque de baja tensión. Algunas de las afirmaciones que vierte en su obra, efectivamente, son del todo incompatibles con la mayoría de los trabajos que estudian la entonación, especialmente en las últimas dos décadas; no tiene sentido, por ejemplo, hablar de entonaciones lingüísticas más que para un oído aficionado (pues de eso se trata, de "oído"): "lo más obvio sobre entonación para el oído de un aficionado es que esta contribuye a interpretar un enunciado como una pregunta, una orden, etc.; y el aficionado tiende a sentir que hay entonaciones *propias de* esos variados tipos de actos de habla. Pero hemos visto que la entonación es demasiado espontánea para ser confinada de ese modo" (1986: 338).

Sin embargo, para el experimentado oído de Bolinger la entonación sí que desempeña una importante función comunicativa: "la entonación es importante para saber quién está hablando, para saber quién hablará el próximo turno, para ver cómo se entiende el acto de habla (...) para ver cómo será considerado el hablante (como nativo, miembro de una determinada clase social, etc.) (...) (y también) es importante para hacer distinciones gramaticales, marcar acotaciones, paréntesis, etc.; para identificar presuposiciones, temas y remas, coordinaciones, la parte y el todo, el alcance de una negación..." (ibid.). Diríamos, por tanto, que Bolinger considera la entonación como un importante operador pragmático y discursivo, cuya función, sin embargo, no es realmente previsible y formalizable, generalizable, sino guiada sólo por la intuición tanto del hablante como del oyente, genéticamente codificada en términos metafóricos, emocionales y actitudinales.

En resumen, la perspectiva de Bolinger consiste en considerar más

adecuado el análisis de configuraciones, dado su carácter impresionista y no formalizable, para describir un fenómeno como la entonación, que forma parte del complejo gestual encargado de transmitir la emoción en el discurso; la entonación, por tanto, no es un fenómeno estrictamente lingüístico, como sí lo es la gramática, aunque participa activamente en la comunicación humana; los significados entonativos son intuitivos y metafóricos, basados en la oposición arriba / abajo; el modo de remarcar las palabras importantes del enunciado es mediante el *accent*, que consiste en una inflexión tonal; hay una serie básica de inflexiones tonales, llamadas "perfiles", que tienen asociado un significado general y metafórico; los perfiles se juntan formando contornos, que no son unidades genuinas sino conjuntos de unidades; no puede elaborarse un lexicón entonativo, pues la entonación es un fenómeno espontáneo y aproximativo, contextual.

2.2.3. *El manual de A. Cruttenden*

En 1986 aparece el libro de cabecera del moderno análisis de configuraciones: *Intonation*, de Alan Cruttenden. En él, por primera vez en la tradición británica, se intenta un compendio de todo cuanto se sabe sobre entonación, en forma de manual para especialistas: un "estado de la cuestión" actualizado. Hasta entonces, las obras generales eran un mero compendio de las teorías de un solo autor (como Bolinger, 1986 y 1989), o bien un "estado de la cuestión" parcial, a modo de prólogo (como en Crystal, 1969, o Ladd, 1980). Lo más común en la tradición, sin embargo, eran los manuales utilitarios de dicción para aprendientes de inglés correcto.

Con todo, el trabajo de Cruttenden no es un compendio objetivo de cuanto se sabe sobre entonación, sino más bien una puesta al día del análisis de configuraciones tradicional, teniendo en cuenta las aportaciones más dispares:

desde las ideas más homologables de Bolinger, a las reglas métricas de Liberman & Prince; desde las descripciones tradicionales de O'Connor & Arnold (1961) para el inglés, a las revolucionarias descripciones de Collier & t'Hart (1981) para el holandés o Gårding (1983) para el sueco. Es más, su intención es generalizar el método hasta el punto de proponer el análisis de las más diversas lenguas o de los distintos dialectos de una lengua, y su comparación.

En este sentido, el manual de Cruttenden supone un punto culminante en la tradición británica del análisis de configuraciones, que con él se absolutiza y se presenta como "el método": capaz de afrontar cualquier descripción entonativa y de asumir en su seno cualquier aportación teórica o de corpus, venga del sitio que venga, aunque tenga su origen en unos presupuestos teóricos incompatibles con los suyos (como ocurre, por ejemplo, con las reglas métricas de acentuación, que sólo tienen sentido en un entorno generativista y que Cruttenden incorpora a su modelo directamente). En realidad, en numerosas universidades europeas se plantea actualmente el análisis de la entonación siguiendo el modelo de Cruttenden, tomándolo como el que engloba a todos los demás. Sin embargo, el modelo de Cruttenden apenas se aparta en lo esencial de los modelos tradicionales británicos y sólo se muestra eficaz en la descripción del inglés, como veremos. Debe considerarse, pues, como la culminación de una tradición parcial, pero no como su superación.

Al igual que Bolinger en su libro coetáneo, Cruttenden comienza distinguiendo los tres usos del tono, asignando, sin embargo, un tipo de lengua a cada uso: así, habla de lenguas "tonales", de lenguas "de entonación" y de lenguas "con acento melódico" (*pitch accent*). Entre estas últimas figura, por ejemplo, el japonés, que es una lengua "de entonación" pero cuyo acento no es de intensidad, "como en inglés", sino debido a una inflexión tonal, es un "accento melódico". Es decir, el acento, tanto en inglés como en la mayoría de las

lenguas de entonación, para Cruttenden es un fenómeno de intensidad, no de tono.

Sin embargo, puesto que es una evidencia que sí existe un "acento melódico" que no es de intensidad, entonces opta por considerar que hay un tipo de lenguas en las que el acento no es *stress* sino *accent*: y, según la tradición, su propia tradición, el inglés no puede ser una de ellas. De este modo se ignora toda la bibliografía científica que ha descrito el carácter tonal del acento en inglés y se opta por seguir los postulados tradicionales, sin más, aportando de paso una nueva tipología lingüística.

Este tipo de inexactitudes y operaciones teóricas más que dudosas son frecuentes en el manual de Cruttenden, en el que, por ejemplo, se cita al español como una lengua de acento fijo (1986: 17) o se postulan cuatro grados de acentuación en inglés, siguiendo sorprendentemente el modelo norteamericano en vez del británico: como, además, se encuentra con que las descripciones de otras lenguas europeas, como el francés, se han basado en el análisis de configuraciones, que propone sólo tres grados, deduce que sólo el inglés tiene cuatro (op. cit.: 22), sin caer en la cuenta de que las descripciones que siguen métodos distintos no son homologables, y que los autores norteamericanos han trabajado sólo con el inglés.

En general, puede observarse a lo largo de todo el manual cierta fascinación por las propuestas norteamericanas más atrevidas, provinientes sobre todo del campo generativista, o tal vez cierto coqueteo que podría explicarse por un ambiente académico especialmente favorable a las ideas chomskianas; sin embargo, Cruttenden no intenta, en ningún momento, llevar a cabo una integración efectiva de ambas tradiciones, y se limita a recoger en el seno de la tradición británica algunos elementos aislados de la norteamericana.

Así, en inglés habría cuatro niveles acentuales: el acento primario (que sería el núcleo de la entonación, y que consistiría en una inflexión tonal), el

acento secundario (que sería una especie de "acento de insistencia", también de origen tonal), el acento terciario (que sería el tradicional acento de intensidad: *stress* o *word-stress*) y las sílabas átonas (en las que ocurre la reducción vocálica, y que ocupan todo el grupo rítmico, hasta la siguiente vocal acentuada).

Según Cruttenden, el enunciado se segmenta en "grupos de entonación" (concepto empleado ya por Palmer), cuya definición, sin embargo, sólo viene dada mediante las marcas de límite que permiten determinarlos. Hay dos modos de marcar el límite de los grupos de entonación: los "criterios externos", que son marcas ajenas a la propia estructura entonativa del grupo, y los "criterios internos" que contiene la propia estructura entonativa.

Los criterios externos son: la pausa, la presencia de anacrusis, el alargamiento de la sílaba final del grupo y los cambios en la melodía que ocurren en las sílabas átonas.

Puede hablarse de dos tipos de pausa, según Cruttenden: la pausa "vacía" y la pausa "llena". La pausa "vacía" es la pausa genuina, que consiste en un silencio más o menos prolongado pero claramente identificable; la pausa "llena", en cambio, es ese tipo de sonidos que se emiten aisladamente como marca de titubeo ("m...", "eh..."). La pausa, generalmente, aparece en el discurso como un separador de los constituyentes sintácticos básicos: el sujeto del predicado, el objeto dislocado del sujeto, etc.; sin embargo, Cruttenden destaca que no es necesaria para separar tales constituyentes, con lo que no siempre puede servir como marca de límite de grupo de entonación. Además, en muchas ocasiones aparece una pequeña pausa ante las palabras de mayor contenido léxico, con lo que el uso de la pausa como marca de límite puede verse muy distorsionado. Otras veces, la pausa aparece tras la primera palabra del grupo de entonación, como un pequeño titubeo: tampoco en este caso sirve para separar el grupo, y

de nuevo puede ser un elemento distorsionador.

Las sílabas inacentuadas previas a la primera sílaba tónica del grupo es lo que se llama "anacrusis" (que, como vimos, para algunos autores constituye la primera parte del contorno: *pre-head* o *anacrusis*). En los grupos en los que hay anacrusis, tales sílabas átonas pueden marcar el inicio de contorno.

Igualmente, en ocasiones la sílaba final del grupo de entonación experimenta cierto alargamiento, que, en los casos en los que ocurre, puede marcar el final del grupo. Sin embargo, muchas veces también hay alargamientos en el interior del grupo, por razones de énfasis, con lo que tampoco puede considerarse como un criterio muy seguro.

Un tipo de juntura muy eficaz como marca de límite de grupo son los cambios en la melodía que ocurren en las sílabas átonas. En principio, las sílabas átonas no deberían experimentar ningún cambio tonal, pues son precisamente "átonas". Sin embargo, en numerosas ocasiones una sílaba átona realiza un salto tonal desde la sílaba anterior, de modo que la línea melódica queda "rota" o interrumpida: cuando esto ocurre, es que en esa precisa sílaba ha comenzado un nuevo grupo de entonación.

Los criterios internos para delimitar los grupos de entonación, por su parte, son, junto al último criterio externo, los fundamentales para delimitarlos efectivamente, aunque generalmente en combinación con los demás criterios externos. En primer lugar, cada grupo de entonación ha de tener necesariamente un acento primario: es decir, en un enunciado habrá tantos grupos de entonación como núcleos puedan detectarse. En segundo lugar, desde ese acento primario, núcleo de un grupo de entonación, debe haber, también necesariamente, una inflexión o cambio tonal: "desde" él o "hacia" él.

En resumen, las marcas de límite de grupo son: un cambio abrupto en la dirección de la melodía ocurrido en una sílaba átona, y la combinación de los demás marcadores (pausa, anacrusis, alargamiento del final, y existencia de un

acento primario en el que ocurra una inflexión tonal).

En un grupo de entonación puede haber varios acentos tonales (acentos secundarios), el más prominente de los cuales será el núcleo del contorno. Las inflexiones en que consisten tales acentos tonales (o melódicos) pueden ser, genéricamente, de cuatro tipos: un ascenso o un descenso tonal, o bien "desde" el acento o bien "hacia" el acento (v. figura 29)

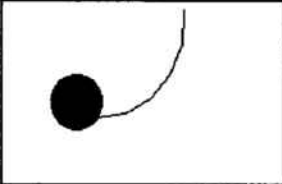

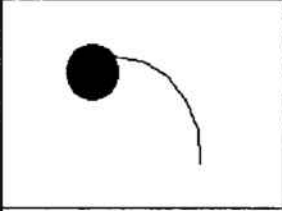
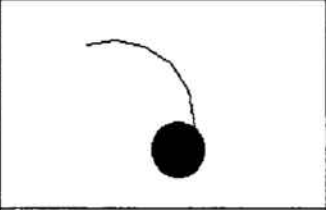
	
ascenso "desde"	ascenso "hacia"
	
descenso "desde"	descenso "hacia"

Figura 29 - Inflexiones tonales, según Cruttenden.

En inglés, el medio más importante de focalizar la información relevante de la frase es, justamente, localizando el núcleo entonativo en la parte que quiere focalizarse: para el entonólogo, por tanto, también es

esa la mejor manera de localizar el núcleo entonativo, buscando el foco informativo de la frase (la parte donde está la información "nueva": el *rema*).

Normalmente, el núcleo del grupo entonativo es el último acento melódico del grupo, es decir, el último elemento léxico de la frase, a no ser que quiera focalizarse cualquier otra parte: como Bolinger cuando habla de su *accent*, Cruttenden también adjudica al núcleo de la entonación un importante valor enfático, restringido, sin embargo, por la idea de que no puede enfatizarse simplemente cualquier parte de la frase, sino que el emplazamiento del núcleo debe regirse, en todo caso, por ciertas reglas sintácticas (siguiendo en este extremo, por tanto, a Halliday).

Con todo, para Cruttenden no siempre coinciden la cláusula sintáctica y

el grupo de entonación, cuya relación es más bien "probabilística": "lo más que puede afirmarse es que la cohesión sintáctica es más estrecha dentro los grupos de entonación que a través de los grupos" (1986: 79).

Pueden distinguirse dos tipos de foco de información: un "foco ancho", que abarca a toda la frase, en los casos en los que toda la frase es informativa (toda la frase es rema); y un "foco estrecho", que abarca sólo a una parte de la frase, pues el resto contiene información ya conocida, *tema*.

Generalmente, en ambos casos el núcleo recae sobre el último elemento léxico del foco:

*John ran all the way to the **station**.*

Sin embargo, Cruttenden señala que en ocasiones el núcleo puede recaer en otro elemento léxico que no sea el último, o bien fuera del foco (en los casos de "foco estrecho"). Por ejemplo, en las "oraciones de suceso" (*event sentences*), el núcleo suele recaer sobre el sujeto, que es la parte más informativa, a pesar de tratarse de un "foco ancho":

*Watch out! That **chimney's** falling down.*

También en los casos en los que el elemento final es un adverbio el núcleo recae sobre otro elemento:

*He didn't know how to **do** it fortunately.*

En los casos en los que haya una partícula interrogativa (*wh- element*) adjetiva, el sustantivo al que acompaña, si cumple la función de objeto, será la

palabra sobre la que recaiga el núcleo:

Which course did you take?

En casos de "foco estrecho" también puede focalizarse, como hemos dicho, un elemento que esté fuera del foco: bien porque el elemento "nuclearizado" opone cierto "contraste" al significado del foco informativo, a pesar de que era información conocida; bien en preguntas o exclamaciones "eco", que lo único que hacen es repetir una frase anterior (con sorpresa, por ejemplo), que es por tanto información conocida; o bien para dar un especial énfasis, de "insistencia", a un elemento léxico que ya era conocido pero que conviene resaltar (1986: 74-88).

Es decir, si bien la norma general es que el núcleo de la entonación recaer sobre el último elemento léxico de la frase, Cruttenden se esmera en detallar todos y cada uno de los casos en los que tal norma es contravenida (y es que este era, precisamente, el tema de su tesis doctoral, 1981).

Nuestra intención al detallar aquí este aspecto concreto del modelo ha sido sólo la de comprobar dos cuestiones que pueden resultar importantes: una, cómo continuamente se hace referencia a categorías gramaticales y a fenómenos sintácticos para definir uno de los aspectos más importantes del análisis entonativo del modelo como es la localización del núcleo (a pesar de que previamente había negado la dependencia sintáctica de la entonación); otra, cómo todos los datos que se ofrecen como argumentos para sostener una "teoría general de la entonación" sólo tiene sentido referidos al inglés, pero no a cualquier otra lengua: en español, por ejemplo, rara vez se focaliza una información por medio del núcleo entonativo, además de que los casos de núcleo entonativo dislocado que acabamos de ver son rigurosamente extraños a nuestra lengua.

Es decir, el modelo de Cruttenden está demasiado ligado aún al inglés como para servir a otras lenguas, ya desde su mismo principio y en la propia definición de sus elementos fundamentales. Esto no es, desde luego, ningún fenómeno extraño, pues es una característica común a todos los modelos de análisis pertenecientes a la tradición británica; simplemente, es un serio impedimento para tomarlo como modelo universal.

Otros factores importantes en la descripción de los grupos de entonación son el "campo tonal", la "tesitura" y el "registro".

El "campo tonal" (*accent range*) se refiere al recorrido tonal que pueden realizar las inflexiones, esto es, a la diferencia de valor de la F_0 entre el comienzo y el final de la inflexión: el ascenso "desde" o "hacia" y el descenso "desde" o "hacia", que pueden ser más o menos alargados.

La "tesitura" (acertada traducción de Mascaró del término original *key*) se refiere a la anchura del recorrido tonal que puede o que suele realizar una persona a lo largo de su alocución, y que puede ser ancha o estrecha, siempre en términos relativos. Por ejemplo, se supone que una entonación enfática de alegría suele usar una tesitura más ancha que una entonación de tristeza, cuya tesitura será más estrecha.

El "registro" (*register*), que se refiere a la altura media de la entonación de una persona, no es exactamente igual al término *register* usado por Bolinger, que se refería a un cambio de timbre para forzar un cambio tonal: voz de falsete, voz profunda, etc. En este caso, tiene que ver con la altura relativa de la línea base de la entonación del hablante, que puede ser alta, media o baja: por ejemplo, se dice que un registro alto favorece entonaciones de cortesía o de inseguridad, mientras que un registro medio suena más objetivo, y un registro bajo favorece entonaciones de autoridad.

El punto más importante e informativo del grupo de entonación es, según el modelo de Cruttenden, el núcleo, de modo que, en realidad, no puede hablarse de "contornos" entonativos completos, como si se tratara de unidades entonativas discretas: esta es una característica diferencial muy importante del análisis de configuraciones británico con respecto a otras tradiciones, como la española, que sí consideran informativo todo el contorno. Para Cruttenden, lo anterior al núcleo es enteramente irrelevante, y ni siquiera vale la pena hablar de "coda" (*tail* o *terminal*) como de un elemento separable, pues en la inflexión tonal del núcleo ya están contenidos todos los movimientos tonales informativos del grupo.

Del mismo modo que Palmer ofrecía una lista de *tone-patterns* con los que se caracterizaban el grupo de entonación, Cruttenden, siguiendo de cerca a O'Connor & Arnold (1961) y a otros autores, como Crystal (1969), ofrece una lista de "tonos nucleares" (*nuclear tones*) del inglés (de nuevo, en la descripción general del modelo de análisis sólo puede ofrecer la descripción de su lengua), que constituyen los patrones generales significativos que pueden seguir las inflexiones tonales en que consisten los núcleos de entonación.

Así, el movimiento tonal desde el principio del núcleo hasta el final del grupo constituye un tono nuclear, que puede definirse por tres factores independientes (similares, según Cruttenden, en todas las lenguas de entonación): a) el movimiento inicial del núcleo, que puede ser un ascenso, un descenso o una suspensión del tono; b) el punto de arranque del movimiento inicial, que puede ser alto o bajo; y c) un segundo movimiento tras el núcleo, que puede dar lugar a tonos nucleares complejos, ya que puede ser un descenso-ascenso, un ascenso-descenso o un ascenso-descenso-ascenso.

En total, Cruttenden ofrece un inventario de siete tonos nucleares en inglés: descenso alto, descenso bajo, ascenso alto, ascenso bajo, descenso-ascenso, ascenso-descenso y suspensión (v. figura 30).

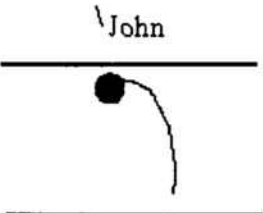
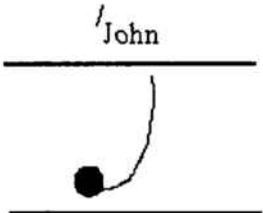
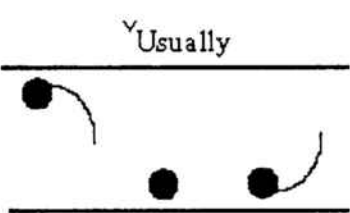
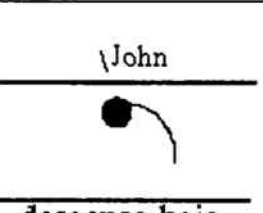
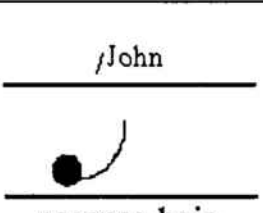
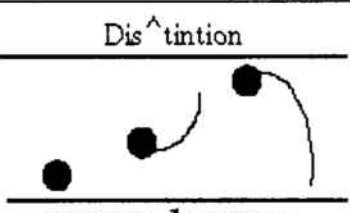
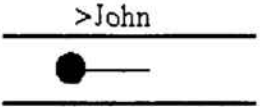
 <p>\John</p> <p>descenso alto</p>	 <p>/John</p> <p>ascenso alto</p>	 <p>^Usually</p> <p>descenso ascenso</p>
 <p>\John</p> <p>descenso bajo</p>	 <p>/John</p> <p>ascenso bajo</p>	 <p>Dis^tinction</p> <p>ascenso descenso</p>
	 <p>>John</p> <p>suspensión</p>	

Figura 30 - Tonos nucleares del inglés.

El tipo de transcripción usado por Cruttenden es similar al de Armstrong & Ward, pero con las inflexiones al modo de Palmer. Esta es la transcripción más extendida entre los autores británicos desde los años 60. Sin embargo, Cruttenden añade algo más: junto a la transcripción intuitiva a base de puntos y rayas, en la línea de escritura (ortográfica y no fonética) se añaden unos símbolos que también reproducen los tonos nucleares; las marcas son (1986: XII):

- | | |
|---------------------------------|--------------------|
| / límite de grupo de entonación | , ascenso bajo |
| \ descenso alto | ∨ descenso-ascenso |
| ∖ descenso bajo | ^ ascenso-descenso |
| / ascenso alto | > medio-sostenido |

De este modo, consigue dos cosas: en primer lugar, puede prescindir de la línea de puntos y rayas en determinadas ocasiones, que además resulta demasiado "escolar"; en segundo lugar, así su notación se parece más a la notación del análisis de niveles, tenida por más "científica", y en la misma línea de adoptar elementos provenientes de la otra tradición que caracteriza el trabajo (algunos elementos están simplemente calcados: " / " , por ejemplo, es la marca de límite de grupo, exactamente igual que en el análisis de niveles).

El significado de la entonación siempre ha sido un aspecto tomado muy en cuenta en la tradición británica, y ya desde Palmer encontramos algo parecido a "lexicones entonativos": los significados asociados a la inflexión del tono en el núcleo del contorno, y en un contexto determinado.

En este punto, Cruttenden sigue muy de cerca a Ladd (1980), y define el significado entonativo como de carácter "discursivo" (término que Cruttenden emplea como sinónimo de "pragmático") y "actitudinal". En ningún caso, dice, hay que asignar un carácter "gramatical" a los significados de la entonación: las oraciones "interrogativas", "declarativas" o "imperativas" lo son por razones gramaticales, pero no entonativas.

Así, los significados de los tonos nucleares (y no de los contornos, que no constituyen realmente una unidad entonativa) son del tipo de "educado", "incrédulo", "quejumbroso", "conspirativo", "desafiante", "tranquilizador", etc.

Cruttenden acepta, sin embargo, y siguiendo la línea de eclecticismo con la que enriquece su modelo, que pueda haber un tipo de significados más "abstractos" asociados a esquemas entonativos generales, al modo de los significados descritos por Bolinger: terminal, incompleto, etc.; acepta, pues, tanto la propuesta de Brazil (1975), según la cual hay dos entonaciones significativas (descenso y descenso-ascenso), como la de Gussenhoven (1983), que propone tres entonaciones (descenso, descenso-ascenso y ascenso).

Sobre tales significados "abstractos" se solapan los significados "locales" de los siete tonos nucleares descritos, que se agrupan genéricamente en "descensos", "ascensos dependientes" (es decir, precedidos de otros grupos de entonación) y "ascensos independientes"; finalmente, el significado concreto de una entonación concreta estará condicionado por los "factores contextuales condicionantes": de carácter sintáctico, como el tipo de frase que contenga la entonación, o de carácter estrictamente pragmático, como el juego de presuposiciones en la conversación, etc.

No parece, pues, que para Cruttenden haya nada parecido a un "signo lingüístico entonativo", posibilidad claramente expresada por Ladd, pero que debería considerarse, en el ámbito de la tradición británica, más como una extravagancia que como un programa de investigación.

En el terreno de la "comparación entre lenguas", sin embargo, de nuevo Cruttenden se deja llevar por corrientes ajenas, y usa abiertamente categorías entonativas que había desdeñado unas páginas atrás (declarativas, interrogativas sí/no, interrogativas parciales, imperativas...); del mismo modo, habla de contornos entonativos y no de tonos nucleares para lenguas que no sean el inglés, etc.

En resumen, el trabajo de Cruttenden consiste en una ambiciosa puesta al día del análisis de configuraciones británico, tratándolo como modelo "general" de análisis de la entonación: para ello emplea el curioso método de aceptar sin réplica o incluso con entusiasmo aspectos del análisis de niveles novedosos (esto es, no tradicionales) que, sin embargo, chocan frontalmente con sus propios presupuestos, como cuando acude a las reglas generativas de acentuación.

Por otra parte, en ocasiones desvirtúa la explicación de un fenómeno o de un método simplemente porque no entran en sus coordenadas ideológicas. Así, por ejemplo, cuando habla del análisis de niveles lo describe como si sólo fuera

un modo distinto de formalizar o de anotar la entonación (y por eso él mismo adopta una notación parecida, con las mismas juntas y con tonos nucleares en vez de niveles); da igual anotar una sucesión de números que una sucesión de alturas relativas, viene a decir, porque lo importante es anotar dónde está el tono nuclear y qué forma tiene la inflexión: es decir, no se da cuenta de que el análisis de niveles no puede contemplar tonos nucleares, porque su enfoque teórico es distinto. El análisis de niveles que presenta (op.cit.: 45 y ss.), pues, es falso, no silábico sino por contornos, y no fonológico: igual al análisis por contornos, pero con números en vez de con puntos y rayas.

Esto, unido a la enorme dependencia de la lengua inglesa que tienen los conceptos y los métodos empleados en su manual, en ocasiones difícilmente generalizables, hace del modelo de Cruttenden, como decíamos, no un modelo general, ni un modelo genuinamente ecléctico, como en ocasiones parece pretender, sino una auténtica culminación y una actualización vigorosa del tradicional análisis de configuraciones británico.

2.3. La tradición norteamericana

2.3.1. El análisis de niveles tradicional

Como en tantos otros aspectos de la lingüística norteamericana, el inspirador de la tradición norteamericana en el análisis de la entonación fue L. Bloomfield, cuyas ideas sobre el acento y la entonación expuestas en su libro *Language* (1933) constituyen la base sobre la que autores posteriores construirían el método de análisis por niveles.

Para Bloomfield, el acento (*stress*) es un fenómeno de intensidad, de fuerza, totalmente independiente del tono y de la entonación. Frente a los autores británicos de la época, que distinguían entre acento y prominencia (la prominencia era considerada un fenómeno tonal: el núcleo del contorno), Bloomfield habla sólo de acento, y considera que toda prominencia no será más que un grado mayor de acentuación. Así, uno de los caballos de batalla de la tradición norteamericana se establece ya desde el propio Bloomfield por contraste a la tradición británica: en ambos casos se distingue el acento de la entonación, como vimos, y en ambos casos el acento es considerado un fenómeno de intensidad, pero en el caso norteamericano la distinción es absoluta, y se incluye en ella también el fenómeno de prominencia, que se considera una manifestación más del acento, del *stress*.

La entonación, por su parte, para Bloomfield es un fenómeno lingüístico, de carácter suprasegmental pero completamente homologable a los fenómenos segmentales, de modo que puede hablarse, igualmente, de "fonemas" y de "morfemas tonales" (*pitch phonemes* y *pitch morphemes*). Los "morfemas tonales", consecuentemente, tendrán un significado propio de la entonación y

distinto del significado de las palabras.

Por "morfemas tonales", en principio, deberíamos entender algo parecido a los "contornos entonativos": una sucesión de "fonemas tonales" que constituyen una unidad significativa.

Por tanto, vemos cómo desde su mismo comienzo la tradición norteamericana se caracteriza por: a) un enfoque no pedagógico sino estrictamente científico en el análisis de la entonación; y b) un enfoque claramente lingüístico, según el cual la entonación está constituida por fonemas y morfemas, paralelos a los fonemas y los morfemas segmentales.

Uno de los primeros autores que se aplican a la tarea de describir los "fonemas tonales" del inglés es R. S. Wells, en su artículo "*The pitch phonemes of English*" (1945). En él, además de la primera crítica importante a un autor británico (a Palmer, de quien dice que no distingue entre los fenómenos del acento y la entonación), aparece lo que después sería la tradicional división en cuatro niveles tonales (Harris, 1944, por ejemplo, había propuesto seis niveles), que serían los cuatro fonemas tonales con los que describir la entonación.

Con ellos, Wells describe hasta 19 frases cuyo significado es diferente: y que constituirían, por tanto, "morfemas tonales" llenos de significado. La intención de Wells en este primer trabajo era trascender inmediatamente el nivel "fonológico" y establecer algo parecido a un "lexicón" de entonaciones, al estilo de los británicos. Sin embargo, el ejemplo no cundió en absoluto, y la mayoría de los autores se dedicó a estudiar y describir exclusivamente el nivel "fonológico" de la entonación, dejando de lado el nivel "morfológico", esto es, su parte significativa. Seguramente, la razón fue que los significados propuestos por Wells eran significados de carácter básicamente emocional, muy poco

sistematizables y, desde luego, muy alejados de la sensibilidad de los lingüistas norteamericanos de la época, seguidores de un estructuralismo formal poco dado a coqueteos con la semántica.

Más importante que Wells, aunque tal vez menos influyente en su propia tradición, es K. L. Pike, con su libro *The intonation of American English* (1945), único ejemplo de manual dedicado a la enseñanza del inglés (aunque no del inglés R.P. sino del inglés americano) basado en el análisis de niveles. Esto, junto a su afán por dilucidar el tipo de significado que aporta la entonación o el hecho de que hable de "contornos" más que de "morfemas", son características que lo acercan mucho a la sensibilidad británica, y ha sido efectivamente un autor muy citado por los estudiosos europeos; sin embargo, su aportación fundamental a la tradición norteamericana ha sido, sobre todo, el concepto de "juntura" (*junction*) como fonema tonal, tan importante, fonológicamente, como los niveles tonales para describir la entonación.

Pike también considera el acento como un fenómeno de intensidad, independiente de la entonación, y comienza a hablar ya de *sentence stress* como del acento más importante de la frase, situado en el comienzo del *primary contour* (lo que para los británicos es *nucleus* y *tail*; como vimos en la fig. 25, Pike divide el contorno en *precontour* y *primary contour*), pero independiente de los movimientos tonales.

Pike también habla de cuatro niveles tonales, aunque les asigna un "nombre" distinto: si en Wells el nivel 1 era el más bajo y el 4 el más alto, para Pike es justamente al revés, 1 es el nivel más alto y 4 el más bajo. Trager & Smith, los auténticos fijadores de la tradición, optaron, sin embargo, por el sistema de Wells, y con ellos la mayoría de los autores posteriores.

Con todo, Pike aporta algunas ideas que luego serán importantes, como el tema de la altura relativa o la altura absoluta de los niveles tonales: para el autor, no tiene sentido hablar de frecuencia fundamental dando los valores absolutos en hertzios, porque el valor del contorno reside en la sucesión de tonos y en su altura relativa con respecto a los niveles anteriores. En cada individuo, dice Pike, varían los valores absolutos, pero el esquema tonal es el mismo.

Junto a los cuatro niveles tonales, Pike considera otros dos elementos fonológicos: la "pausa final" (*final pause*) y la "pausa provisional" (*tentative pause*). Las pausas marcan el final de contorno: la pausa "final" supone el descenso tonal característico de las entonaciones declarativas; la pausa "provisional", por su parte, es algo más corta que la pausa final, y puede ser de dos tipos, la que supone un final de contorno sostenido y la que supone un final ascendente, y, en general, corresponde a entonaciones que incluyen una actitud de incerteza o "no-finalidad" (idea que luego tomaría con fuerza Bolinger, como hemos visto).

Con todo, y aunque sólo describa dos finales fonológicos de contorno, para Pike no puede hablarse sólo de dos entonaciones, porque cada contorno significativo está constituido por una serie de niveles tonales: es, por tanto, la estructura de niveles que constituye el contorno lo que lo caracteriza, también junto al tipo de final empleado. Pero no es exclusivamente el final, como en el análisis de configuraciones. Nótese que, en contraste con el análisis de configuraciones británico, en el que la *única parte informativa* de la entonación es el núcleo, en el análisis de niveles el núcleo del contorno no merece ninguna consideración especial.

Por su parte, el análisis de configuraciones no británico, como el

empleado por Navarro Tomás para el español, considera importante la melodía *de todo el contorno*, así como el final ascendente o descendente de la entonación y, por tanto, en este sentido se parece más al análisis de niveles que al análisis británico: razón por la cual los autores europeos han sido especialmente dados a combinar los dos sistemas.

Por otra parte, para Pike tampoco puede hablarse de marcas de límite de contorno demasiado seguras: la pausa, por ejemplo, aparece al final del contorno, pero también puede aparecer en el interior del contorno; y la unidad rítmica, que en un principio parece asociada a los contornos entonativos (una unidad rítmica equivaldría a un contorno de entonación), pronto descubre que no es una unidad comparable, pues un contorno puede contener más de una unidad rítmica, del mismo modo que una unidad rítmica puede contener más de un contorno.

En cuanto al significado de la entonación, Pike considera que es un significado superpuesto al significado de las palabras, un especie de "matiz de significado" (*shade of meaning*) añadido o superpuesto al significado léxico, y que básicamente refleja la actitud del hablante. Pike aboga, pues, por abandonar todo tipo de descripciones que partan de una base léxica o gramatical para definir los contornos entonativos: toda definición significativa debe contemplar la "actitud del hablante", porque este es el único contenido que aporta la entonación por sí misma.

El análisis de niveles en los términos en los que suele conocerse, sin embargo, fue fijado por Trager & Smith en su libro *Outline of English Structure* (1951). En él se definen todos los elementos de la entonación considerados tradicionalmente fonológicos, la terminología empleada, el tipo de transcripción

y, sobre todo, el enfoque radicalmente "fonológico" del análisis de la entonación, alejado de todo intento "morfológico" o "semántico". Con Trager & Smith, por tanto, la tradición norteamericana abandona, en cierto modo, el tema de los significados de la entonación, y opta por un análisis explícitamente formal.

Trager & Smith siguen a Wells en la caracterización de los niveles tonales como los "fonemas tonales" más importantes: son cuatro, del 1 (el nivel más bajo) al 4 (el más alto). Siguiendo a Pike, en cambio, renuncian a cuantificar el valor absoluto de los niveles tonales, a los que ni siquiera consideran "márgenes" de dispersión tonal: para ellos se trata, simplemente, de cuatro niveles relativos y fácilmente reconocibles por los hablantes.

Por otra parte, consideran otros cuatro "fonemas tonales" más: tres "junturas terminales", que marcan el final de una frase entonativa, y una "juntura interna", que separa las palabras en el interior de la frase. Las junturas terminales son:

[#], que viene a ser como la "pausa final" de Pike, y que se corresponde a un descenso tonal característico en entonaciones declarativas;

[//], que corresponde a un ascenso tonal, característico en las frases interrogativas; y

[/], que corresponde a un sostenimiento final del tono, a una suspensión tonal, característico de frases dubitativas o de incerteza. Las junturas [//] y [/] se corresponderían a la *tentative pause* de Pike.

La juntura interna, por su parte, cumple una función demarcativa dentro de la frase y constituye un punto de unión entre la fonología segmental y la suprasegmental. Se transcribe con el símbolo [+].

Aunque el análisis se llame "de niveles", las junturas son elementos de suma importancia que tienen dos finalidades bien distintas: marcar el final de

la frase entonativa, que generalmente se considera una pausa; y señalar el tipo de inflexión que ocurre en el final de la entonación (y que en el análisis de configuraciones era el único punto informativo).

Trager & Smith describen también cuatro niveles acentuales, que no deben considerarse, en ningún caso, "fonemas tonales": se trata, simplemente, de los cuatro grados en que puede manifestarse el fenómeno del acento, considerado, como en el resto de los autores, un fenómeno de intensidad e independiente.

El "acento primario" [´] es el *sentence stress*, en el suele recaer una inflexión tonal ajena, sin embargo, a su propia naturaleza acentual. El "acento secundario" [^] suele recaer al comienzo de la frase. Hay, también, un "acento terciario" [`] y "un acento débil" [~] que, en realidad, corresponde a las sílabas no acentuadas. Este mismo esquema, como vimos, es el que incorpora Cruttenden a su propia tradición.

La transcripción empleada por Trager & Smith consiste en una serie de diacríticos y símbolos especiales colocados en la misma línea de la escritura (ortográfica y no fonética): al comenzar la frase, se marca (en superíndice), el número del nivel tonal con el que comienza (sólo cuando haya un cambio de nivel se marca un nuevo número); al final de la frase o de la "cláusula" se coloca la juntura terminal correspondiente; la juntura interna se marca, cuando aparezca, al final de la palabra; finalmente, sobre las vocales acentuadas se colocan los diacríticos correspondientes a su grado de acentuación. Es una transcripción, por tanto, en la que no pueden marcarse las inflexiones tonales, consideradas secundarias: la única información sobre inflexiones la da la juntura, como hemos visto.

Un aspecto que llama poderosamente la atención del sistema de Trager & Smith es cómo, a pesar de su carácter explícitamente fonológico y formal, la

unidad entonativa considerada no es el contorno, como hiciera Pike dentro de la tradición norteamericana, sino la "cláusula" (*phonemic clause*), unidad que depende de la solidaridad sintáctica de las palabras de la frase: la *phonemic clause*, por tanto, es el correlato fonológico del sintagma gramatical. Ello es consecuencia, sin duda, del intento de evitar las cuestiones semánticas asociadas tradicionalmente a la unidad "contorno" que, considerada independientemente de la sintaxis (como hace, por ejemplo, el mismo Pike), necesariamente ha de contar con un significado que le dé carácter de "signo". Trager & Smith, por el contrario, optan por supeditar el nivel entonativo al nivel sintáctico para poder describirlo en términos estrictamente fonológicos y desentenderse, de momento, del significado entonativo.

En resumen, la tradición norteamericana de análisis de la entonación por niveles se inicia, casi un tercio de siglo después de la británica, con las ideas de Bloomfield sobre la fonología suprasegmental y su carácter segmentable. El análisis por niveles propiamente dicho, sin embargo, lo inauguran los trabajos de Wells y de Pike, los dos autores que más influyeron en Trager & Smith, auténticos fijadores de la tradición.

Puede decirse que el análisis de niveles ha sido el método más exitoso y el que más seguidores ha tenido, a pesar de haber sido el que menos aportaciones teóricas ha ido recibiendo a lo largo de los años y a pesar también de las feroces críticas a que lo sometieron Bolinger y otros autores europeos. Los repetidos intentos de conjuntar ambas tradiciones, a los que ya nos hemos referido, en realidad surgen de la fascinación que sobre numerosos autores europeos (seguidores, en principio, de un análisis de configuraciones) ha ejercido el aspecto formal y "objetivo" del análisis de niveles, fascinación que ha llegado

incluso a un autor como Cruttenden que, como vimos, es el responsable de la máxima expansión de objetivos del análisis de configuraciones británico.

2.3.2. *La irrupción generativa*

Cuando Chomsky & Halle inauguran "oficialmente" la fonología generativa con su libro *The sound pattern of English* (1968), apenas le dedican atención al análisis de la entonación, en el que siguen de cerca el método de Trager & Smith. Aunque el análisis de niveles, hasta ese momento, podría considerarse un método inequívocamente estructuralista, a partir de entonces se consagra también como el método generativista para analizar la entonación.

Chomsky & Halle, sin embargo, realizan una aportación importante en la caracterización y el tratamiento del acento, fundamental para el posterior desarrollo del análisis generativo de la entonación, plenamente integrado en la tradición del análisis de niveles.

Consideran estos autores que no hay que hablar de sólo cuatro niveles acentuales, como hacen Trager & Smith, porque no hay ningún límite teórico en el número de grados acentuales posibles, pues un acento siempre está en contraste con otro, este con otro, etc. (siguiendo, también aquí, la impronta binarista que embarga todo el libro). Así, /1/ sería el acento más importante de la frase y, paulatinamente, /2/, /3/, etc., los demás acentos gradualmente más débiles. Chomsky & Halle manifiestan que no les interesa la realidad fonética en sí, ni el hecho de que sea imposible (además de perceptivamente innecesaria) una sucesión "infinita" de grados acentuales: su propuesta es puramente teórica y los resultados de sus análisis son únicamente "constructos fonológicos abstractos".

Es más, a Chomsky & Halle ni siquiera les interesa el aspecto propiamente fonológico del acento, sino las "reglas de dependencia sintáctica" que lo generan: en el *lexicón* está la información necesaria para emplear cada unidad léxica, pero entre esta información no figura el acento, que sólo será asignado dentro de una oración en función de la categoría gramatical de la palabra (nombre, verbo, etc.), en función de la posición de la palabra dentro de la oración y según unas "reglas de generación del acento" que determinan el esquema de la frase.

Es decir, si en la escuela generativista es patente el lugar central que ocupa la sintaxis, vemos cómo incluso un fenómeno fonológico como el acento se ve completamente condicionado por la forma sintáctica de la frase.

Se trata, sin duda, de una perspectiva sorprendente y ajena a cuanto se había hecho anteriormente en fonología: para Chomsky & Halle, como para todos los autores generativistas posteriores, la forma sintáctica (la estructura profunda o, más modernamente, la estructura P) de la frase es previa, condiciona y genera la forma fonológica (llamada "forma fonética" en los modelos de *Rección y Ligamento*), que también es previa y genera al sonido concreto. Las sucesivas perspectivas generativas en fonología (fonología natural, fonología auto-segmental, fonología métrica, fonología léxica, etc.) no son más que variaciones que hacen hincapié en un aspecto u otro del proceso de generación de la forma fonológica (o "forma fonética").

Así, en realidad se trata de un paradigma teórico nuevo que no mantiene la diferencia fonética-fonología en los mismos términos y que opera justamente a la inversa que la fonología "clásica": si esta parte del sonido concreto para encasillarlo en una clase funcional abstracta (el fonema), cuya sucesión forma las unidades léxicas (morfema, palabra), que se relacionan entre sí (sintaxis) y

constituyen una unidad de sentido completa (oración), para la fonología generativa el principio del proceso es la propia oración, cuya estructura sintáctica abstracta se concreta mediante una serie de reglas de dependencia (sintácticas y fonológicas) que condicionan, por una parte, la forma lógica y, por otra, la forma fonológica concreta del discurso.

La fonología generativa, por tanto, no necesita de la unidad "fonema" sino de las "reglas de dependencia sintácticas" o "reglas generativas"; por tanto, tampoco la división segmental-suprasegmental tiene tanto valor como en la fonología clásica y pueden afrontarse los fenómenos fonotácticos con restricciones exclusivamente gramaticales, pero no fonológicas; paralelamente, la independencia relativa de la fonología se pone en entredicho y surgen nuevas líneas de investigación híbridas sintáctico-morfo-fonológicas.

Con todo, el análisis de la entonación siempre ha sido una excepción dentro de la fonética y la fonología clásicas, pues los sistemas propuestos nunca han coincidido perfectamente con los sistemas de análisis fonológicos estructuralistas: resultaban demasiado imprecisos, y los fonólogos preferían considerar la entonación como un fenómeno no lingüístico y no fonológico.

El análisis de niveles, por tanto, y a pesar de su génesis estructuralista, se adapta inmediatamente a las ideas generativistas, que encuadran el acento y la entonación en un ámbito de inequívoco rendimiento lingüístico. En realidad, y en lo que se refiere a los procedimientos, el generativismo en fonología siempre se ha mostrado, más que como una ruptura, como una superación del estructuralismo.

Con las reglas de generación del acento, por tanto, Chomsky & Halle inauguran una vía que irá a dar, una década después, a la fonología métrica; pero también imprimen un cambio de rumbo al análisis de niveles, reubicándolo

en el ámbito generativista sin apenas modificaciones.

Trabajos tan citados como el de Stockwell (1972) constituyen una imagen perfecta de esta reubicación: el autor se plantea la cuestión de la "representación" de la entonación (contornos o niveles, en lo que viene a ser una crítica del análisis de configuraciones), la cuestión del "acento nuclear" (*nuclear-stress*: nótese el uso de *stress*, en contraste con el término *nuclear-tone* empleado por Cruttenden: Stockwell, obviamente, pertenece a la tradición norteamericana y no concibe más acento que el de intensidad, a pesar de que en su trabajo intenta unificar, en cierto modo, los criterios de ambas tradiciones), la cuestión del "límite" de la frase entonativa (es decir, la cuestión de la "juntura") y la cuestión del "significado" de los contornos (que para Stockwell es una continuación del significado "gramatical"), todo ello siempre desde una perspectiva de dependencia sintáctica, tratando de desvelar las reglas de proyección fonológico-entonativas de la estructura gramatical de la oración.

P. Lieberman, un año antes del trabajo de Chomsky & Halle, había publicado *Intonation, perception and language* (1967), en el que se llevaba a cabo el primer tratamiento de la entonación desde un punto de vista netamente generativista, aunque muy distinto del que harían después los autores citados. El libro de Lieberman fue, y continúa siendo, una obra importante de referencia, muy influyente en su momento, aunque no precisamente entre los demás autores generativistas.

En este sentido, su propuesta teórica debe considerarse un relativo fracaso, aunque es un buen ejemplo de cómo podrían haber sido las cosas en la tradición norteamericana de análisis de la entonación si la actitud de Chomsky & Halle hubiera sido más radical en el terreno de la fonología y menos respetuosa con

la tradición anterior a ellos (y, en concreto, con el análisis de niveles). Como veremos, Lieberman va mucho más allá que Chomsky en la aplicación fonológica de algunos preceptos chomskyanos, aunque en ningún momento vislumbra la posibilidad de inaugurar un nuevo paradigma fonológico basándose en ellos.

El trabajo de Lieberman parte de un enfoque fisiológico, articulatorio, de la producción de los rasgos tonales. La unidad entonativa, explica, es el "grupo espiratorio" (*breath group*), que es una realidad física universal: la espiración del aire dura unos dos segundos, durante los cuales hay salida de aire y presión infraglótica sobre las cuerdas vocales. Por tanto, Lieberman no habla de "contorno" ni de "frase entonativa", etc., sino únicamente de "grupo espiratorio".

El centro (o núcleo) del grupo espiratorio es una "prominencia" (*prominence*) producida por un aumento momentáneo de la presión infraglótica sobre una vocal. El acento, entonces, es un fenómeno a la vez de intensidad y de tono, pues es producto de la presión infraglótica, que provoca ambos parámetros simultáneamente. Sin embargo, para Lieberman no puede hablarse más que de dos niveles acentuales (*stressed - unstressed*) en vocales aisladas; en el habla conectada, la sensación de más grados acentuales se debe al juego que se establece entre vocales reducidas (que no pueden ser acentuadas) y vocales plenas (que son las que pueden estar acentuadas o no): la sensación, pues, puede ser de tres grados (*reduced vowel - full vowel: unstressed - full vowel: stressed*, esquema parecido al propuesto por Bolinger).

Normalmente, el tono desciende al final de la espiración como consecuencia de la disminución de la presión infraglótica, a no ser que el hablante tense en ese momento las cuerdas vocales, lo cual da como resultado un aumento "artificial" del tono; así, Lieberman considera dos tipos de "grupo

espiratorio", semejantes a los TUNE I y TUNE II de Armstrong & Ward: el "grupo espiratorio no marcado" [-BG] (de final descendente) y el "grupo espiratorio marcado" [+BG] (de final ascendente). Por otra parte, el límite final de un grupo espiratorio siempre es una pausa, que es el momento en el que se inspira de nuevo el aire, en el que no hay fonación.

Este sería un esquema universal y provocado por la propia forma de producirse el tono, como fenómeno motivado por el juego entre la presión infraglótica y la tensión de las cuerdas vocales. Por tanto, para Lieberman no tiene sentido hablar de la entonación como de un fenómeno lingüístico independiente, sino que debe considerarse un fenómeno motivado, innato y universal.

Puesto que la entonación no es un fenómeno lingüístico *per se*, el grupo espiratorio ha de coincidir con la oración gramatical, como portador natural de la misma en el habla: esta sería, en todo caso, su función principal. Sin embargo, Lieberman señala que no siempre coinciden los límites de la oración gramatical con los del grupo espiratorio, y aventura dos posibles motivos: en primer lugar, si la oración es demasiado larga (cosa que ocurre a menudo) el hablante ha de respirar, por lo que segmentará la unidad gramatical en dos o más grupos espiratorios; en segundo lugar, en ocasiones es necesario desambigüizar una oración cuya estructura sintáctica superficial pueda corresponder a dos estructuras profundas distintas (por ejemplo, en frases como "no me gusta" y "no, me gusta", cuya única diferencia es que constituyen uno o dos grupos espiratorios, fenómeno al que suele llamarse "función demarcativa" de la entonación): esta también sería una función específicamente lingüística de la entonación, entendida como un procedimiento gramatical complementario.

Es decir, la entonación, como fenómeno no lingüístico y dependiente de

la estructura gramatical de la oración, no puede ser significativa: "el grupo espiratorio es un rasgo fonológico. No tiene un estatus semántico independiente. La interpretación semántica de una oración se sigue de la estructura profunda..." (1967: 121).

Para Lieberman, por tanto, un enfoque generativista del análisis de la entonación debe contemplar: a) su naturaleza innata, es decir, universal; b) su dependencia directa del nivel sintáctico; y c) el carácter no lingüístico y no significativo del fenómeno. Es decir, un enfoque mucho más cercano al análisis de configuraciones (y a autores como Bolinger) que a la tradición norteamericana.

Chomsky & Halle, sin embargo, toman otro camino, que sólo coincide con el emprendido por Lieberman en la dependencia de la entonación del nivel sintáctico, como vimos; por lo demás, estos autores no niegan su carácter significativo, aunque tampoco desarrollan el tema, ni insisten en la universalidad de las entonaciones (aunque sí, indirectamente, en la universalidad de las reglas de dependencia sintáctica).

Por otra parte, Lieberman se aplica a desmoronar con una severa crítica el sistema de análisis por niveles de la entonación, entendiendo que se trata del método estructuralista, es decir, el adversario contra el que debe ir su interpretación; sin embargo, Chomsky & Halle no sólo no atacan el sistema de niveles sino que lo asumen sin apenas modificaciones. Obviamente, Chomsky no es un militante de sus propias ideas tan rígido como algunos de sus seguidores, y no considera que sus trabajos de sintaxis puedan ser mecánica y directamente extensibles a la fonología; por el contrario, espera a contar con un colaborador de la talla de M. Halle (que también fue colaborador, anteriormente, de R. Jakobson) para atreverse a lanzar su propuesta renovadora, más

epistemológica que de detalle: en todo aquello que no afecta directamente a su planteamiento general, por tanto, se muestra modesto y acepta buena parte de la fonología clásica (el sistema binarista de oposiciones, la vuelta a los rasgos articulatorios, el análisis de niveles, etc.).

En resumen, el intento de Lieberman consiste en una interpretación de la entonación a partir de los postulados generales de la teoría sintáctica generativa, mientras que el trabajo de Chomsky & Halle constituye un cambio de paradigma fonológico (equivalente al cambio de paradigma gramatical y coherente con él) a partir del cual se crea una nueva fonología que paulatinamente irá sustituyendo los objetivos, los métodos y los conceptos de la fonología clásica por otros propios del nuevo paradigma; esta tarea, sin embargo, Chomsky & Halle no podrían llevarla a cabo ya en su primer trabajo, razón por la cual asumen buena parte de las descripciones y los métodos de la fonología clásica, entre ellos el análisis de niveles.

Como veremos, aunque la fonología generativa ha abandonado por completo, a lo largo de los años setenta, la tradición fonológica no generativa y se ha erigido casi en una disciplina nueva, en el campo del análisis de la entonación los autores generativistas no sólo no han abandonado el método tradicional norteamericano de análisis por niveles, sino que se han planteado su trabajo conscientemente como una continuación y una profundización del mismo, enriqueciéndolo con la metodología generativa pero también con otras aportaciones.

2.3.3. *La perspectiva métrica*

Los trabajos de Liberman & Prince "On stress and linguistic rhythm" (1977) y de Liberman (autor que no debe confundirse con P. Lieberman, citado en el apartado anterior) *The intonational system of English* (1978) inauguran una nueva perspectiva fonológica centrada en la interpretación del acento y la entonación como fenómenos rítmicos: la llamada "fonología métrica". Dicha perspectiva constituye una continuación natural de las reglas de acentuación propuestas por Chomsky & Halle (1968) y, aunque es en buena parte deudora de autores ajenos o anteriores a él, se enmarca claramente en el nuevo paradigma de la fonología generativa. También se erige, por otra parte, en la heredera moderna de la tradición norteamericana de análisis por niveles, aunque en sus trabajos Liberman & Prince en ningún momento citen explícitamente ni a Pike ni a Trager & Smith.

La idea fundamental de la perspectiva métrica es que el acento no viene dado por los parámetros acústicos propios del segmento acentuado, sino que se desprende de la posición que dicho segmento ocupa en la estructura rítmica de la frase. Es decir, no puede hablarse simplemente de acento de intensidad (*stress*) ni de acento melódico (*accent*), sino de "accento como fenómeno rítmico" (*stress-as-rhythm*).

Por otra parte, si el sistema fonológico generativo supone que la estructura sintáctica abstracta (estructura P) es previa y condiciona la estructura fonológica, Liberman & Prince profundizan en la idea afirmando que, aunque se trata de estructuras teóricamente independientes, la estructura sintáctica de superficie (estructura S) se planifica sobre la estructura rítmica: razón por la cual ambas estructuras son tan congruentes.

El primer autor que lanzó la idea del *stress-as-rhythm* fue Householder (1957), para quien "la gente oye cosas que no están ahí". Realizó un experimento en el que presentaba una serie de sonidos indistintos, de igual tono, intensidad, duración, etc., y preguntaba a los oyentes cuál era el sonido acentuado; estos, sorprendentemente, le señalaban siempre los mismos sonidos (siguiendo patrones binarios o ternarios), aunque no había ningún parámetro acústico que los distinguiera de los demás. En conclusión, Householder dedujo que el parámetro fundamental para percibir el acento es, precisamente, el esquema rítmico que el oyente proyecta sobre la frase. Como en el habla real el acento suele estar asociado a algún parámetro físico, cuya función sería justamente la de realzar el esquema rítmico, la sensación que podemos tener es que son la intensidad o el tono los responsables de la percepción del acento, pero en realidad sólo son los responsables de que detectemos el esquema rítmico general, no cada acento en particular.

Liberman & Prince llevan esta idea a sus últimas consecuencias y elaboran una serie de reglas métricas de asignación de acento que parten de la información gramatical que se encuentra en el lexicón de cada palabra empleada: se trata, pues, de reglas, como las de Chomsky & Halle, de dependencia categorial (*LCPR: Lexical Category Prominence Rule*, *NSR: Nuclear Stress Rule* y *RPPR: Relative Prominence Projection Rule*).

Ciertamente, al tratarse de reglas generativas, dan cuenta del proceso de constitución rítmica de la frase y de la asignación de la prominencia; el proceso de percepción, en cambio, queda peor explicado: según estos autores, los parámetros acústicos de inflexión tonal, duración, intensidad y cualidad vocálica, suelen estar asociados al esquema rítmico proyectado desde la estructura P, aunque no es necesario que lo estén; el oyente, apenas detecta una parte de esa

información física (en principio redundante), percibe el esquema rítmico abstracto y proyecta su conocimiento de él sobre la frase percibida; la prominencia, por tanto, está "en la cabeza" del oyente (diseñada para percibir selectivamente esquemas jerárquicos) y no en el propio sonido.

Como los parámetros acústicos "suelen" estar asociados al segmento prominente, los fonetistas creen, erróneamente, que son estos la causa de su percepción: y es que ningún investigador había tenido en cuenta hasta entonces (excepto Householder) la importancia del esquema rítmico para la percepción del acento; es decir, todos los experimentos sobre el tema estaban mal planteados y sus resultados, por tanto, son erróneos.

De este modo, sin duda, la controversia sobre el "acento melódico" y el "acento de intensidad" (*accent - stress*) queda superada por nuevas evidencias teóricas (sólo asumibles, sin embargo, desde una perspectiva generativa, en la que la estructura sintáctica sea previa, como hemos visto, a la estructura fonológica y a la forma fonética) y puede mantenerse la división tradicional entre *stress* y entonación, formulada como *stress-as-rhythm* y entonación, a la vez que pueden explicarse las "supuestas" evidencias instrumentales según las cuales el acento no es de intensidad.

Nótese, sin embargo, que el término empleado no es *accent-as-rhythm*, sino *stress-as-rhythm*: en última instancia, siempre es el parámetro de intensidad el más informativo, como se demuestra en música, en la que melodía y duración de las notas son aspectos independientes del ritmo, que está marcado, por tanto, por la intensidad relativa de las primeras notas de cada compás.

El trabajo de Liberman & Prince, sin embargo, está centrado sólo en las reglas métricas de acentuación. El tratamiento métrico de la entonación en su

conjunto lo llevará a cabo Liberman un año después (1978). En este trabajo, después de plantear la cuestión del *stress-as-rhythm* como eje central de su descripción, Liberman adopta un análisis de la melodía entonativa mediante niveles tonales. El grupo rítmico define al grupo de entonación, cuya melodía constituye un "morfema tonal" distintivo, significativo: así, Liberman introduce en su trabajo elementos característicos de la tradición británica, como el concepto de "lexicón entonativo" y la lista de significados que ofrecen O'Connor & Arnold (1969) y Crystal (1969), a los que sigue de cerca, y cuyos contornos pretende describir mediante niveles (no se trata, por tanto, de un intento original como el de Pike, sino que en esta ocasión ya se parte de una lista previa de significados y el objetivo es únicamente su interpretación formal).

Los niveles son fonemas tonales definidos mediante dos rasgos tonales distintivos: [alto] y [bajo] (*[high]* - *[low]*). Liberman, como los demás autores de la tradición, describe cuatro niveles o fonemas tonales:

H (*high*), cuyos rasgos serían: *[+ high - low]*.

HM (*high-mid*), con los rasgos: *[+ high + low]*.

LM (*low-mid*), con los rasgos: *[- high - low]*.

L (*low*), con los rasgos: *[- high + low]*.

Sin embargo, no se trata exactamente de los mismos cuatro niveles de Trager & Smith, pues para estos autores el nivel más alto estaba reservado al tono *overhigh*, menos usual que los otros niveles, mientras que para Liberman los cuatro niveles que propone juegan un papel igualmente importante en la descripción de los contornos.

Aunque el carácter fonemático de la descripción de Liberman es bien

patente desde el principio, en realidad otorga tanto valor a los propios fonemas tonales como a los dos rasgos que los caracterizan: de modo que puede describir la melodía de los morfemas entonativos (lo significativo de la entonación no es, para Liberman, ni el núcleo ni la juntura, sino la melodía completa) no sólo como una sucesión de fonemas, sino también como una sucesión de rasgos que permite ciertas variaciones fonemáticas. Por ejemplo, la melodía del morfema de "sorpresa o redundancia" es definida como la sucesión de los rasgos:

[- high] [+ high] [- high]

Puede haber, por tanto, ocho realizaciones concretas del morfema, pues [- high] puede ser **L** o **LM** y [+ high] puede ser **H** o **HM**: cada realización concreta sería una especie de "alomorfo entonativo". Esta posibilidad de describir las unidades entonativas con verdaderos márgenes de dispersión constituye una novedad absoluta en el análisis de niveles (en Trager & Smith, cada secuencia concreta de niveles tenía que ser significativa y distintiva) y una aportación fundamental al estudio de la entonación.

El trabajo de Liberman sirvió de base para que Pierrehumbert elaborara un sistema de generación de modelos de entonación basada en la generación de modelos acentuales en *The phonology and phonetics of English intonation* (tesis leída en 1980, año con el que a menudo se la cita, pero publicada en 1987).

Pierrehumbert parte de una representación métrica de la frase semejante a la propuesta por Liberman & Prince (1977) y Liberman (1978), sobre la que hace recaer una serie de tonos altos (*high*: **H**) o bajos (*low*: **L**) para describir la melodías que ocurren en inglés: se trata, por tanto, de formalizar los datos que ofrecen otros autores, ofreciendo una representación subyacente de la entonación de las frases, sin detenerse en ningún análisis semántico y sin reparar en lo

"distintiva" que pueda resultar la melodía formalizada. A Pierrehumbert le interesa, fundamentalmente, describir el modo de asignar una melodía determinada a un texto determinado, es decir, las reglas mediante las cuales un texto dado genera una entonación determinada, y cómo textos de distinta duración y estructura pueden proyectar melodías similares.

En este sentido, el trabajo de Pierrehumbert es la culminación del análisis generativo de la entonación, el punto al que tendían los trabajos anteriores, pues constituye un modelo realmente generativo y no descriptivo de la entonación, y porque asume completamente algunos de los conceptos básicos de la fonología generativa aún no bien asumidos por los demás autores: por ejemplo, en Pierrehumbert se concibe una estructura gramatical completa y previa a la estructura fonológica (Lieberman sostenía aún que la estructura S es independiente y posterior al esquema rítmico); la entonación es considerada como un fenómeno estrictamente fonológico, y por tanto no significativo (Lieberman añade a su análisis un tratamiento de los significados similar a los de la tradición británica); la dependencia categorial afecta tanto a la melodía como al acento; etc.

El sistema de niveles propuesto por Pierrehumbert es extremadamente económico: únicamente dos niveles (Lieberman hablaba de dos rasgos, pero aún conservaba los tradicionales cuatro niveles), cuya altura es siempre relativa y dependiente del nivel anterior. Es decir, no se trata de dos niveles más o menos estables, sino de dos movimientos tonales: ascenso o descenso. Toda melodía consiste, pues, en una sucesión de ascensos y descensos, es decir, de niveles **H** y **L**, cuya altura relativa tiene que ver con el nivel del tono precedente, con el grado de prominencia del tono (si es el centro de un acento o es un tono adyacente) y con la "línea base" (*baseline*) del hablante.

Esta "línea base" es el nivel más bajo posible de producción tonal del hablante en un momento determinado del sintagma entonativo; este nivel más bajo no es constante a lo largo de todo el sintagma, sino que va descendiendo gradualmente a lo largo de la emisión de voz, fenómeno conocido como "declinación": a lo largo de la emisión desciende, fundamentalmente, el nivel de los sucesivos acentos de la frase, aun en el caso de que el hablante pretendiera emitirlos al mismo nivel. Sin embargo, para Pierrehumbert el grado de "declinación" entonativa no es necesariamente proporcional a la longitud del texto, y no hay, por tanto, lo que se ha dado en llamar "preplanificación" (concepto según el cual el hablante planificaría la estructura de la entonación en función de la longitud de su emisión: Ladd, 1984).

Además de los dos niveles tonales relativos, Pierrehumbert añade otros elementos al sistema. En primer lugar, y una vez ha quedado claro que el acento no es un fenómeno tonal sino rítmico, puede describirse su participación en la melodía en términos tonales: los "acentos melódicos de frase" (*phrase accent*, que vienen a ser como los "núcleos" de la melodía) pueden ser, así, de un solo tono (**H*** y **L***) o bien bitonales, es decir, consistentes en una inflexión tonal: **H** + **L***, **L*** + **H**, **H*** + **L**, **L** + **H*** y **H*** + **H** (el tono con asterisco es el centro del acento y los tonos **H** **L** son tonos "de arrastre"). También se habla de un "acento de frase" (*phrase stress*) que consiste en la última parte de la melodía, desde el acento melódico o núcleo hasta el final: una especie de *tail*. Finalmente, se consideran dos límites o junturas (**H%** y **L%**), en la última sílaba de la frase.

Se trata de un sistema, pues, que en cierto modo añade al análisis de niveles tradicional una parte del análisis británico basado en tonos nucleares: los *phrase accent* bitonales son similares, por ejemplo, a los *nuclear-tone* que vimos

en Cruttenden; aunque con la diferencia, muy relevante, de que no son tratados como una unidad sino descompuestos en dos niveles sucesivos, consecuentemente con el estilo norteamericano de no atender a las gradaciones tonales sino a la dirección y al nivel resultante del movimiento tonal.

2.4. La escuela holandesa

2.4.1. El enfoque perceptivo

En 1957 se creó en Eindhoven el Instituto de Investigación Perceptiva (IPO, según las siglas en holandés), donde, desde principios de los años 60, se lleva a cabo un programa de investigación dedicado a la caracterización "psicoacústica" de la entonación. Después de más de veinte años de investigación en equipo, durante los cuales se publicaron algunos trabajos parciales sobre la entonación fundamentalmente del holandés, los investigadores más relevantes del centro, J. 't Hart (director), R. Collier y A. Cohen (las tres figuras centrales de lo que ha dado en llamarse la "escuela holandesa" de análisis de la entonación) publicaron, en 1990, *A perceptual study of intonation*, compendio de todos los trabajos anteriores y síntesis de las líneas maestras de su enfoque.

Genéricamente, podría decirse que la escuela holandesa se inscribe dentro del marco del "análisis de configuraciones", aunque constituye una línea de investigación muy diferente a la tradición británica. En realidad, constituye una línea de investigación diferente a cualquier otra, cuyas características principales serían: el trabajo en equipo, encuadrado en un centro específico de investigación; la investigación centrada en la dimensión perceptiva de la entonación; y el método de análisis experimental, estrictamente fonético.

Por primera vez en la historia, por tanto, un método de análisis de la entonación no depende de las opiniones de un solo autor independiente más o menos ligado a una tradición, sino que es el patrimonio de un equipo de investigación coherente y diversificado, dotado de medios instrumentales y que

sigue una línea de trabajo unitaria.

Por otra parte, el hecho de seguir una metodología experimental, centrada en la percepción, y de atender exclusivamente a los fenómenos fonéticos de la entonación resulta sorprendente en un campo en el que la representación, y el propio análisis, siempre han sido "impresionistas", lejos de toda voluntad de exactitud, porque el objetivo era "fonológico": describir la forma y el uso de las entonaciones distintivas, o significativas. Para 't Hart, Collier & Cohen, por el contrario, lo importante es describir las entonaciones más comunes, y no entrar en disquisiciones fonológicas o semánticas que, según su opinión, sólo sirven para empobrecer los datos con los que implementar, por ejemplo, un sintetizador de voz.

Para describir y formalizar una curva entonativa han de eliminarse todas las variaciones de tono imperceptibles para el oyente (las variaciones "micromelódicas") pero que complican el análisis acústico hasta el infinito: hay que extraer, pues, los "movimientos tonales" (*pitch movements*) estrictamente necesarios para que el oyente pueda reconocer exactamente la misma curva de entonación.

El método central del enfoque perceptivo (del "enfoque *IPO*", como sus propios autores lo llaman) es la "estilización" de las curvas entonativas: considerando el "movimiento tonal" (ascenso/ descenso) como el elemento mínimo de la melodía, se trata de obtener el esqueleto de movimientos tonales perceptibles e identificables por los oyentes, mediante un sistema progresivo de reducción de las curvas.

El sistema consiste en una primera medición instrumental, tras la cual se sintetiza un "copia ajustada" (*close-copy*), casi exacta, de la curva, de la que se

han eliminado sólo las variaciones micromelódicas: es decir, las fluctuaciones de la melodía debidas a la influencia tono intrínseco de las vocales y a fenómenos de coarticulación (v. figura 31).

Por definición, la *close-copy* es auditivamente exacta a la curva original, no puede distinguirse de ella, y debe ser validada, por tanto, por una prueba de equivalencia perceptiva en la que un número importante de oyentes

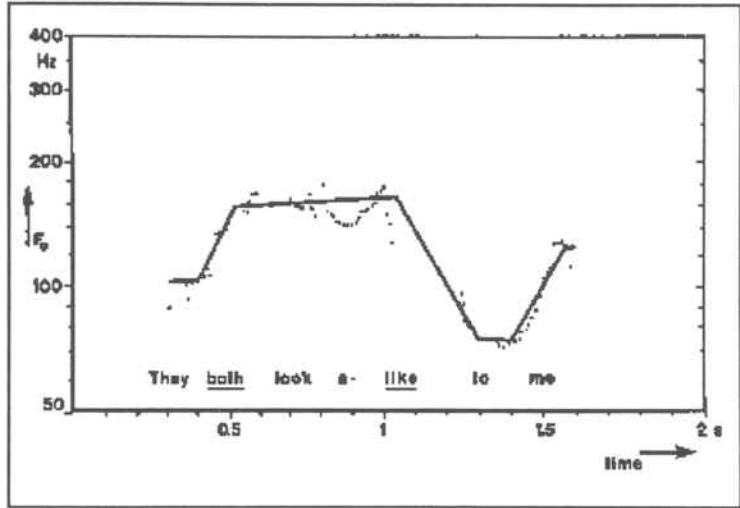


Figura 31 - "Copia ajustada" de una entonación.

independientes la reconozcan como "igual" a la original.

El siguiente paso es elaborar la "estilización estandarizada" de la curva. Para ello se realizan "copias ajustadas" de la misma curva grabada a varios hablantes, se estandarizan la duración, el recorrido tonal (*size*) y la posición de los movimientos tonales correspondientes de cada hablante y se crea en el sintetizador una nueva curva con el promedio de todos los valores (v. figura 32).

La "estilización estandarizada" resultante se somete de nuevo a una prueba auditiva, con la cual

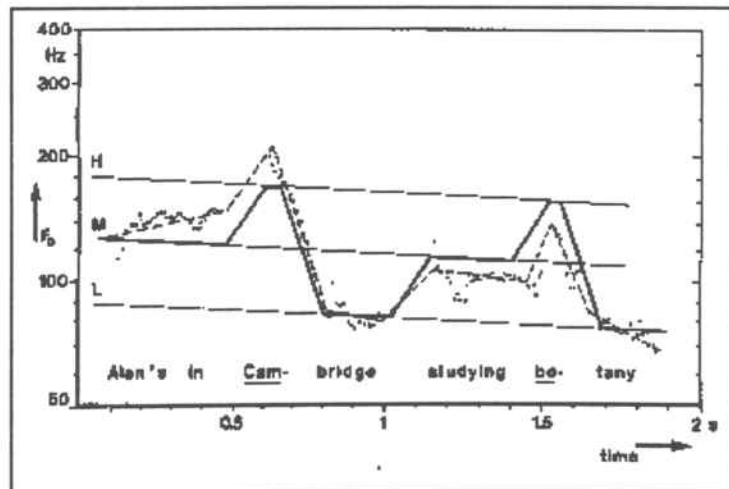


Figura 32 - "Estilización estandarizada" de una entonación.

se valida su equivalencia perceptiva con las curvas originales. El proceso de elaboración y síntesis de las curvas estilizadas es, obviamente, largo y complejo, pero el resultado es sorprendente, pues se obtienen curvas sintetizadas en las que están todos los elementos relevantes de las curvas originales emitidas por los informantes, pero no las fluctuaciones melódicas irrelevantes.

La estilización que llevan a cabo 't Hart, Collier & Cohen es exclusivamente melódica y no considera más elementos que las meras variaciones de F_0 , es decir, no contempla la intervención del acento. Frente a los demás métodos de análisis, muy preocupados por el papel que desempeña el acento, que definen la entonación como un producto condicionado por el esquema acentual de la frase, o bien que consideran al acento como el núcleo de la entonación y su parte más informativa, etc., la escuela holandesa se aplica a describir la entonación únicamente como la sucesión de valores tonales, la melodía del habla. En realidad, en ningún momento se excluye teóricamente la posibilidad de que haya otros factores que interesen a la descripción lingüística de la entonación: simplemente, a los autores sólo les interesa la melodía como fenómeno físico y sus elementos perceptivamente relevantes. Así, el carácter no fonológico del enfoque holandés es paralelo al abandono de los factores no melódicos (acento, junturas, etc.) en la descripción de las curvas entonativas.

Coherentemente con el objetivo de describir la melodía del habla, la unidad de medida empleada es el "semitono": unidad relativa con la que cuantificar la distancia entre un valor frecuencial y el valor anterior. El análisis en valores absolutos, ciertamente, no podría ser eficaz, pues no interesa la altura concreta de cada punto de la melodía, sino la propia melodía, es decir, la altura relativa de cada punto. En este sentido, el avance que supone la escuela holandesa no reside en describir la melodía como una sucesión de valores

relativos (en todos los enfoques subyace un procedimiento similar) sino en cuantificar con datos precisos tales valores relativos. Hasta el momento, todas las descripciones eran impresionistas, y ningún autor se había aventurado a señalar valores precisos: el análisis de configuraciones, por ejemplo, representa una línea no ubicada en ninguna parrilla de valores numéricos; el análisis de niveles, por su parte, representa la entonación como una sucesión de valores numéricos que, sin embargo, no tienen ninguna correspondencia con los valores frecuenciales de la curva, ni en términos absolutos ni en términos relativos. Por tanto, la descripción de 't Hart, Collier & Cohen es el primer intento serio de cuantificar y definir con exactitud las melodías del habla. En realidad, las melodías están definidas con tal precisión que pueden incluso sintetizarse melodías perceptivamente iguales.

2.4.2. Principios teóricos

't Hart, Collier & Cohen ofrecen en su libro una serie de proposiciones o principios con los cuales definen su posición teórica con respecto a la entonación, en lo que viene a ser una exposición sistemática de su "teoría de la entonación".

En primer lugar, afirman que en la percepción del tono en el habla conectada el oyente sólo es sensible a los cambios de F_0 que el hablante ha producido intencionadamente. Es decir, todas las variaciones de F_0 producidas colateralmente, en las que no interviene la voluntad del hablante, son variaciones "micromelódicas", entonativamente despreciables, e imperceptibles.

Todos los rasgos de los movimientos tonales, según los autores, están

controlados directamente por los músculos laríngenos. No son, pues, fruto de la coarticulación o de la casualidad, sino que obedecen a la voluntad manifiesta del hablante. De este modo, el planteamiento estrictamente fonético en la descripción de la entonación no supone una negación de su carácter fonológico, sino simplemente su no consideración: en realidad, en la afirmación de que todo movimiento tonal perceptivamente relevante es intencionado subyace una concepción fonológica del fenómeno.

Generalmente, tales cambios tonales ocurren en forma de *glides*, de "ligaduras" tonales, y no de "saltos". Así, si en la tradición británica se distingue bien entre *glide* y *jump*, atribuyendo a cada uno una función distinta; y si en la tradición norteamericana se considera irrelevante la diferencia, suponiendo que los cambios de nivel generalmente son "saltos" tonales; en la escuela holandesa se considera la *glide* tonal como el fenómeno más común, que por tanto se toma como base para sintetizar las frases estandarizadas.

Por otra parte, la unidad más pequeña de análisis perceptivo es el "movimiento tonal", que puede descomponerse en "rasgos tonales" tales como la dirección (*direction: rise / fall*), su duración con respecto al límite silábico (*timing: early / late / very late*), la velocidad (*rate of change: fast / slow*) y el recorrido tonal (*size: full / half*).

		<i>Transcription symbol</i>									
		1	2	3	4	5	A	B	C	D	E
<i>Direction</i>											
<i>rise</i>		x	x	x	x	x					
<i>fall</i>							x	x	x	x	x
<i>Timing</i>											
<i>early</i>		x				x		x			x
<i>late</i>				x			x				
<i>very late</i>			x							x	
<i>Rate of Change</i>											
<i>fast</i>		x	x	x		x	x	x	x	x	x
<i>slow</i>					x						x
<i>Size</i>											
<i>full</i>		x	x	x	x		x	x	x	x	
<i>half</i>						x					x

Figura 33 - Movimientos tonales del holandés.

En la figura 33 (op. cit.: 73), se describen los

movimientos tonales del holandés (designados con los símbolos "1, 2, 3, 4, 5, A, B, C, D, E") según sus rasgos perceptivos.

Cada uno de los movimientos tonales descritos, a pesar de su aparente parecido con las transcripciones tradicionales (es numérico como el análisis de niveles y emplea letras mayúsculas como el análisis de Bolinger) está cuantificado y es el resultado de numerosas mediciones y estandarizaciones melódicas: así, por ejemplo, el movimiento tonal "1" se caracteriza por un ascenso (*rise*) de 5 semitonos, una duración de 120 ms. y un *timing* de 50 ms. (es decir, entre el comienzo de la vocal y el pico tonal de la inflexión que ocurre en ella pasan 50 ms.).

Los movimientos tonales pueden combinarse según una serie de restricciones, propias de cada lengua, que podría considerarse la "gramática" de la entonación. Los movimientos tonales combinados forman las "configuraciones", que tienen tres partes: *Prefix* (uno o más en cada configuración, que aparecen de un modo optativo), *Root* (la parte más relevante, en ocasiones constituye ella sola la configuración) y *Suffix* (también de carácter optativo).

A su vez, las configuraciones se combinan entre sí, obedeciendo también a ciertas restricciones "sintácticas", y forman los "contornos". Hay un número ilimitado de contornos en cada lengua, que son la manifestación concreta de un número de modelos de entonación básicos.

Puede considerarse, pues, que los autores sí contemplan, finalmente, un tipo de abstracción fonológica, los "modelos básicos", de los que, sin embargo, no se especifica ninguna función comunicativa o lingüística concreta.

Ya hemos dicho que este tipo de análisis puede encuadrarse en el bloque

de "análisis de configuraciones": en la figura 34, en la que se presentan los modelos de las configuraciones del holandés, puede observarse la correspondencia entre los símbolos que se refieren a los movimientos tonales, los rasgos de la configuración y la curva estilizada.

Ya hemos visto que en la escuela holandesa la cuestión del acento no recibe un tratamiento comparable al que recibe en las otras tradiciones. 't Hart, Collier & Cohen llaman la atención, únicamente, sobre el fenómeno, observado experimentalmente, de que sólo ocurren movimientos tonales en vocales acentuadas. Sin embargo, no entran en la disputa sobre si hay un acento melódico y otro dinámico, etc., porque les basta con esta observación: los movimientos tonales ocurren en las vocales tónicas.

Puesto que lo que la entonación reside en la melodía, el hecho de que esta "repose" sobre la estructura de las vocales tónicas de la frase no invalida su

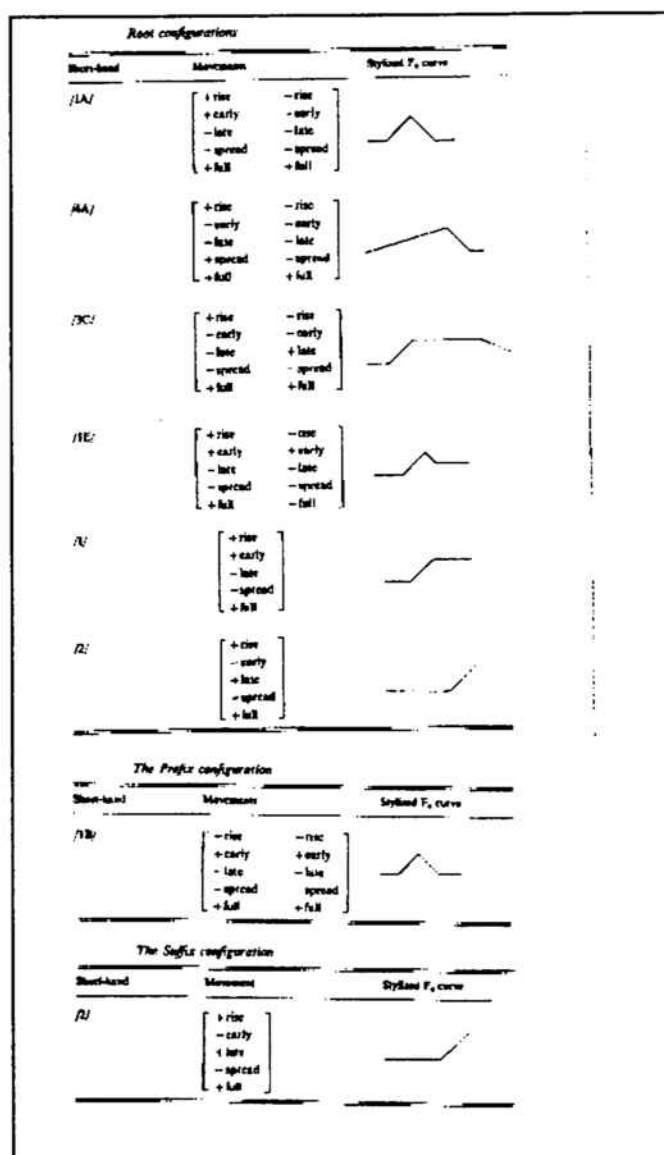


Figura 34 - Modelos básicos de la entonación del holandés.

preponderancia sobre el fenómeno de acentuación: un acento puede estar aquí o allá, y justamente sobre él ocurrirá cada movimiento tonal; sin embargo, el que la vocal acentuada esté antes o después, sea una o sea otra, no modifica en absoluto la estructura melódica de la configuración. Es decir, la misma melodía puede ocupar frases cuya estructura acentual sea completamente distinta.

Nótese la notable diferencia de enfoque que esto supone con respecto a los análisis métricos, en los que era la estructura acentual la que generaba la estructura entonativa.

Por otra parte, si en el análisis británico la única parte informativa de la melodía es el núcleo de la configuración, en el análisis holandés nos encontramos con que toda la melodía de la configuración es igualmente relevante: así, el análisis "nuclear" (como el británico) y el análisis "melódico" u "holístico" (como el holandés) son las dos principales modalidades que puede adoptar el "análisis de configuraciones", como veremos más adelante (2.6.).

Según el enfoque *IPO*, la forma concreta de una configuración tiene que ver con la longitud de la frase y con el fenómeno de la "preplanificación": puesto que el tono de la frase "declina" a lo largo de la emisión, el hablante planifica la longitud de la frase y sitúa el primer movimiento tonal a una altura suficiente para que la paulatina declinación en la que se inscriba la melodía permita llegar al final en condiciones de aceptabilidad. El fenómeno de la declinación es importante para el modelo de curvas estilizadas, porque es lo que permite sintetizar una entonación y que no suene muy artificial: a lo largo de toda emisión, los límites tonales entre los que se inscribe la melodía descienden paulatinamente. Es decir, una serie de movimientos tonales "a la misma altura" en realidad estarán escalonados, cada uno más bajo que el anterior.

El hablante, por tanto, tiende a planificar su entonación y a integrar en ella los fenómenos de coherencia sintáctica y acentual.

En cuanto a la correspondencia entre la entonación y la sintaxis, con todo, los autores señalan que ciertamente es muy común, por la solidaridad que suele establecerse entre los contornos y los sintagmas gramaticales, y aceptan que en ocasiones pueda ser la entonación la encargada de desambigüizar una frase. No obstante, puntualizan que no se trata de una relación obligatoria ni exclusivista: es decir, la entonación no depende en ningún caso de la sintaxis, ni tiene asociada curvas específicas a estructuras gramaticales específicas. Por tanto, toda relación entre ambos fenómenos será "casual", opcional, pero no sistemática. En una serie de experimentos, por ejemplo, demuestran cómo otros fenómenos fonológicos no entonativos pueden contribuir igualmente a desambigüizar una frase o a marcar los límites de un sintagma, y cómo la entonación puede no cumplir estas funciones.

Finalmente, es interesante llamar la atención sobre una de las características más importantes del análisis de la escuela holandesa, pues define bien el carácter rigurosamente instrumental de su enfoque: la resistencia a tratar sobre el tema del significado de la entonación. La entonación, afirman, no conlleva ningún significado intrínseco en el nivel en el que se describen las configuraciones.

En realidad, se trata de una actitud coherente con el enfoque estrictamente fonético que llevan a cabo, en el que se desiste de entrada hacer una interpretación fonológica o funcional y donde no caben, por tanto, explicaciones semánticas de tipo "actitudinal" o "pragmático".

En resumen, los trabajos de la escuela holandesa constituyen un paso importante en la descripción fonética de la entonación, que hasta entonces sólo había recibido interpretaciones más o menos impresionistas no contrastadas experimentalmente. En ellos se realiza una tarea de cuantificación precisa y de reducción de los datos (estilización) que permite enfrentarse no ya a un terreno selvático y pantanoso sino a un jardín cuidado, en el que los senderos por los que transitar aparecen limpios. Sin embargo, ese carácter de descripción exclusivamente fonética también conlleva algún lastre, como es no considerar el rendimiento lingüístico de las melodías descritas, las unidades funcionales que puedan constituir, así como su posible significación independiente.

2.5. La escuela española

2.5.1. La doctrina de T. Navarro Tomás

En 1918, T. Navarro Tomás publica su *Manual de pronunciación española*, obra de referencia y fuente principal de toda la moderna fonética española (y aun de la fonología, como vimos en el apartado 1.2.2.) que, sin embargo, fue pensada con fines estrictamente pedagógicos y utilitarios: con indicaciones de uso, ejercicios prácticos, textos fonéticos, glosario, etc. En él encontramos ya un capítulo completo dedicado a la entonación, que se caracteriza por la riqueza de datos ofrecida en tan breve espacio, y en el que se apuntan algunas de las características de la doctrina que su autor desarrollaría años después: análisis de la entonación por configuraciones, perspectiva fonética y eminentemente didáctica, empleo de medios instrumentales (quimógrafo y grabadora), especial atención a la "entonación literaria", etc. Sin embargo, dicho capítulo no es más que una primera aproximación, muy somera, y en la que aún Navarro Tomás no había dado forma definitiva a su doctrina. Casi treinta años después, en 1944, publicó su *Manual de entonación española*.

Presentado como una obra en la que se desarrollan "las nociones anticipadas en el *Manual de pronunciación española*, en el capítulo concerniente a la entonación" (1944: 7), en realidad es una reformulación completa de tales "nociones", y constituye una obra original y unitaria. Por otra parte, a sus fines didácticos añade, por primera vez en su obra, un objetivo eminentemente científico: "uno de los propósitos de este trabajo ha sido precisamente presentar en forma organizada y metódica una serie de cuestiones que puedan promover el interés hacia un campo de estudio de naturaleza tan rica como poco cultivada"

(ibid.). Así, si por un lado se trata de un manual dedicado a la enseñanza práctica de la lengua (como ocurría con los primeros trabajos de la tradición británica, hecho que, como vimos, condicionó su posterior desarrollo), por otro lado contempla también la dimensión científica y se presenta como un tratado teórico de la entonación española (como los trabajos de la tradición norteamericana, siempre de carácter teórico y que rara vez perseguían fines didácticos): la escuela española de análisis de la entonación, pues, nace con una obra compleja, en la que se expone una doctrina teóricamente muy bien trabada y en la que se ofrece, a la vez, un corpus muy nutrido de entonaciones analizadas.

Sin embargo, no han sido pocos los autores que han considerado la obra de Navarro Tomás más como una fuente de datos que como una fuente de conceptos, y se ha buscado en otras tradiciones el aparato teórico que no se supo encontrar en ella o, simplemente, que no se supo repensar y reformular en el momento adecuado. En realidad, si bien prácticamente todos los autores españoles siguen la tipología de frases propuesta por Navarro Tomás, su doctrina apenas ha sufrido variaciones desde que fue formulada y ha sido o bien repetida sin mucha convicción o bien ignorada completamente.

Un ejemplo de ello lo tenemos en el *Esbozo de una nueva gramática de la lengua española* (1973), obra firmada por la Real Academia pero redactada por S. Gili Gaya y S. Fernández Ramírez, en cuyo capítulo dedicado a la entonación se sigue de cerca la tipología de frases de Navarro Tomás, algo abreviada, frases en cuya descripción concreta, sin embargo, se emplea un curioso método de análisis por niveles, como veremos. Algo parecido ocurre en la *Gramática española* (1975) de J. Alcina y J. M. Blecua, en la que también se sigue la tipología de Navarro Tomás, pero que opta por describir la

entonación de cada frase siguiendo a J. H. Matluck (del que trataremos en el apartado siguiente), es decir, otra vez según el método de análisis por niveles.

La obra de Navarro Tomás, con todo, es muy apreciable teóricamente y pudo haber tenido más predicamento del que a la postre ha tenido, pues los conceptos que ofrece son de una entidad comparable al menos a los que por aquel entonces se formulaban en las dos tradiciones mayores ("tradiciones", pues, y no "escuelas", porque no dependen del magisterio de un solo autor, sino del trabajo encadenado de numerosos autores a lo largo de varias décadas).

El primer rasgo que consideraremos de la doctrina de Navarro Tomás es el carácter independiente que atribuye a los fenómenos del acento y la entonación. En español, el acento de la palabra es un fenómeno de intensidad (en el *Manual de pronunciación española*, el capítulo dedicado al acento se titula, precisamente, "Intensidad"); el "acento dinámico o de intensidad" divide las sílabas, pues, en "fuertes" y "débiles". Navarro Tomás señala, sin embargo, que la sílaba "fuerte" suele ser también "tónica" (es decir, de un tono más alto) y la sílaba "débil" suele ser "átona": la razón de que esto sea así es puramente fisiológica, ya que ambos fenómenos dependen de la tensión de las cuerdas vocales, y ocurre siempre en las palabras aisladas, pues en la frase la melodía es un fenómeno superpuesto e independiente del acento de intensidad.

Así, para Navarro Tomás no es necesario postular un acento de frase o núcleo en el que recaiga la inflexión fundamental de la entonación, etc., con todos los problemas que ello supone a los autores británicos (para quienes dicho núcleo es un fenómeno entonativo) y norteamericanos (para quienes es un fenómeno acentual): Navarro Tomás separa acento y entonación netamente, como fenómenos independientes, uno de intensidad y otro de melodía, y evita

la controversia causada por el concepto "acento de frase", simplemente ignorándolo.

En realidad, la tradición norteamericana, en la que el concepto de *sentence stress* no juega ningún papel determinante dentro de la estructura propiamente melódica de la frase, también podía haber ignorado el problema, y la controversia sólo tiene sentido si se considera: a) que los análisis norteamericanos siempre fueron a remolque de las descripciones británicas, cuyos *corpora* emplearon sistemáticamente; y b) que los autores norteamericanos siempre han intentado validar su enfoque metodológico en oposición teórica a los trabajos británicos. Sin embargo, en el análisis de niveles la descripción de las entonaciones considera toda la melodía como una unidad estructural, y no hay un verdadero núcleo, como en el análisis británico de configuraciones: lo natural hubiera sido ignorar la concepción "nuclear" de la entonación.

En el análisis de configuraciones europeo, que es el que aplica Navarro Tomás, tampoco se define la entonación en función de un núcleo, sino que se considera la melodía de la frase como un todo. Por tanto, y puesto que Navarro Tomás elabora su método de análisis independientemente de otras tradiciones, en ningún momento se deja llevar por inercias ajenas a su propio enfoque. Sí detecta la concordancia sílaba fuerte-tónica / sílaba débil-átona; pero, al no postular ningún núcleo entonativo, los fenómenos de acento y entonación pueden considerarse independientes sin estridencias.

Por otra parte, Navarro Tomás siempre negó las evidencias experimentales del carácter tonal del acento en español: respondió con mucha energía (1964), por ejemplo, al trabajo de H. Contreras "Sobre el acento en español" (1963), en el que se ponían de manifiesto tales evidencias. Lo cierto es que, aunque en la comunidad científica nadie pone ya en tela de juicio dicho carácter tonal del

acento, aún se mantiene la terminología empleada por Navarro Tomás incluso en autores tan significados como A. Quilis, en cuyos trabajos se muestra, por ejemplo, la poca relevancia que tiene la intensidad en el acento español (1971, 1981: 330-332), a la vez que se postula, paradójicamente, que "los fonemas acentuales ... en español son dos: el **fuerte** /ˈ/, que es el que se marca, y el **débil** /ˌ/, que normalmente no se señala" (el subrayado es nuestro) (1981: 416; cfr. también 1975, 1985): la impronta de Navarro Tomás es formidable.

Para Navarro Tomás, pues, la entonación se circunscribe a la estructura melódica de la frase, al margen de los fenómenos acentuales; se trata, por tanto, de una perspectiva similar a la que vimos en la escuela holandesa: ambos enfoques, en efecto, se encuadran en el análisis de configuraciones europeo, "melódico", frente al análisis de configuraciones británico, "nuclear".

También, como en la escuela holandesa, la unidad de medición empleada por Navarro Tomás es el semitono; sin embargo, en el autor español no se considera una medida relativa, con la cual definir la relación entre los puntos sucesivos de una melodía, sino una medida absoluta: en este sentido, la comparación de la melodía del habla con la melodía musical en Navarro Tomás va más allá que en ningún otro autor moderno. Ya vimos en el apartado 1.3.3. cómo la escala musical es un producto cultural muy marcado que no puede servir de referencia para medir las sutiles variaciones de la entonación; sin embargo, Navarro Tomás la emplea para determinar con exactitud la amplitud del movimiento melódico entre cuyos límites se enmarca la entonación de una lengua o un dialecto: es el concepto de "campo de entonación".

"Cada lengua -dice el autor- desarrolla habitualmente sus inflexiones dentro de una zona musical de determinada extensión" (1944: 24). El concepto

de "campo de entonación", por tanto, no debe confundirse con el de "campo tonal" (*accent range*) que se maneja en la tradición británica y que vimos en el apartado 2.2.3.: mientras este se refiere al recorrido tonal de la inflexión del núcleo del contorno (y sólo tiene sentido, pues, en un análisis "nuclear" de la entonación), aquel se refiere a los límites tonales entre los que suele moverse la entonación (y estaría más cerca, por tanto, del concepto de "tesitura" *-key-*, aunque referido no a una sola persona sino al propio código en su conjunto).

Un fenómeno semejante al "campo de entonación" es la división musical de las voces en el canto: soprano, de Do_3 a La_4 ; contralto, de La_2 a Sol_4 ; primer tenor, de Fa_3 a Re_4 ; segundo tenor, de Do_2 a La_3 ; barítono, de La_1 a Fa_3 ; y bajo, de Fa_1 a Re_3 (íbid.).

Así, para la entonación española, Navarro Tomás indica que el campo de entonación, "en el modo ordinario de la conversación", oscila entre Mi_1 y Sol_2 , mientras que "la pronunciación declamatoria o enfática, dilatando principalmente las inflexiones agudas, llega a alcanzar hasta dos octavas, $Re_1 - Re_3$ " (op.cit.: 25). En la voz femenina, añade, estos valores son similares, aunque una octava más alta, aproximadamente.

Sin embargo, "ni los tonos agudos ni los graves son los más frecuentes en el desarrollo del discurso. Dentro del campo de la entonación la línea melódica se produce en su mayor parte alrededor de una altura determinada que aparece como base o eje de los movimientos de la voz" (op.cit.: 27): es lo que se llama "tono normal", y "responde en cada individuo a la tendencia al equilibrio entre los músculos tensores y distensores de los órganos de fonación" (íbid.). En español, el tono normal en la voz masculina oscila entre Mi_2 y Fa_2 , "es decir, entre la altura de bajo y la de barítono" (op.cit.: 28).

El concepto de "tono normal" tiene varios equivalentes en otros enfoques,

fundamentalmente el concepto de "línea base" (*baseline*), empleado en la tradición norteamericana, a partir del cual se elabora el concepto de "declinación" (v. 2.3.3.).

En su *Manual de pronunciación española*, Navarro Tomás define el "grupo fónico" como "la porción de discurso comprendida entre dos pausas o cesuras sucesivas de la articulación" (1918: 30), y lo considera la unidad entonativa básica: "simplificará la enseñanza práctica de nuestra entonación tomar como guía la entonación del grupo fónico, la cual, sobre ser por sí misma más fácil de distinguir que la de otras unidades fonéticas, es en realidad la que más influye en la determinación de la estructura musical propia de cada frase" (op.cit.: 212).

Sin embargo, en el *Manual de entonación española* propone como unidad entonativa la "unidad melódica": "la unidad melódica es la porción mínima de discurso con forma musical determinada, siendo al propio tiempo una parte por sí misma significativa dentro del sentido total de la oración" (1944: 29). Con todo, ello no implica una verdadera contradicción con sus afirmaciones anteriores: "los límites de la unidad melódica coinciden en español con los del grupo fónico. El grupo fónico, de capital importancia para el tratamiento sintáctico de los sonidos, sirve asimismo de base a la forma melódica de la frase" (op.cit.: 31).

La unidad melódica, que sería una unidad entonativa equivalente al "contorno", está constituida por varios "grupos rítmico-semánticos": el grupo rítmico-semántico se define por ser la unidad más pequeña con acento espiratorio (una palabra acentuada, a la que pueden asociarse palabras inacentuadas: "casa", "la casa", "de la casa", etc.) y constituye la unidad rítmica

de la frase (comparable, por tanto, a "las notas de un compás" dentro de la melodía musical). Sin embargo, "los límites de la unidad melódica no van determinados en español por el efecto del acento espiratorio, sino por las circunstancias del sentido y por el orden y armonía del conjunto musical" (op.cit.: 30).

En cuanto al efecto de la pausa, que es el elemento con el que se define el grupo fónico, Navarro Tomás señala que las unidades melódicas "no van siempre marcadas por verdaderas pausas. Con frecuencia el paso de una unidad a otra se manifiesta solamente por la depresión de la intensidad, por el retardamiento de la articulación y por el cambio más o menos brusco de la altura musical, sin que ocurra real y efectiva interrupción de las vibraciones vocálicas" (op.cit.: 31). Si comparamos este punto con los "criterios externos" y los "criterios internos" que expone Cruttenden (1986) para delimitar los "grupos de entonación" (v. 2.2.3.), podremos comprobar la agudeza del análisis de Navarro Tomás y la enorme modernidad de algunos aspectos de su doctrina, inexplicablemente ignorados por sus propios discípulos.

La extensión de la unidad melódica en español, según nuestro autor, oscila entre las siete y las once sílabas, más largas generalmente en la lengua escrita y más cortas, pero también más irregulares, en el diálogo hablado. Señala que sobre un 25% de las unidades melódicas es de 7 / 8 sílabas, razón que explicaría la preponderancia del heptasílabo (seguidillas y canciones) y el octosílabo (romances) en la lírica tradicional española. Sin embargo, señala también que no hay medidas fijas y que sólo pueden anotarse tendencias.

La unidad melódica, con todo, no es la única unidad entonativa que contempla Navarro Tomás: la unidad mayor es la frase, que "consta

regularmente de una parte tensiva y otra distensiva. La primera, prótasis, estimula y reclama la atención; la segunda, apódosis, completa el pensamiento respondiendo al interés suscitado" (op.cit.: 40). Estas dos partes de la frase se llaman "ramas de la frase": "cada una de las dos ramas de la frase puede estar formada por una o varias unidades melódicas" (íbid.).

Así, "el grupo melódico es unidad de comunicación de orden superior a la palabra e inferior a la rama de la frase..." ("palabra" debe entenderse, aquí, como equivalente a "grupo rítmico-semántico") "...en el orden de relación de estas unidades, las palabras desempeñan su función en el grupo melódico que las contiene, los grupos en la rama a que corresponden y las ramas en el conjunto de la frase" (op.cit.: 42): este esquema jerárquico de unidades es lo que él llama "constitución fonológica de la frase".

La constitución fonológica de la frase, por lo demás, implica la solidaridad de los elementos léxicos y gramaticales que constituyen cada una de las unidades jerárquicas entonativas; es decir, Navarro Tomás, como otros autores, también contempla la relación entre entonación y sintaxis, pero desde el punto de vista de la integración de los elementos en la frase: "cada uno de los elementos indicados desarrolla su actividad, sus enlaces y relaciones, entre unidades de su mismo plano. El adjetivo, por ejemplo, afecta al sustantivo a que acompaña en el mismo grupo melódico y no al que le precede o le sigue en otro grupo inmediato. El complemento circunstancial, que suele formar un grupo por sí solo, se refiere al sujeto o predicado que se halla en la misma rama de la frase, no al que figura en la rama opuesta..." (op.cit.: 42).

En este sentido, su aportación consiste en contemplar la relación entre entonación y sintaxis no como fenómenos intercambiables, que cumplan en ocasiones la misma función (por ejemplo, señalar la modalidad de la oración),

y que puedan llegar a ser incluso incompatibles, es decir, no como actores en el mismo plano, sino como fenómenos que actúan en planos distintos: básicamente, la entonación sería el contenedor y el ordenador de las relaciones sintácticas entre los elementos léxicos, sin cuya participación no podrían determinarse tales relaciones. ¡Y esto en 1944!

En resumen, el análisis melódico se circunscribe a la unidad melódica, y el análisis funcional se ciñe a la frase; por su parte, hay tipos de frases que sólo pueden ser descritos convenientemente teniendo en cuenta la unidad intermedia: las ramas de la frase. La palabra o el grupo rítmico-semántico, en cambio, apenas se tiene en cuenta y no se considera unidad entonativa: "dentro de la frase, la palabra no tiene entonación propia" (op.cit.: 22).

Sin duda, la parte más conocida de la doctrina de Navarro Tomás es el análisis de la unidad melódica, que a menudo se presenta como su única aportación en el terreno teórico. Ya hemos visto que, en realidad, en cada uno de los aspectos teóricos que toca hay alguna aportación de cierta relevancia; aunque, en efecto, el análisis de la unidad melódica es uno de los puntos más originales de su enfoque.

En primer lugar, Navarro Tomás divide la unidad melódica en tres partes: la "inflexión inicial", el "cuerpo" y la "inflexión final". Algunos autores (vgr.: Alcina & Blecua, 1975: 458; R. Seco, 1930: 299, en el apéndice de "fonética" agregado por M. Seco en la 3ª ed., de 1958) confunden en sus resúmenes las "partes de la unidades melódica" con las "ramas de la frase" y, así, hablan de "rama inicial", "cuerpo" y "rama final", con lo cual no pueden hablar ya propiamente de "ramas de la frase" y se empobrece y desvirtúa el análisis en su

conjunto; otros, en su caracterización del análisis de configuraciones simplemente desprecian la "inflexión inicial" contemplada por Navarro Tomás, y hablan sólo de dos partes: "cuerpo" y "final" (vgr.: Quilis, 1981: 373).

Esta división de la unidad melódica en tres partes es muy similar a la que se lleva a cabo en la tradición británica. Así, la "inflexión inicial", que corresponde a las primeras sílabas de la unidad melódica hasta el primer acento, es equivalente al *Anacrusis* o *Prehead*; el "cuerpo" melódico se correspondería con el concepto de *Head*; la "inflexión final", por su parte, se correspondería con el *Nucleus* de la tradición británica según lo emplea Ladd (1980), es decir, sin *Tail* (v. figura 25 en el apartado 2.2.1.), y comprendería las últimas sílabas de la unidad melódica, desde el último acento: es decir, una, dos o tres sílabas, según la última palabra sea aguda, llana o esdrújula.

Frente a la inflexión inicial y el cuerpo, la inflexión final es la parte más informativa de la unidad melódica: "la forma y desarrollo del tono en el principio y cuerpo de la unidad (melódica) ofrecen menos valor significativo que idiomático. El valor expresivo de la unidad enunciativa depende sobre todo de su inflexión final" (1944: 51). (Al decir "idiomático", como veremos, Navarro Tomás se refiere a la entonación "dialectal", al "acento" característico de cada dialecto).

Así, puesto que las pocas sílabas que constituyen la inflexión final son la parte más informativa de la unidad melódica, Navarro Tomás se aplica a describir sus posibles movimientos tonales: "el principio y el cuerpo de dicha unidad son elementos cuya estructura varía relativamente poco; el fin, por el contrario, aparece bajo diversas formas, las cuales en español responden esencialmente a cinco tonos, distintos entre sí por su altura musical y por su función expresiva" (íbid.).

Los "cinco tonos" a los que se refiere son: la Cadencia, la Anticadencia, la Semicadencia, la Semianticadencia y la Suspensión. Navarro Tomás los define así:

Cadencia: "Terminación grave, a unos ocho semitonos aproximadamente por debajo de la línea del cuerpo del grupo. Expresa la terminación absoluta de la frase enunciativa. En elocución enfática el descenso alcanza y hasta sobrepasa a veces la octava inferior" (íbid.)

Anticadencia: "Terminación alta a cuatro o cinco semitonos por encima del cuerpo del grupo. Constituye el fin de la rama tensiva, factor de papel esencial en el relieve predicativo y en el contraste y oposición de conceptos, cima máxima en la línea musical de la oración enunciativa" (íbid.).

Semicadencia: "Terminación descendente, menos grave que la de la cadencia, a tres o cuatro semitonos por debajo del cuerpo de la unidad. Expresa el concepto o proposición en serie semántica, la aseveración insegura, la idea insuficientemente definida" (íbid.).

Semianticadencia: "Terminación menos alta que la de anticadencia, dos o tres semitonos sobre el nivel medio de la unidad. Corresponde a unidades interiores de sentido continuativo y señala oposiciones y contrastes de carácter secundario" (íbid.).

Suspensión: "Terminación en el mismo nivel que el cuerpo del grupo. Aunque la uniformidad tónica no sea matemáticamente exacta, su impresión es la de una

interrupción de la voz sin elevación ni descenso apreciables. Expresa el sentido inclompleto, el corte de la idea pendiente de continuación" (íbid.).

Nótese que, frente a la mayoría de sus autores contemporáneos, Navarro Tomás no sólo no es reacio a cuantificar las unidades que describe sino que ofrece unos valores relativos determinados con una gran exactitud. Tales valores muy precisos, sin embargo, rara vez han sido tomados en consideración para el trabajo experimental: simplemente, se han ignorado.

También es de notar que Navarro Tomás se aplica a describir el significante formal de cada tono y la función "expresiva" que desempeña: lejos, pues, de la afirmación de que "su metodología carecía lógicamente de los principios funcionales de la lingüística de hoy" (Quilis, 1981: 414); por el contrario, su descripción pretende ser una descripción de los signos lingüísticos que constituyen la entonación española, con una forma y una función definidas: "como elementos típicos en que concurren caracteres representativos y funcionales, estos tonos que cierran la línea de las unidades enunciativas pueden recibir propiamente el nombre de *tonemas*. El tonema final sella el carácter fonológico de la unidad en que se halla" (1944: 51). Es decir, Navarro Tomás es plenamente consciente de que las unidades melódicas son unidades fonológicas con una función propia: el "valor expresivo" que atribuye en su definición a los "tonemas".

Este carácter manifiestamente fonológico de su análisis de la entonación ha sido pasado muchas veces por alto, pero explica en buena medida el candor con que algunos autores españoles posteriores han buscado en el análisis de niveles norteamericano una formalización más precisa y objetiva que la ofrece el análisis de configuraciones. Por lo demás, las ideas de Navarro Tomás

constituyen una notable excepción en el campo teórico del análisis de configuraciones, pues difícilmente un autor europeo ha postulado con tanta claridad dicho carácter fonológico de la entonación. En este sentido, se encuentra mucho más cerca de la sensibilidad norteamericana que de la británica.

Sin embargo, no es posible identificar los cinco tonemas de Navarro Tomás con cinco niveles tonales, al estilo norteamericano (como apunta Hoyos Andrade, 1970: 21): no se trata de niveles con los que definir toda la unidad melódica, sino movimientos finales con los que caracterizarla fonológicamente.

Es decir, algo muy parecido, más bien, a los *nuclear tones* de la tradición británica, y así lo concebía el propio Navarro Tomás: "refiriéndose en conjunto a la entonación inglesa, sin separar la elocución enunciativa de la interrogativa o emocional, señaló Palmer cuatro tipos de tonemas, llamados por él *nucleus tones*. Tres de estos tonemas *-falling, high risings, low risings-* tienen cierta semejanza con la cadencia, anticadencia y semianticadencia del español" (op.cit.: 52).

Es error generalizado considerar los cinco tonemas indicados como los únicos definidos por Navarro Tomás en su trabajo, y creer que sirven para la descripción de cualquier unidad melódica. Por el contrario, *cadencia, anticadencia, semicadencia, semianticadencia* y *suspensión* son tonemas con los que se caracteriza sólo la "entonación enunciativa": "sobre la base de la forma y valor de su tonema final, la unidad melódica en la entonación enunciativa española presenta cinco tipos diferentes. Como unidad fonológica con una determinada función expresiva, el conjunto de tonos e inflexiones reunidos en la línea musical del grupo de entonación podría ser llamado *sintonema*. En

inglés se ha usado *tune* para designar esta misma unidad melódica. Se puede extender a cada unidad melódica o sintonema la denominación del tonema final que le caracteriza" (op.cit.: 52).

Así, la entonación enunciativa está formada en español, según NavarroTomás, por cinco tipos de unidades melódicas o "sintonemas": Unidad de cadencia (C), Unidad de anticadencia (A), Unidad de semicadencia (c), Unidad de semianticadencia (a) y Unidad de suspensión (s). En la figura 35 se puede ver un ejemplo de entonación enunciativa con los cinco sintonemas característicos, y su representación gráfica: la frase "En el pardo lomo del cerro (c), entre los viejos olivares (A), la casa solariega (s), con su amplia portalada (a), ofrecía franco reposo (C)" (op.cit.: 209).

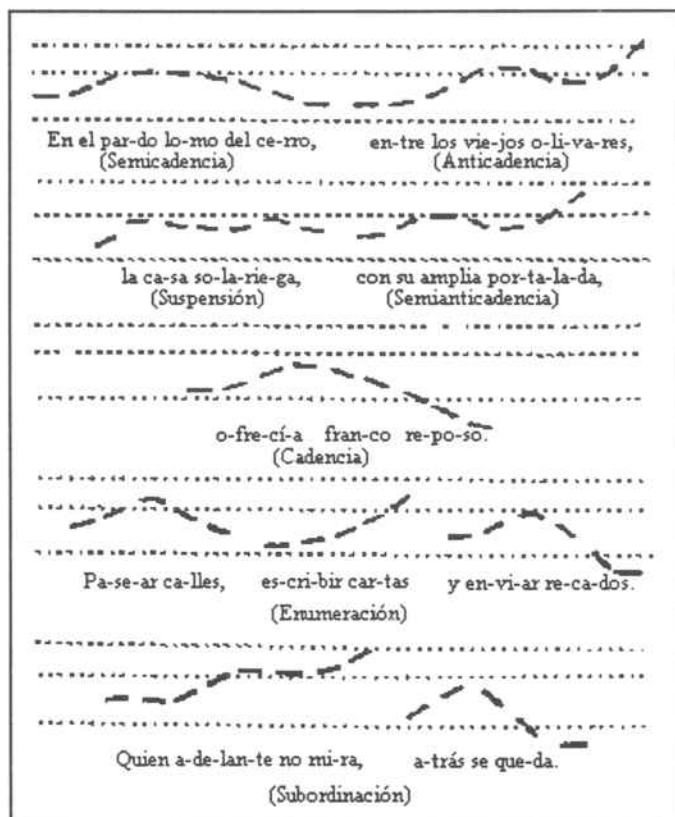


Figura 35 - Ejemplos de entonación enunciativa.

Pero Navarro Tomás también describe las unidades melódicas propias de la "entonación interrogativa":

"El grupo melódico interrogativo presenta en español numerosas formas. No consiste simplemente como suele creerse en una línea de tono medio terminada con elevación de la voz. Tampoco consiste en una variedad de

inflexiones sin molde ni medida, inspiradas por las circunstancias del momento. La multiplicidad de sus manifestaciones se reduce en líneas generales a los tipos siguientes:

Unidad de interrogación absoluta: Tono descendente en el cuerpo del grupo e inflexión final ascendente. Se emplea como forma simple en las preguntas de una sola unidad y como grupo final en las de varias unidades.

Unidad de interrogación relativa: Tono perceptiblemente alto y sostenido en el cuerpo de la unidad y terminación circunfleja. Se emplea como unidad independiente o como grupo final en preguntas en que se presume la contestación.

Unidad de interrogación aseverativa: Inflexión descendente en el cuerpo del grupo y nuevo descenso aún más marcado en la terminación. Se usa como unidad independiente o final en los casos en que el sentido de la pregunta se inclina a la afirmación y de manera más general cuando las frases empiezan con pronombre o adverbio interrogativo.

Unidad de interrogación intensificativa: Movimiento ascendente en el cuerpo del grupo y terminación aguda. Corresponde a las preguntas de sentido reiterativo y en especial a las que dentro de ese mismo carácter envuelven cierto

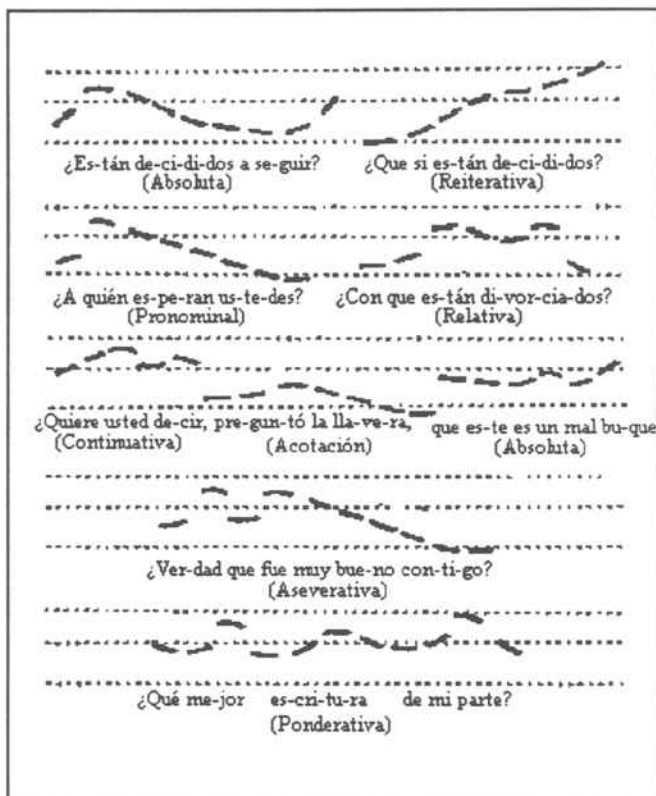


Figura 36 - Ejemplos de entonación interrogativa.

valor de exclamación.

Unidad de interrogación continuativa: Tono descendente en el cuerpo del grupo, como en la unidad de pregunta absoluta, e inflexión final circunfleja más reducida y más baja que la de la pregunta relativa. Sirve en general para las unidades no finales en preguntas de varios grupos." (op.cit.: 100-101).

En la figura 36 puede verse la representación gráfica de cada una de estas unidades melódicas propias de la entonación interrogativa.

El sentido de esta larga cita es poder comprobar cómo los cinco tonemas que suelen darse como universales en las reseñas del trabajo de Navarro Tomás ni siquiera aparecen en la descripción de las unidades melódicas interrogativas, pues corresponden exclusivamente a la entonación enunciativa. Sin embargo, el autor reconoce que pueden establecerse ciertas correspondencias entre los sintonemas de ambas modalidades entonativas:

"Hay una visible relación entre los tipos esenciales de las unidades interrogativas y enunciativas. A la unidad de anticadencia (A) corresponde la de interrogación absoluta (IA); a la de cadencia (C), la de interrogación aseverativa (IC). Análoga relación se observa entre la unidad de suspensión (s) y la de continuación interrogativa (Is). La unidad de interrogación relativa (Ir) ofrece semejanza con las unidades intermedias de semicadencia y semianticadencia. La unidad de interrogación intensificativa o exclamativa, sin correspondencia aparente, puede representarse por las iniciales (Ie)" (íbid.).

En realidad, Navarro Tomás divide la entonación en cuatro tipos generales: entonación lógica, entonación volitiva, entonación emocional y

entonación idiomática. Esta última, la "entonación idiomática", se refiere a "la cadencia habitual del habla de cada país, el dejo característico de cada dialecto y hasta el sello ... de cada individuo" (op.cit.: 10). La "entonación volitiva" y la "entonación emocional", por su parte, son consideradas por el autor como modalidades generales sin una forma realmente determinada y sin una función constante: hoy diríamos que corresponden al plano expresivo de la lengua; reciben un tratamiento pedagógico, pues se trata de un manual práctico de entonación, pero no se describe en ningún momento la forma determinada de sus unidades melódicas.

Sólo la llamada "entonación lógica" recibe un tratamiento formal serio y bien definido, puesto que "el conjunto de las unidades enunciativas e interrogativas constituye el sistema melódico de la entonación lógica" (op.cit.: 101).

Es decir, y como resumen, por "entonación lógica" debe entenderse "entonación fonológica", según hemos visto, que Navarro Tomás reduce a entonación enunciativa o interrogativa, cada una de las cuales consta de una serie de unidades melódicas o sintonemas, de carácter fonológico, que se combinan en la ramas de la frase, etc.

Cada unidad melódica se define fundamentalmente (aunque no exclusivamente) por la forma de la inflexión final. Las unidades enunciativas pueden ser: de cadencia (C), de anticadencia (A), de semicadencia (c), de semianticadencia (a) y de suspensión (s). Las unidades interrogativas, por su parte, pueden ser: interrogativas absolutas (IA), interrogativas aseverativas (IC), interrogativas continuativas (Is), interrogativas relativas (Ir) e interrogativas intensificativas o exclamativas (Ie).

Finalmente, entre las unidades melódicas enunciativas e interrogativas

pueden establecerse ciertas relaciones, expresadas en el siguiente esquema:

<u>Unidades interrogativas</u>			<u>Unidades enunciativas</u>	
Interrog. absoluta	<i>IA</i>	→	Unidad de anticadencia	<i>A</i>
Interrog. aseverativa	<i>IC</i>	→	Unidad de cadencia	<i>C</i>
Interrog. continuativa	<i>Is</i>	→	Unidad de suspensión	<i>s</i>
Interrog. relativa	<i>Ir</i>	→	Unidad de semianticadencia	<i>a</i>
			Unidad de semicadencia	<i>c</i>
Interrog. intensificativa	<i>Ie</i>	→	- - -	

Por tanto, es falso que Navarro Tomás describa la entonación española basándose sólo en cinco tonemas: decir eso equivale a detenerse en el segundo capítulo del *Manual de entonación española* y a deformar y empobrecer la doctrina que en él se expone, de una manera radical e injusta.

Es más, el autor habla incluso de otros dos tonemas más en español, no propios ya de la entonación lógica, sino de las otras entonaciones: el tonema volitivo o desiderativo ("la entonación desiderativa, en sus diferentes grados de petición, ruego o súplica, se sirve de un tonema característico, de suave y modulado giro musical" -op.cit.:185-) y el tonema emocional ("la entonación afectiva utiliza como signo propio una inflexión circunfleja, de variable altura y amplitud según el grado y calidad de la emoción" -íbid.-).

Para concluir nuestro breve recorrido, en fin, conviene hacer una referencia a la tipología de frases ofrecida por Navarro Tomás en su libro, que, como dijimos, ha sido seguida, en mayor o menor medida, por la mayoría de los autores que han tratado la entonación española.

En el ámbito de la entonación enunciativa, se distinguen en el manual tres

apartados bien definidos: las aseveraciones (ordinaria, categórica, dubitativa e insinuativa), las enumeraciones (completa, interior, acumulativa, intensificativa, calificativa, descriptiva, etc.) y otros fenómenos entonativos de la enunciación (características entonativas del complemento circunstancial, de la aposición predicativa, de las locuciones, de los paréntesis, etc.).

En el ámbito de la entonación interrogativa, el manual distingue entre preguntas absolutas, relativas, aseverativas, pronominales, reiterativas, restrictivas, hipotéticas, alternativas, etc.

En el ámbito de la entonación volitiva, se describe la entonación de mandato, de invitación, de exhortación, de recomendación, de ruego, de súplica, etc.; pero también la entonación de cortesía, la entonación de los proverbios, etc.

Finalmente, en el ámbito de la entonación emocional se distingue la entonación afectiva descendente, ascendente, ondulada, aprobatoria, desaprobatoria, etc.

Puede observarse, por tanto, cómo hay una notable confusión de criterios en cuanto a la tipología de las frases: criterios semánticos, gramaticales, pragmáticos, discursivos, actitudinales... Tal vez esta misma confusión, al menos aparente, haya ahuyentado a los estudiosos de la doctrina que encubren; sin embargo, no deja de ser curioso que sea este punto, precisamente, el que haya encontrado mayor predicamento entre los demás autores.

En cualquier caso, no hay que olvidar que el *Manual de entonación* es justamente eso: un manual. Es decir, la tipología de las frases estudiadas y los comentarios de su autor sobre tal o cual uso estaban indicados para la enseñanza de la entonación española culta y literaria, y especialmente para la enseñanza a los extranjeros (el libro fue editado originalmente en Nueva York). El objetivo didáctico de la obra, por tanto, si bien no impedía la precisión del análisis o la

originalidad de los conceptos empleados, en cierto modo sí que obligaba a una mayor libertad tipológica en cuanto a las frases cuyo estudio se proponía.

El estudio de la entonación ha de fijarse únicamente, pues, en la clasificación de "sintonemas" enunciativos e interrogativos: las demás consideraciones entonativas, sin duda interesantes y jugosas desde una perspectiva didáctica, escapan ya del ámbito estrictamente lingüístico y fonológico de la entonación. Y así lo entendía, según se desprende de cuanto hemos podido ver, el propio Navarro Tomás.

2.5.2. Análisis por niveles de la entonación española

En los años 50 y 60, algunos lingüistas americanos se aplicaron a analizar la entonación del español americano según el sistema de niveles, acentos y junturas propio de su tradición. Se trata, sin duda, de un fenómeno atípico y curioso, pues el español es la única lengua, además del inglés, que cuenta con estudios sistemáticos de su entonación siguiendo la tradición norteamericana de análisis por niveles.

Tales estudios partieron, generalmente, del análisis de corpus propios, grabados a hispanoablantes latinoamericanos: la obra de Navarro Tomás, por tanto, apenas les servía de referencia, pues ni los datos (español peninsular y literario) ni el método empleado (análisis de configuraciones) era homologable a los suyos propios.

En realidad, el alcance de la lingüística española rara vez ha llegado con fuerza hasta la otra parte del Atlántico: en Latinoamérica, la lingüística y, en concreto, la fonología, normalmente han seguido los criterios y las directrices de la lingüística y de la fonología norteamericanas, que inmediatamente se han

considerado como propias. Ello es así no sólo por la relativa pobreza conceptual de la lingüística española (que se manifiesta en el escaso número de autores traducidos a otras lenguas), o por el eurocentrismo hispánico de los estudios lingüísticos realizados en España (hasta hace poco tiempo, "americanismo" era sinónimo *de facto* de "vulgarismo"), sino sobre todo por la enorme riqueza de ideas, autores y métodos originales de la lingüística norteamericana: de ningún modo el reducido y casi provinciano ámbito de la lingüística española podría haberle hecho sombra ni ejercer mayor influencia sobre el continente americano de habla hispana.

Así, el fenómeno del análisis de niveles aplicado al español por autores de los dos continentes americanos no es lo chocante, sino el predicamento que tuvieron tales trabajos en el ámbito estricto de la fonología de nuestro país. La influencia, pues, no fue de España a América sino todo lo contrario: de América a España.

Los principales análisis por niveles de la entonación española fueron los realizados por J.D. Bowen (1956), R.P. Stockwell, J.D. Bowen & I. Silva-Fuenzalida (1956), J.D. Bowen & R.P. Stockwell (1960), I. Silva-Fuenzalida (1957), D. Cárdenas (1960) y J. H. Matluck (1965). En concreto, *Patterns of Spanish pronunciation* (1960), de Bowen & Stockwell, y "Entonación hispánica" (1965), de J. H. Matluck, fueron los trabajos más influyentes, hasta el punto de eclipsar temporalmente la obra de Navarro Tomás, como hemos dicho, y de condicionar el desarrollo posterior de la "escuela española" de análisis de la entonación.

En general, estos trabajos no aportan ninguna novedad teórica al propio análisis de niveles y se limitan a adaptarlo a la lengua española. La novedad

estriba, desde luego, en hablar, para el español, de "juntura terminal", de "morfemas entonativos", o de "fonemas acentuales"; pero en el conjunto de la tradición norteamericana son irrelevantes.

Otra característica importante de estos trabajos es la continua referencia que se hace a la lengua inglesa para describir la entonación española, que sin duda se debe a un cierto condicionamiento lingüístico de los autores norteamericanos: condicionamiento de su propia competencia en español, seguramente (pues en ningún caso se siguió un proceso de contrastación perceptiva experimental de los modelos propuestos), pero sobre todo del peso de la tradición lingüística norteamericana, centrada exclusivamente en el inglés (frente a las tradiciones europeas, mucho más abiertas a la incorporación de datos y al análisis de otras lenguas). Sin embargo, y curiosamente, en toda la bibliografía española sobre el tema sólo Hoyos Andrade alude a este hecho: "el defecto principal (de tales trabajos) consistió en haberse fiado en exceso del oído y de los análisis similares hechos para el idioma inglés ... el oído del mejor de los fonetistas está condicionado de una u otra forma por su propia lengua" (1970: 18).

Normalmente, los distintos autores consideran que la entonación del español puede describirse con sólo tres niveles tonales: alto (3), medio (2) y bajo (1), entre los que el nivel alto (3) se considera poco usual. Tales niveles son pertinentes, a lo largo del grupo de entonación, al menos en tres posiciones: primera y última sílaba de la frase, y última sílaba acentuada.

Matluck (1965) describe tres juntas terminales: ascendente (\uparrow), descendente (\downarrow) y suspensiva (\rightarrow); Bowen & Stockwell (1960), en cambio, describen dos juntas terminales: ascendente (\uparrow , que incluye la suspensión

final) y descendente (\downarrow); y una juntura interna (|), que marca el límite de los grupos no finales y que se considera una suspensión.

Finalmente, se describen dos fonemas acentuales: el acento fuerte (´) y el acento débil (˘); aunque, en ocasiones, también se habla de un acento "secundario" o "de insistencia".

Los modelos propuestos por Bowen & Stockwell (1960) son los siguientes (según Hoyos Andrade, 1970: 16-17):

/1211 \downarrow / : Enunciado normal, órdenes, preguntas con partícula interrogativa.

/1231 \downarrow / : Enunciado enfático, pregunta pronominal enfática.

/1222 \uparrow / : Pregunta absoluta.

/1231 \downarrow / : Pregunta relativa.

/2321 \downarrow / : Pregunta con partícula interrogativa y matiz de interés.

/2321 \uparrow / : Pregunta muy cortés.

/1211 | 122 \uparrow / : Pregunta que pide confirmación.

/1122 \uparrow 2211 \downarrow / : Pregunta selectiva.

/1222 | 1121 \downarrow / y /1222 | 2211 \downarrow / : Afirmación deliberada.

/1222 | 1122 | 2211 \downarrow / : Oración compuesta.

El trabajo de J. H. Matluck (1965), sin embargo, encontró mucho más eco entre los autores españoles, y fue tomado como modelo, por ejemplo, en trabajos de investigación como el de Hoyos Andrade (1970) y en tratados como la *Gramática española* (1975) de J. Alcina y J. M. Blecua.

Matluck describe los quince siguientes patrones o "morfemas" entonativos (según Hoyos Andrade, 1970: 33 y ss.):

- I / (12)11 ↓ / : Afirmación simple ("*está contento*").
Pregunta pronominal simple ("*¿por qué no me lo dices?*").
- II / (12)22 ↑ / : Pregunta absoluta ("*¿está contento?*")
Pregunta pronominal con matiz de impaciencia, petulancia o fastidio ("*¿por qué no me lo dices?*").
- III / (12)31 ↓ / : Afirmación enfática o contrastante ("*¡está contento!*")
Pregunta pronominal enfática ("*¡¿por qué no me lo dices?!*").
Pregunta reiterativa o pregunta eco ("*¿que se está contento?*").
- IV / (12)31 → / : Pregunta relativa ("*¿no vienen en barco?*" "...no").
- V / (12)11→(12)22↑ / : Pregunta confirmativa ("*vienen mañana ¿verdad?*").
- VI / (23)21 ↓ / : Afirmación llamativa ("*¡pero fíjate cómo baila!*").
Pregunta pronominal, con matiz de sorpresa, inoportunidad o interés especial ("*¿por qué molestas?*" -inoportunidad-).
- VII / (23)11 ↓ / : Pregunta pronominal con matiz de molestia o de fastidio ("*¿por qué molestas?*" -fastidio-)
- VIII / (23)11 ↑ / : Pregunta con matiz de condescendencia ("*¿por qué no vienes?*").
- IX / (12)33→1211↓ / : Grupo suspensivo mayor -fin de la prótasis de un enunciado- ("*ya son las ocho*" "...y no llegan").

- X / (12)33 ↑ / : Pregunta reiterativa intensificada ("*¿que se están decididos?*" "...¡claro que sí!").
- XI / (12)22 → / : Grupo suspensivo menor ("*llegamos a su cuarto*" "...y para poder entrar tuvimos que llamar al mozo").
- XII / (11)22↑(21)11↓/: Pregunta alternativa o de selección ("*se va mañana o pasado mañana?*").
- XIII / (12)22→1211↓/: Afirmación deliberada ("*es estudiante, de primer curso*").
- XIV / → (11)11 ↓ / : Aposición final ("*lo siento...*" "*señor Rodríguez*").
- XV / (12)21 ↓ / : Aposición final con otros matices expresivos -frialdad, indiferencia...- ("*muy bien, gracias*").
- Saludos y fórmulas de cortesía ("*buenos días*").

El trabajo citado de Hoyos Andrade (1970 - resumen de la tesis doctoral de su autor, dirigida por A. Quilis) consistió en un análisis espectrográfico de las frases propuestas por Matluck como ejemplo de sus patrones entonativos, para confirmar, de modo experimental, la validez de su descripción; sin embargo, Hoyos Andrade se encontró con numerosos problemas de identificación de los niveles tonales y lo que comenzó como una muestra más del entusiasmo que despertaba el análisis de niveles en los autores españoles acabó como una vuelta al análisis de configuraciones.

El enfoque previo del trabajo de Hoyos Andrade sin duda pecaba de cierta ingenuidad, pues no en vano ningún autor norteamericano se había atrevido a cuantificar el valor de unos niveles tonales que proponían sólo como unidades relativas y abstractas; dice el autor español: "se me ocurrió que sería útil e

interesante encontrar la confirmación científica que faltaba en los análisis de los hispanistas norteamericanos" (op.cit.: 5). Naturalmente, no sólo no encontró confirmación científica alguna, sino que concluye:

"Creo haber demostrado, por lo menos, aunque esto parezca un resultado negativo, que el sistema de niveles de tono y junturas que se ha querido presentar como descripción objetiva de la entonación hispánica carece de valor científico por no fundarse en la realidad fonética de la lengua, ni por lo tanto en ningún análisis instrumental previo.

»Los niveles de tono numerados, las flechas indicadoras de la dirección de la voz en las terminaciones podrán ser quizás útiles en algún caso determinado como ayudas pedagógicas para enseñar la lengua, pero no pueden tomarse como descripción lingüísticamente objetiva de la entonación castellana.

»Creo que la descripción científica de la entonación debe orientarse definitivamente por el sistema de las configuraciones tal como lo proclama Bolinger y tal como lo han venido haciendo los fonetistas españoles guiados por las sabias enseñanzas de Don Tomás Navarro Tomás" (op.cit.: 76).

Estos párrafos finales constituyen todo un manifiesto a favor de la escuela española tradicional, y muestran cómo la irrupción del análisis de niveles fue algo parecido a una moda científica cuyos presupuestos teóricos y cuyas implicaciones fonológicas nunca acabaron de asumirse por completo.

Un buen ejemplo de ello es el curioso método propuesto en el *Esbozo de una nueva gramática...* (1973) de la Real Academia Española (redactado, como hemos dicho, por S. Gili Gaya y S. Fernández Ramírez), que intenta conjugar el moderno análisis de niveles con la doctrina tradicional de Navarro Tomás: "el número 2 corresponde al llamado tono *semigrave* por Navarro Tomás -*Manual de entonación española*-, separado aproximadamente por cuatro semitonos del

tono normal. El número 1 al tono grave, unos cuatro semitonos por debajo del 2. El número 4 corresponde al llamado por el mismo autor tono agudo, superior musicalmente en cuatro semitonos al 3, que es el normal. El tono agudo puede elevarse todavía alrededor de ocho semitonos por encima del tono normal. No empleamos en este caso un número especial, que sería el 5, por la razón de que la nota, representada por 4, pueda incrementarse tonalmente en casi todos los casos, según el grado de intensidad de la voz, de la tensión de los órganos articulatorios y según la naturaleza del momento expresivo concomitante" (R.A.E., 1973: 105-106, en la nota 5 a pie de página).

Es decir, la original propuesta de los académicos consiste en describir la entonación española según un sistema de niveles tonales cuantificados con la mayor precisión y que además han sido extraídos de la tradicional doctrina de Navarro Tomás: la cuadratura del círculo.

En realidad, los cinco "niveles tonales" a los que se refiere el *Esbozo...* son los cinco tonemas con los que Navarro Tomás caracteriza la inflexión final de las unidades melódicas enunciativas: el tono grave (1) es la cadencia (C), el semigrave (2) la semicadencia (c), etc. Por tanto, los autores ignoran la descripción que hace Navarro Tomás de la entonación interrogativa (que, como vimos en el apartado anterior, está constituida por otros cinco sintonemas diferentes) y emplean los tonemas enunciativos para describir incluso las frases interrogativas. La confusión es de peso y se enmarca en la mala lectura que se ha hecho tradicionalmente de la obra de Navarro Tomás.

También Hoyos Andrade incurrió en su trabajo en el mismo error (y no es descartable que fueran precisamente sus palabras las que influyeran en la posterior interpretación de los académicos): "los cinco tonemas de Navarro Tomás no suelen ser generalmente admitidos por los lingüistas americanos como

otras tantas unidades funcionales. Para ello habría que admitir cinco niveles de tono fonológicamente distintos y esto contradice a la ya casi tradicional teoría de señalar sólo tres niveles de tono funcionalmente oponibles en español. Pero si uno se atiene a las enseñanzas de Navarro Tomás, habría que concluir que son cinco y no tres los niveles pertinentes, ya que cada tonema (y cada sintonema) tiene una función expresiva diferente" (1970: 21). Es decir, Hoyos Andrade no se da cuenta de que los tonemas enunciativos sólo caracterizan a las unidades melódicas enunciativas; ignora, asimismo, los sintonemas interrogativos; desprecia el carácter estrictamente final de los tonemas, que sólo sirven para caracterizar la inflexión final, y supone que podrían caracterizar igualmente toda la melodía (con lo cual desvirtúa por completo el sentido del análisis); pero, sobre todo, cada tonema, que engloba un movimiento o inflexión tonal complejo (comparable, como vimos, a los *nuclear tones* británicos), aquí es identificado como un nivel tonal simple, como un "nota" musical sencilla o una altura relativa y continua.

Así, la cuantificación que hace la Real Academia de los "niveles tonales" basados en Navarro Tomás constituye un error doble: por una parte, los tonemas, que Navarro Tomás considera inflexiones complejas, son considerados en el *Esbozo...* niveles simples; por otra, la altura precisada de tales niveles se refiere, justamente, a la altura del movimiento interno de tales inflexiones complejas o tonemas (no es la altura del tonema con respecto a otro punto, sino la altura de un punto con respecto al anterior, y ambos componen el tonema).

Hemos visto que Hoyos Andrade, finalmente, se decantó en las conclusiones de su tesis por el análisis de configuraciones. Lejos de confirmar los patrones por niveles de Matluck, incluso propone dos tipos fundamentales

de configuraciones en español, según parece desprenderse de sus datos: la "configuración descendente" y la "configuración descendente-ascendente", que coinciden con la entonación enunciativa y la interrogativa, según el esquema general de Navarro Tomás.

Sin embargo, el propio director del trabajo, A. Quilis, no debió de quedar muy convencido de tales conclusiones, pues unos años después comienza a significarse como uno de los abanderados de la aplicación del análisis de niveles al español (v., por ejemplo, Quilis, 1975: "Las unidades de entonación").

En 1981, Quilis publica su *Fonética acústica de la lengua española*, cuyo capítulo referido a la entonación es un magnífico resumen de cuanto habían publicado él y sus colaboradores sobre el tema (aunque ni siquiera cita, curiosamente, el trabajo de Hoyos Andrade). En él encontramos la primera descripción sistemática de la entonación española empleando el análisis de niveles y llevada a cabo por un autor español.

Quilis habla, así, de tres niveles tonales (bajo /1/, medio /2/ y alto /3/), de dos fonemas acentuales (fuerte /' / y débil /˘ /) y de tres junturas terminales (descendente /↓/, ascendente /↑/ y suspensiva / | /), aunque finalmente las reduce a dos (descendente /↓/ y ascendente /↑/ -que incluye la suspensión, como en Stockwell & Bowen, 1960-), y dice que todos ellos son "fonemas suprasegmentales". También habla de "morfemas entonativos" y, en general, sigue el método de análisis de niveles coherentemente, considerando que la entonación es un fenómeno fonológico, etc.

Sin embargo, el autor considera que el análisis de niveles no es incompatible con el análisis de configuraciones (como vimos en el apartado 2.1.1.), aduciendo que "a) mirando una curva de entonación, se distinguen

claramente una sucesión de ascensos, descensos, deslizamientos a todo lo largo del enunciado; pero para un oído normal, la melodía del lenguaje no reside en estas sucesivas variaciones frecuenciales, sino en una consecución de niveles tonales cada unos de los cuales es más alto, más bajo, o está a la misma altura que el precedente; b) el análisis configurativo incluye también los niveles al describir sus materiales: alto ascendente y descendente, bajo ascendente y descendente; y también incluye junturas, puesto que los tonos finales ascendentes, descendentes y suspensivos no pueden manejarse si no se perfilan unas unidades más pequeñas" (1981: 366-367). Es decir, Quilis aboga decididamente por el análisis de niveles, pues incluso su percepción del análisis de configuraciones está claramente mediatizada por aquel: no entiende que en el análisis de configuraciones haya un solo núcleo relevante; no concibe que pueda manejarse la inflexión (por ejemplo, el *nuclear tone*) como una unidad tonal "si no se perfilan unas unidades más pequeñas"; sólo concibe la percepción de la entonación como "percepción de niveles"; identifica la "inflexión final" (o el *nucleus*) como una "juntura"; etc.

Con todo, en su descripción de la entonación española ofrece igualmente la curva entonativa y la notación por niveles y junturas, todo a la vez: "partiremos del análisis de configuraciones para establecer después los niveles pertinentes de la frase que se estudia" (op.cit.: 415). Esto, unido a la tipología de frases que emplea, en la que sigue de cerca (como todos los autores) a Navarro Tomás, parece indicar que se trata de una auténtica perspectiva integradora de ambos métodos; así es, al menos, como lo pretende el autor. Sin embargo, esta "integración" entre ambos métodos, en la que se emplean y se contemplan las unidades y las características de uno y se ignoran las unidades y las características del otro, resulta a la postre una descripción por niveles en

toda lid, aunque ilustrada con el dibujo de las curvas entonativas. Así, en Quilis (1985), el autor prescinde ya de todo lastre configurativo y propone un análisis inequívocamente de niveles.

Finalmente, y a pesar de que en su *Tratado de fonología y fonética españolas* (1993) apenas aporta información nueva y se limita a repetir y a resumir el capítulo dedicado a la entonación en Quilis (1981), encontramos alguna modificación importante: en primer lugar, los niveles con los que describe las mismas frases varían notablemente en las dos obras, aunque las curvas que ofrece son exactamente las mismas; en segundo lugar, se añade un apartado de entonaciones dialectales ("funciones de la entonación en el nivel sociolingüístico") que no se describen, sin embargo, mediante niveles tonales, y de las que sólo se ofrece el dibujo de la curva y una descripción intuitiva y aproximativa, al estilo del análisis de configuraciones.

De cualquier modo, la aportación principal de Quilis al estudio de la entonación es su planteamiento general sobre las distintas funciones que pueden desempeñar los fenómenos entonativos (1981: 376 y ss.; 1981b). Así, distingue tres niveles en los que actúa la entonación: el nivel lingüístico, el nivel expresivo y el nivel sociolingüístico.

Dentro del nivel lingüístico, señala que la entonación desempeña la llamada "función integradora": "para algunos entonólogos la entonación tiene como función primordial y única la de integrar las palabras para formar una oración" (1981: 384). También desempeña la llamada "función delimitadora" o "demarcativa": "la entonación delimita los enunciados y segmenta el continuum de discurso ... por razones fisiológicas ... por razones de comprensión del mensaje ... o por otros motivos lingüísticos" (op.cit.: 388). Finalmente, dentro

del nivel lingüístico la entonación cumple la "función distintiva": "la función distintiva de la entonación reside en los movimientos descendente o ascendente de la frecuencia fundamental al final de un enunciado ... un enunciado afirmativo terminaría con una frecuencia fundamental descendente, mientras que un enunciado interrogativo lo haría con un frecuencia fundamental ascendente" (op.cit.: 381).

En el nivel expresivo, la entonación sería el transmisor por excelencia de los contenidos afectivos o emocionales en la comunicación, lo cual plantea al estudioso dos cuestiones de mucho interés: por un lado, si se trata de un fenómeno lingüístico o motivado; por otro, si se trata de un fenómeno cultural o universal.

Finalmente, y dentro del nivel sociolingüístico, "la entonación comunica dos tipos de información: a) la información relacionada intrínsecamente con el individuo, es decir, la que comunica sus características personales, como la edad, el sexo, el temperamento, el carácter, etc.; b) la información propiamente sociolingüística, es decir, la que comunica las características del grupo al que pertenece el individuo, como el origen geográfico, el medio social, el grado de cultura, etc." (op.cit.: 389).

Para el español, la lista de entonaciones que ofrece Quilis es la siguiente (según 1993, que ofrece patrones más abstractos que en 1981, pues sólo se fija en el final del contorno -siguiendo, de nuevo, a Bowen & Stockwell, 1960-):

Nivel lingüístico - Función distintiva:

Enunciado declarativo: / 2 1 ↓ / "Aquí viene carmen"

Enunciado interrogativo absoluto: / 1 2 ↑ / "¿Había mucha gente?"

Enunciado interrogativo pronominal: / 2 1 ↓ / "¿Por qué no me lo dices?"

("un patrón semejante al del enunciado declarativo; la diferencia entre éste y el enunciado interrogativo pronominal viene dada por la presencia del elemento gramatical interrogativo" (1993: 431).

Nivel lingüístico - Función demarcativa no distintiva:

Pregunta disyuntiva: / 1 2 ↑ 2 1 ↓ / "¿Usa azul o negro?"

Pregunta copulativa: / 1 2 ↑ 2 1 ↓ / "¿Usa azul y negro?"

En este apartado hay también referencias y descripciones aproximativas, pero sin ningún patrón tonal por niveles, de los siguientes tipos de frases:

Enumeración completa ("El caballo es fuerte, rápido y negro"),

enumeración incompleta ("El caballo es fuerte, rápido, negro"),

complemento hiperbolizado ("Sobre el caballo, con el sombrero en la mano, cabalga el jinete"),

coordinación de primer grado ("Ni come ni bebe"),

coordinación de segundo grado ("Viene todas las tardes, a pesar del frío"),

enunciado parentético ("Cuando venía, me contaba Juan, amaneció")

y subordinación.

Nivel expresivo:

Afirmación enfática:	Aparece el nivel /3/ y el final es descendente. "¡No quiero nada!": /3 2 1 2 1 ↓/
Pregunta pronominal enfática:	Íd. "¿Por qué no me lo dices?" /1 3 2 1 2 1 ↓/
Pregunta pronominal con matiz de cortesía:	Hay un elemento gramatical interrogativo y el final es ascendente. "¿Por qué no me lo dices?" /1 2 2 1 1 2 ↑/
Pregunta de cortesía:	Fuerte descenso desde la última sílaba tónica, y luego ascenso. "¿Ha terminado?" /2 1 2 1 2 ↑/
Pregunta reiterativa:	Elemento gramatical interrogativo, aparición del nivel /3/ y final ascendente. "¿Que por qué no me lo dices?" /1 3 2 1 2 1 ↓/
Pregunta relativa:	Nivel tonal /3/, final descendente y "se utiliza para cerciorarse de algo de lo que sólo se tiene una idea" (op.cit.: 451). "¿Estudia mucho?" /1 2 1 3 1 ↓/
Pregunta confirmativa:	"se caracteriza por la presencia en el final del enunciado de una palabra prácticamente lexicalizada (<i>sic</i>) -no, verdad- con fundamebntal ascendente" (op.cit.: 451-452). "Es bonita ¿verdad?" /2 1 2 1 ↓ 1 2 ↑/

Pregunta imperativa:	"caracterizada por mantener un fundamental muy alto" (op.cit.: 453).
	"¿Quieres venir?" /2 2 3 ↑/
Exclamación:	"caracterizada por un descenso muy acusado del fundamental desde la primera sílaba acentuada" (íbid.).
	"¡Qué vas a hacer!" /3 2 1 2 1 ↓/

Como puede comprobarse, la clasificación fraseológica de Quilis es una nueva ordenación de los tipos de frase de Navarro Tomás, y en ese sentido constituye un avance. Sin embargo, también puede comprobarse que en su descripción se mezclan los criterios formales, por niveles o no, semánticos, gramaticales y discursivos y que, finalmente, se limita a ofrecer ejemplos inequívocos y a señalar su pronunciación.

Quilis sólo ofrece, por tanto, dos patrones entonativos claramente formalizados: el final del enunciado declarativo /21↓/ y el final del enunciado interrogativo absoluto /12↑/; es decir, el descenso y el ascenso finales. Otro dato que llama la atención de su descripción es que el nivel tonal /3/ queda circunscrito al ámbito de las entonaciones expresivas. En resumen, y a pesar del aparato teórico, del instrumental empleado y de la erudición desplegada por este autor, la información formal que realmente aporta resulta, a la postre, muy escasa.

2.5.3. *Tendencias actuales*

Las últimas aportaciones al estudio de la entonación en nuestro país suponen un cambio de dirección o, más bien, una vuelta al sentido original de lo que llamamos la "escuela española": análisis de configuraciones y magisterio de Navarro Tomás. A estas características hay que añadir una nueva y fundamental: el análisis instrumental de los datos y la metodología experimental que siguen los trabajos. En este sentido, merecen destacarse las investigaciones realizadas por V. Salcioli (1988) y por J. M. Garrido (1991).

El trabajo de V. Salcioli "Estudio fonético-experimental de la entonación interrogativa catalana" (1988, resumen de la tesis doctoral presentada por la autora un año antes) todavía presenta una cierta influencia del análisis de niveles, aunque tanto el planteamiento teórico como la representación que ofrece de las entonaciones analizadas se encuadran en un análisis de configuraciones coherente con la doctrina de Navarro Tomás: por ejemplo, uno de los objetivos del trabajo es "encontrar el campo entonativo del catalán" (op.cit.: 39); la unidad de medida empleada es el semitono; en cuanto a "la tipología de las preguntas (analizadas) sigue, dentro de lo posible, la de Navarro Tomás" (op.cit.: 40); etc. Por lo demás, Salcioli considera compatibles ambos métodos y llega a identificar (como Hoyos Andrade) los tonemas enunciativos de Navarro Tomás con niveles tonales: "Navarro Tomás, como es sabido, realiza una subdivisión de cinco niveles" (op.cit.: 39).

Las frases analizadas, pertenecientes a diversos informantes, se produjeron de forma espontánea en una entrevista guiada. Para su análisis, se empleó un analizador de tono controlado por ordenador (el *Visi-Pitch* de *KAY El.*, al que nos referimos en el apartado 1.3.3.). Con los datos extraídos del análisis se

elaboró un esquema de la curva entonativa en el que se marca el "tono normal" del emisor y los movimientos tonales relevantes. Paralelamente, se ofrece una notación en la que se marcan los niveles tonales de la frase.

Sin embargo, la notación por niveles que ofrece Salcioli consiste más bien en una división del campo de entonación de la lengua, cuantificada con precisión, dado el carácter experimental del trabajo: tres semitonos de distancia entre cada nivel o "índice fonético". Después de efectuar su análisis, dice la autora: "el campo de entonación de la pregunta ... se sitúa alrededor de la octava" (op.cit.: 65). Y continúa: "Una división del campo entonativo podría ser:

Superagudo	3 s/t
Agudo	3 s/t
Semiagudo	3 s/t
Normal	-----
Semigrave	3 s/t
Grave	3 s/t
Supergrave	3 s/t

En consecuencia, podríamos hablar de 7 índices fonéticamente necesarios para describir la entonación interrogativa del catalán:

- 7 - Superagudo
- 6 - Agudo
- 5 - Semiagudo
- 4 - Normal
- 3 - Semigrave
- 2 - Grave
- 1 - Supergrave" (op.cit.: 66).

Si se tratara de un auténtico análisis de niveles, una gradación de siete niveles tonales podría resultar sorprendente. Pero la autora no hace ninguna referencia a junturas, ni a los acentos, ni habla de morfemas entonativos, ni de fonemas suprasegmentales, etc.: su única intención es encontrar una manera de formalizar con objetividad las configuraciones halladas, para lo que se sirve de los siete niveles con los que divide el campo de entonación, junto con la representación esquemática de la curva (v. figura 37, en la que puede verse un ejemplo del tipo de representación de la curva entonativa que ofrece Salcioli).

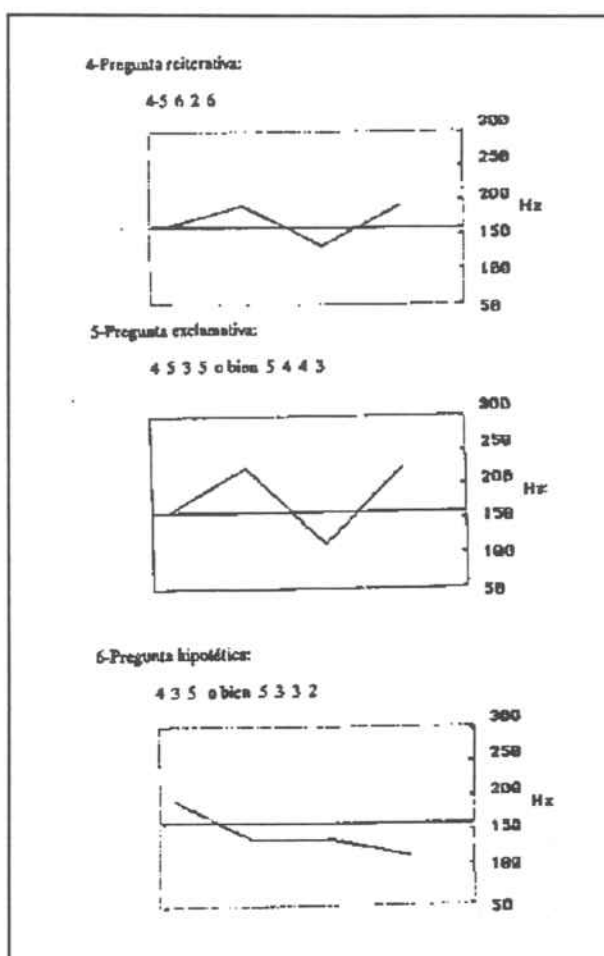


Figura 37 - Representación estandarizada de las curvas entonativas (Salcioli, 1988).

La tipología de frases interrogativas estudiadas y su descripción por niveles es la siguiente:

Pregunta de indiferencia o desconocimiento del emisor:	
"Que has vist en Pere?"	4-5 3 2-1
Pregunta relativa:	
"No deuran ser pas els amics de sempre?"	5 5 6 3
Pregunta restrictiva:	
"És demà quan s'en van?"	5 6 4 5
Pregunta reiterativa:	
"Oi que vindràs demà?"	4-5 6 2 6
Pregunta exclamativa:	
"Però encara sou aquí?"	4 5 3 5 o bien 5 4 4 3
Pregunta hipotética:	
"Serà veritat el cas de la Maria?"	4 3 5 o bien 5 3 3 2
Pregunta disyuntiva:	
"Què vols, aigua o «graciosa»?"	4 7 5 3

Por su parte, el trabajo de J. M. Garrido *Modelización de patrones melódicos del español para la síntesis y el reconocimiento de habla* (1991) supone un paso más allá en la formalización de las entonaciones.

En realidad, se trata de un trabajo preliminar, en el que se establece un corpus de frases y el modelo entonativo de las mismas, siguiendo un proceso de análisis acústico, como un primer paso para un trabajo más amplio que incluya también criterios perceptivos que validen tales modelos. El corpus de frases entonativas se establece según la tipología ofrecida por Navarro Tomás ("para la determinación de las modalidades oracionales incluidas en este estudio se ha tomado como referencia el trabajo de Navarro" -op.cit.: 45-), controlando el número de sílabas y la posición de los acentos; posteriormente, las frases, leídas

por los informantes, se grabaron y se analizaron en el programa *MACSPEECH LAB*.

Las curvas obtenidas en el analizador de tono son sometidas, entonces, a un proceso de estilización que consiste en: estandarizar la duración de las distintas emisiones de la misma frase; medir sólo el valor de la F_0 en el inicio de la curva, en cada "punto de inflexión" y en el final ("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 más" -op.cit.: 41-); y "normalizar" las frecuencias medidas, para trabajar con valores relativos, y no absolutos ("el sistema utilizado ... 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" -op.cit.: 42-). El resultado es un esquema de la curva entonativa del que se han eliminado las variaciones "micromelódicas" y las variaciones personales de cada informante.

Los patrones obtenidos son el resultado, pues, de frases gramaticalmente completas, en las que sólo se contempla la línea melódica, a pesar de que Garrido define la entonación como un "fenómeno lingüístico relacionado con la sensación perceptiva que produce la variación a lo largo de todo un enunciado de tres parámetros físicos -frecuencia fundamental, amplitud y duración" (op.cit.: 7). Los tipos o modalidades oracionales contemplados son:

Enunciativas ("Ramón llegó en avión")

Interrogativas absolutas ("¿Viene hoy Ramón?")

Interrogativas pronominales ("¿Dónde vivió Ramón?")

Interrogativas relativas ("¿Ramón llegó en avión?")

Exclamativas ("¡Ramón llegó en avión!")

Volitivas ("Llama a Ramón")

Todas las frases, pues, comparten similares características acentuales y de número de sílabas, y todas constituyen un único grupo de entonación.

Los patrones melódicos básicos obtenidos son:

Patrón descendente: "propio de frases enunciativas, de algunas interrogativas pronominales, de bastantes exclamativas y de la mayoría de las volitivas" (op.cit.: 96), y que equivaldría a la "semicadencia" de NavarroTomás.

Patrón ascendente-interrogativo: "el patrón típico de las frases interrogativas absolutas, de algunas pronominales y de las interrogativas relativas" (ibid.), que equivaldría a la "anticadencia".

Patrón ascendente-enunciativo: "propio de frases con contenido expresivo, pero que también guardan relación con la enunciación. Así ocurre con las exclamativas y volitivas, concretamente" (op.cit.: 100). Equivaldría a la "semianticadencia".

En la figura 38 se puede observar el tipo de descripción que ofrece Garrido de estos patrones (en concreto, del patrón ascendente-enunciativo).

El autor también obtiene otros patrones, que llama "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" (op.cit.: 100). Tales formas superpuestas serían: el "esquema ondulado", que consiste en un mayor número de picos; el "aumento del rango frecuencial", que consiste en que se aumentan las diferencias entre el punto más elevado y el más bajo de la curva; el "esquema circunflejo", en la última sílaba de la frase; la "elevación de la altura tonal o la pendiente del primer pico"; y la "variación de la pendiente en el segmento final".

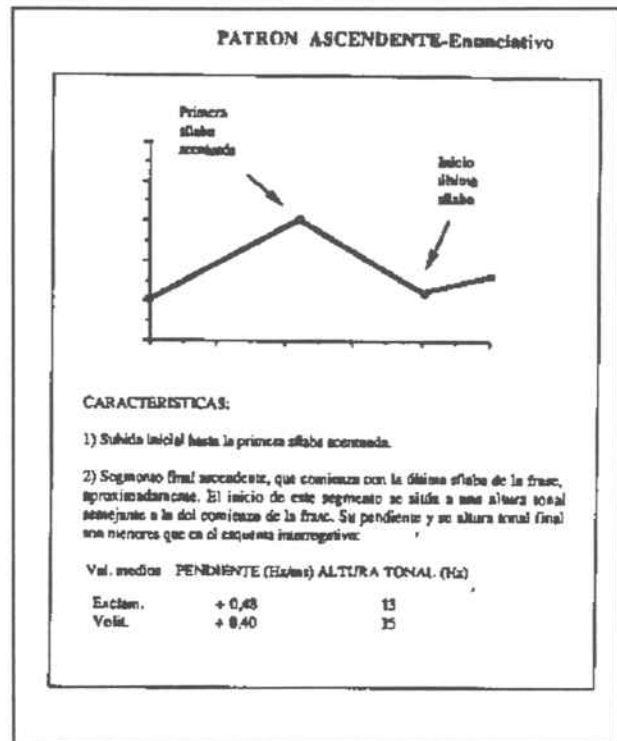


Figura 38 - Descripción del "Patrón ascendente-enunciativo" (Garrido, 1991).

Los resultados de su trabajo, dice Garrido, "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" (op.cit.: 106). Tal afirmación, sin duda, es importante por cuanto manifiesta el magisterio de Navarro Tomás, incluso tratándose de un estudio experimental, etc.; sin embargo, también aporta otro tipo de información: y es que el corpus analizado por Garrido está tan apegado a la descripción y a los elementos conceptuales propuestos por Navarro Tomás que no es extraño que sus conclusiones concuerden. El mismo hecho de que las frases analizadas no fueran producidas espontáneamente sino preparadas

ad hoc y leídas aisladamente (o en un contexto mínimo, también leído) ayuda a esta coincidencia.

En resumen, Garrido considera que el "segmento final" ascendente o descendente son la marca de la enunciación y la interrogación, respectivamente (aunque en sus patrones considera toda la melodía como una unidad). Sus patrones son equivalentes, según su opinión, a la "semicadencia", la "anticadencia" y la "semianticadencia" de Navarro Tomás, y añade no encuentra datos que validen la "cadencia"; el tonema continuativo, con buen criterio, lo considera propio de grupos no terminales, aunque él no los contemple en su estudio: "las tesis de Navarro Tomás parecen confirmarse, frente a las de Quilis y otros autores, en cuanto al número de tonemas finales en español" (op.cit.: 107). Sin embargo, y siguiendo la tradición, esto supone que también Garrido ignora los modelos interrogativos propuestos por el autor del *Manual de entonación* en su capítulo 4.

Por último, Garrido observa, a la vista de sus resultados, que no hay diferencias importantes entre las frases exclamativas y 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 F_0 en el segmento final, la utilización en ambos casos tanto de patrones ascendentes como descendentes, la aparición de finales circunflejos en uno y 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" (op.cit.: 108).

2.6. Conclusión: cuestiones abiertas en el estudio de la entonación

A lo largo de este capítulo hemos podido comprobar que la simple distinción entre análisis de niveles y análisis de configuraciones constituye una visión un tanto pobre, aunque esencialmente correcta, de los diversos métodos con los que se ha intentado concebir y analizar la entonación. En este último apartado, por tanto, trataremos de reformular y completar dicha distinción básica.

Por otra parte, hemos visto también cómo los distintos enfoques parten de concepciones de la entonación parcialmente distintas, que manejan incluso conceptos no homologables. Nuestro propósito es, pues, señalar los aspectos teóricos no resueltos por las distintas tradiciones en particular, ni por el cruce entre todas ellas. Como veremos, y a pesar de la numerosa bibliografía que ha tratado el tema durante varias décadas, el estudio de la entonación se encuentra aún en un estadio de desarrollo muy incipiente, en el que, al parecer, las únicas certezas incuestionables son las que, previamente, procura el propio enfoque teórico.

En general, y a tenor de lo que hemos visto, podría hablarse de cinco enfoques metodológicos distintos, que se engloban en los tradicionales análisis de niveles y análisis de configuraciones.

El análisis de niveles, en principio, es un método exclusivo de la tradición norteamericana (por más que algunos autores europeos puedan considerarse seguidores más o menos afortunados o coherentes); sin embargo, y dentro de esta misma tradición, observamos dos enfoques claramente diferenciados.

Por una parte, el análisis de niveles tradicional, *melódico*, de carácter

estructuralista, cuyo principal exponente es la obra de Trager & Smith (1951): en ella, como vimos, se considera toda la frase como una unidad entonativa completa, un "morfema" entonativo formado por tres tipos de "fonemas" suprasegmentales, que son los niveles tonales, los acentos y las junturas. Toda la frase entonativa, por tanto, constituye una unidad melódica definida en su conjunto: según este punto de vista clásico, cualquier diferencia "fonemática" implicaría una diferencia "morfológica" o, lo que es lo mismo, dos patrones similares como /1211↓/ y /1221↓/ constituirían dos unidades entonativas distintas. Es decir, el *análisis de niveles melódico* contempla como forma entonativa la melodía completa de la frase.

Por otra parte, la irrupción de la fonología generativa en la tradición norteamericana de análisis de la entonación supuso, como vimos, la conservación del método de análisis por niveles y el nacimiento de un nuevo enfoque teórico: la fonología métrica, creada por Liberman & Prince (1977). Así, frente al análisis de niveles melódico surge el *análisis de niveles rítmico* o *métrico*, caracterizado por una perspectiva en la que la melodía es una dimensión superpuesta y generada por el esquema rítmico de las frases: si para el análisis melódico el acento era un mero constituyente más de la entonación, para el análisis métrico la estructura acentual de la frase es previa (generada por el propio nivel segmental -y léxico- de la frase), y es la que condiciona la estructura melódica de la misma, según una serie también de reglas generativas. El ejemplo más influyente de este enfoque lo tenemos en Pierrehumbert (1987).

En ambos enfoques, el carácter fonológico de la entonación queda fuera de toda duda, así como su rendimiento netamente lingüístico, aunque en raras ocasiones se hace una descripción sistemática de los significados o las condiciones de uso de los "morfemas" entonativos descritos formalmente.

Dentro del análisis de configuraciones, hemos visto que la tradición británica constituye una rama propia y bien diferenciada, que podríamos nombrar como *análisis de configuraciones nuclear*. En resumen, este enfoque consiste en considerar que toda la información entonativa reside en el núcleo del contorno (que no ha de coincidir necesariamente con una frase gramatical completa) cuya naturaleza es a la vez acentual y entonativa: es el acento principal del contorno, pero también su núcleo entonativo, y consiste en una inflexión tonal, de cuyas características depende el significado entonativo y el uso de dicho contorno. La descripción de los contornos entonativos, pues, se reduce a la descripción formal de los "patrones nucleares" o *tone-patterns* y a sus significados o condiciones de uso. La perspectiva de Bolinger, expuesta en sus dos obras más generales (1986 y 1989), es una versión un tanto personal de este enfoque, que encuentra en la obra de Cruttenden (1986) su formulación más completa y sistemática, en la que se recoge lo más relevante de toda la tradición.

El análisis de configuraciones, sin embargo, y al margen de la tradición británica, ha observado un desarrollo importante en el resto de Europa contemplando no el núcleo del contorno sino la melodía del mismo en su conjunto: es lo que podríamos llamar *análisis de configuraciones melódico*, para el cual el acento principal de la frase únicamente determina la inflexión final del contorno, pero no contiene por sí mismo toda la información entonativa relevante. Es más, ni siquiera se considera un elemento propiamente entonativo: es decir, el *análisis de configuraciones melódico* no es comparable, por tanto, y a despecho de algunos autores europeos, al *análisis de niveles melódico*, para el cual el "fonema acentual" es un elemento constituyente de la melodía. En cualquier caso, este enfoque también tiende a considerar la frase gramatical como una unidad entonativa.

El *análisis de configuraciones melódico*, con todo, no constituye un enfoque unitario, y en él se distinguen dos versiones bien diferenciadas entre sí: la que contempla la melodía en su conjunto como un todo indivisible, que podríamos llamar *holística*; y la que, aun contemplando toda la frase como una unidad melódica completa (y en ello se diferencia del análisis *nuclear*), considera que es en la inflexión final en la que hay una mayor información entonativa, versión que podríamos llamar *de inflexión final*. El *análisis de configuraciones melódico holístico* encuentra en la escuela holandesa su representación más influyente y, en concreto, en la obra de 't Hart, Collier & Cohen (1990), cuyo planteamiento perceptivo y experimental constituye una novedad considerable en el campo de la entonación. Por su parte, el *análisis de configuraciones melódico de inflexión final* encuentra su principal representación en la escuela española y, en concreto, en la obra de Navarro Tomás (1944).

Los tres tipos de análisis de configuraciones comparten una acusada tendencia a la descripción fonética pormenorizada (que en ocasiones recuerda el método atomístico de los neogramáticos, sobre todo en la tradición británica) y, en cierto modo, declinan la responsabilidad de afrontar el carácter fonológico de la entonación, cuando no lo niegan explícitamente. Lo cual no es óbice, paradójicamente, para que se planteen permanentemente el tema del significado entonativo y de las condiciones de uso de las configuraciones descritas, como si se tratara de un signo lingüístico (excepto en la escuela holandesa, más preocupada por la descripción de la forma fonética, y que niega el carácter significativo de la entonación).

En resumen, los cinco enfoques metodológicos con los que se afronta el estudio y el análisis de la entonación son: el *análisis de niveles melódico*

(tradición norteamericana estructuralista), el *análisis de niveles métrico* (tradición norteamericana generativa), el *análisis de configuraciones nuclear* (tradición británica), el *análisis de configuraciones melódico - holístico* (escuela holandesa) y el *análisis de configuraciones melódico - de inflexión final* (escuela española).

En cada enfoque, el propio concepto de "entonación" no es homologable al de los demás. Así, para el *análisis de configuraciones melódico (holístico y de inflexión final)* la entonación es únicamente la línea melódica del habla, es decir, sólo se considera entonación las variaciones de F_0 a lo largo del discurso; para el *análisis de configuraciones nuclear*, en cambio, la entonación es un fenómeno complejo que también incluye al acento, aunque sólo al acento núcleo de contorno; por su parte, para el *análisis de niveles melódico* la entonación es un fenómeno suprasegmental en el que participan el tono (cifrado en niveles), el acento (que se considera de intensidad) y las junturas (que pueden ser inflexiones tonales, pero también pausas, etc.); finalmente, el *análisis de niveles métrico* considera la entonación como un fenómeno en el que la melodía depende de la estructura rítmica, que a su vez depende de la estructura acentual, que ha sido generada por la estructura léxico-fonética de la frase: es decir, la entonación es un fenómeno complejo, del que la melodía sólo es la parte más superficial.

En cambio, en lo que todos los enfoques parecen estar de acuerdo tácitamente es en el carácter "dinámico" del acento: el acento se considera de intensidad, en todos los casos, forme o no parte de la entonación. Incluso en la tradición británica, en la que precisamente se define la entonación en función del "acento melódico", que es un fenómeno tonal, los demás acentos siguen considerándose de intensidad, como si la naturaleza entre el acento de la frase

y el acento de la palabra fuera originariamente distinta.

La evidencia experimental de que el acento es un fenómeno de tono, así, o bien se ignora, simplemente, o bien se considera que tal evidencia afecta sólo al llamado "accento melódico", o bien se rechaza sin más contemplaciones: así, por ejemplo, el análisis métrico niega la validez de los trabajos experimentales que afirman el carácter tonal del acento, basándose en que tales trabajos no contemplan la estructura rítmica de la lengua en cuestión ni el carácter rítmico de la propia percepción del acento, por lo que los resultados están necesariamente falseados. Y es que el acento como fenómeno de intensidad es uno de los fundamentos comunes a las distintas teorías de la entonación, pues sólo si el acento se considera de intensidad puede mantenerse la tradicional distinción entre los fenómenos del acento y la entonación.

Llegados a este punto, observamos que los distintos enfoques existentes dejan una serie de cuestiones abiertas y sin respuesta: en unos casos, porque no se ha afrontado directamente su solución y, en otros, porque las diversas soluciones ofrecidas resultan contradictorias. Tales cuestiones abiertas podríamos reunir las, muy sumariamente, en cuatro grandes temas, que nos servirán, en adelante, de guía ideológica y de norte en nuestra propia descripción del fenómeno:

Qué relación mantienen acento y entonación:

Por una parte, parece claro el carácter tonal del acento, y en especial de la percepción del acento (v. 3.1.1.3.), lo cual implica que tanto el acento como la entonación son fenómenos dependientes del mismo parámetro físico: las variaciones de frecuencia fundamental (F_0), según afecten a un segmento o a

varios. Como hemos visto, todos los enfoques teóricos prefieren considerar, tradicionalmente, el acento como un fenómeno de intensidad para mantener la independencia de ambos fenómenos, pero esto contradice las evidencias experimentales de un modo flagrante.

Por otra parte, las explicaciones métricas sólo pueden aceptarse plenamente desde una perspectiva generativa, en la que, por ejemplo, la forma léxico-sintáctica se considera previa a la propia forma fónica del discurso: un planteamiento, sin duda, teóricamente interesante pero muy alejado de la percepción de la realidad que subyace en cualquier otro enfoque teórico y, por tanto, y difícilmente asumible desde fuera.

Es decir, las relaciones entre acento y entonación aceptando plenamente el carácter tonal de ambos fenómenos constituye uno de los puntos de mayor transcendencia en el estudio de la entonación, pero también uno de los puntos peor tratados por las distintas tradiciones. Está en juego la propia definición del fenómeno.

La entonación es un fenómeno lingüístico o únicamente expresivo, y qué nivel de análisis le corresponde:

En muchas ocasiones, las descripciones de la entonación ni siquiera se detienen en este punto: tratan la entonación como un fenómeno estrictamente fonético y no se plantean en ningún momento su rendimiento lingüístico. En otras ocasiones, a la afirmación de su carácter fonológico le sigue una descripción estrictamente formal y se remata con una descripción de las entonaciones en términos "morfológicos" o "léxicos", de modo que no queda claro qué tipo de signo lingüístico se está describiendo. Es decir, no sabemos si la entonación debe describirse sólo en un nivel fonético, o bien en un nivel

fonológico, o bien en un nuevo nivel "morfológico" o "léxico".

En cualquier caso, a la indefinición en este punto se le añade una relación extraña y desigual con el nivel gramatical, con el nivel textual y con el nivel pragmático: muchos autores se refieren, por una parte, a la estrecha relación que mantienen sintaxis y entonación (en ocasiones, para negarle a esta carácter lingüístico); por otra parte, la naturaleza integradora de la entonación la convierte en un mecanismo de realce de los núcleos informativos del discurso, es decir, en un mecanismo de cohesión textual; y, por otra aún, el carácter actitudinal de su empleo parece conferirle un carácter más bien pragmático, dependiente del contexto y de la intención del hablante más que de un sistema de funcionamiento independiente, lo cual nos lleva, de nuevo, a plantear su naturaleza motivada, expresiva y posiblemente asistemática.

Y todo esto sin que tengamos claro de qué modo enfrentarnos a su posible estatus fonológico, pues la fonología suprasegmental no deja de ser un apéndice molesto dentro de la fonología clásica, que es una disciplina que se ocupa, por definición, del funcionamiento lingüístico de los segmentos fónicos, y que queda claramente desbordada por la realidad continua y cambiante de la entonación.

Así, conviene discriminar, en primer lugar, lo que es propiamente lingüístico en la entonación (esto es, sistemático, que constituya un signo lingüístico) de lo que no puede considerarse lingüístico, ya sea expresivo, contextual o *idiomático*.

En segundo lugar, hay que establecer el nivel de análisis que corresponde al fenómeno: si se trata de un fenómeno fonológico, como en principio parecería lógico suponer; o si excede el ámbito estrictamente fonológico y se trata, más bien, de un fenómeno de carácter léxico, o de un mecanismo de compensación gramatical, de integración discursiva, etc., y ver cómo se articula

todo ello coherentemente; o si se trata, tal vez, de un fenómeno que requiera de un nivel de análisis propio.

En tercer lugar, y por último, es importante tener muy presente el carácter *hablado* de la lengua para no perder de vista el carácter previo y *sine qua non* de la entonación, que no es, por tanto, un "añadido" a la lengua (identificada en nuestra imaginación más con la "lengua escrita" que con la "lengua oral": visión que explica, por ejemplo, perspectivas como la fonología métrica, en las que la entonación va "superpuesta" a una frase ya creada en la mente); en este sentido, la relación que pueda haber entre el nivel léxico-gramatical del enunciado y su entonación ha de contemplar el carácter de "contenedor fónico" de esta, omnipresente e ineludible en la lengua real.

Qué tipo de significado aporta la entonación:

Íntimamente ligado al tema anterior, la cuestión del significado alude también al carácter lingüístico del fenómeno: pues si hay una entonación lingüística no meramente integradora o expresiva, esta ha de constituirse con una forma definida y un significado estable, esto es, con un signo lingüístico propio.

Sin embargo, los distintos enfoques sólo han sabido identificar este supuesto significado con tipos de significado que ya aportan otros fenómenos lingüísticos: con significados léxico-semánticos (declaración, pregunta, imperativos), con significados aportados por la sintaxis (partículas interrogativas, pregunta disyuntiva o copulativa), dependientes del contexto discursivo (pregunta relativa, restrictiva) o situacional (cortesía, duda, ironía...).

Es razonable, pues, que algunos autores prefieran no entrar en el carácter lingüístico de la entonación, o que le nieguen sin ambages dicho carácter, pues no encuentran un nivel de significación que le sea propio, a no ser la matización

afectiva o actitudinal del enunciado: lo cual más bien parece apoyar la idea contraria, según la cual la entonación es un mero vehículo, no lingüístico, de la expresividad y la emoción, como lo es el gesto facial o el movimiento de las manos: un vehículo difícilmente sistematizable, motivado y universal.

La afirmación del carácter plenamente lingüístico de la entonación, por tanto, ha de pasar por la determinación de un nivel de significación propio, independiente de otros niveles de significación, y asociado sistemáticamente a una forma fónica definida.

Cuál es la unidad estructural de la entonación:

Una de las diferencias principales entre los distintos enfoques metodológicos es, precisamente, la unidad estructural que determinan. La opción más extendida es identificar la "frase entonativa" con la "frase gramatical"; sin embargo, esto supone subordinar el nivel entonativo al nivel gramatical, con los problemas que vimos que esto conlleva. Incluso el análisis de niveles melódico, con su descripción fonológica de la entonación, hace depender la unidad entonativa, el "morfema entonativo", de la unidad gramatical (la cláusula o sintagma gramatical), tendencia que se verá más acentuada aún con el análisis de niveles métrico.

El análisis de configuraciones también ha hecho coincidir tradicionalmente el contorno entonativo con la oración gramatical, aunque la propia división del contorno en partes (*pre-head / head / nucleus / tail* en la tradición británica; "inflexión inicial / cuerpo / inflexión final + rama inicial / rama final de la frase", en la escuela española) parece intentar una distinción entre el nivel estrictamente gramatical y un nivel entonativo propio.

Así, y paralelamente al tema del significado, una entonación plenamente

lingüística sólo puede concebirse si la descripción estructural de sus unidades no depende de otros niveles de análisis lingüístico, como el nivel gramatical. Es preciso encontrar una unidad propiamente entonativa, que muy bien podría ser el "contorno", siempre que este no se haga coincidir necesariamente con ninguna unidad sintáctica. En este sentido, tal vez sea la tradición británica la que ha ido más lejos, postulando unos límites de contorno que no tienen que ver necesariamente con la coherencia sintáctica de las palabras que contiene.

3. EL ACENTO Y LA ENTONACIÓN PRELINGÜÍSTICA

3. EL ACENTO Y LA ENTONACIÓN PRELINGÜÍSTICA

3.1. Acento paradigmático y ritmo

3.1.1. Definición de acento

3.1.1.1. Rasgo o fenómeno suprasegmental

El *acento* es un elemento lingüístico que suele definirse, desde dos perspectivas que a veces se confunden, o como un "rasgo" o como un "fenómeno" suprasegmental.

Como "rasgo suprasegmental" (o prosódico), se distinguiría de los "rasgos segmentales" en dos sentidos: por una parte, no afecta a un solo segmento de la cadena sonora sino a más de uno, a varios segmentos a la vez; por otra parte, no es opositivo sino contrastivo, esto es, no se opone a otro "rasgo suprasegmental" para definir una unidad superior (como ocurre con los rasgos segmentales, que se oponen entre sí para definir los fonemas), sino que sirve para contrastar unidades ya definidas, sobre las que actúa. Así, según Quilis, "el acento es un rasgo prosódico que permite poner de relieve una unidad lingüística superior al fonema (sílabas, morfemas, palabras, sintagmas, frases; o un fonema, cuando funciona como unidad de nivel superior) para diferenciarla de otras unidades lingüísticas del mismo nivel" (1981:310).

Los rasgos segmentales, como vimos en la introducción, son verdaderamente "rasgos" del timbre de los sonidos, que, desde un punto de vista fonético y positivo, sirven para segmentar el continuum sonoro en segmentos definidos; y, desde un punto de vista fonológico y opositivo, sirven para definir los fonemas como unidades funcionales. El *acento* como "rasgo suprasegmental",

por su parte, estaría relacionado con la intensidad, el tono y la cantidad. En realidad, y por tanto, los "rasgos segmentales" serían verdaderos "rasgos" de una característica compleja del sonido (el timbre), mientras que el acento, concebido como "rasgo suprasegmental", incluiría en bloque a las demás características del sonido, que no podrían, así, ser consideradas "rasgos", a su vez, ni mucho menos estar definidas por una serie de "rasgos" propios, como el timbre.

Es decir, considerar el acento como un "rasgo" suprasegmental equiparable a los "rasgos" segmentales supone equiparar el conjunto de las características de intensidad, tono y cantidad a un solo rasgo tímbrico. Sin embargo, sabemos que el timbre constituye una realidad unitaria (aunque podamos descomponerla en rasgos: que son los rasgos segmentales) del mismo nivel que, por ejemplo, el tono o la intensidad (realidades físicas que, por tanto, también podríamos descomponer en rasgos).

En resumen, de ningún modo puede considerarse el *acento* como un "rasgo suprasegmental" equiparable a los rasgos segmentales, porque estos son rasgos de timbre.

Históricamente, una variante de esta perspectiva ha sido la de considerar tres tipos de *acentos*: el "acento de intensidad" (o "dinámico", "expiratorio", o "de energía"), el "acento de tono" (o "tónico", "tonal", "musical", "melódico", "de altura", etc.) y el "acento de cantidad" (o "de duración"). Esta posición (expuesta, por ejemplo, por Navarro Tomás, 1918) considera que el acento es un rasgo prosódico, pero que hay distintos "acentos" porque hay distintos tipos de "rasgos prosódicos": así, cada una de las características del sonido, excepto el timbre, se consideraría un rasgo prosódico independiente.

La otra perspectiva, esencialmente distinta aunque muy cercana a esta última variante, es la más extendida y considera el *acento* como un "fenómeno

lingüístico" informado por los parámetros físicos de intensidad, cantidad y tono, que serían considerados, por tanto, también como rasgos suprasegmentales, aunque más bien interdependientes. Así, Martínez Celdrán dice del acento (en el capítulo "Los rasgos fonéticos suprasegmentales o rasgos prosódicos"): "el término *-acento-* hace referencia a un procedimiento fónico por el que una parte de la cadena hablada se pone de relieve, es decir, se destaca a expensas de las otras partes que la rodean" (1984: 243).

Como consecuencia de esta perspectiva, también se emplea "acento" para referirse a una realidad positiva: cuando se habla "acento primario", "acento secundario", "acento débil", etc., o bien de "acento de insistencia", "acento enfático", etc., el término se usa para referirse a los productos concretos del *acento* como fenómeno lingüístico.

Con todo, también desde esta perspectiva se consideran "rasgos" suprasegmentales a las otras características del sonido, oponiéndolas en su conjunto, de nuevo, a los rasgos (segmentales) del timbre, como si este constituyera una realidad aparte. Por el contrario, y como sabemos, el timbre es una característica más, que además depende de la producción tonal igual que el propio tono, etc.

Es decir, no se ha sabido establecer, en primer lugar, una jerarquía clara entre "característica del sonido" (tono, intensidad, timbre y cantidad) y "rasgos" en los que puede descomponerse (aunque, tradicionalmente, sólo el timbre se ha descompuesto en rasgos, mientras las otras características del sonido se han considerado en bloque).

En segundo lugar, el carácter "segmental" de los rasgos tímbricos no es tan absoluto como parece pues, como vimos en el apartado 1.2.1., los rasgos de timbre se extienden a lo largo del continuum sonoro afectando en ocasiones a

varios segmentos, lo cual se manifiesta en fenómenos fonotácticos como la asimilación, etc.; e incluso hay fenómenos que los convierten en "suprasegmentales" (como por ejemplo la armonía vocálica, que puede extender un rasgo segmental como "abertura" o "nasalidad" a varios segmentos vocálicos del continuum -v. Martínez Celdrán, 1989: 68 y ss.-). El propio concepto de "segmento" (que, como vimos, es una abstracción intermedia entre el concepto de fonema y la propia realidad sonora) más que definir el carácter (segmental) de los rasgos tímbricos es definido por tales rasgos para poder concebirse aisladamente.

Finalmente, la característica de "cantidad" vimos que no era homologable a las otras tres características del sonido, sino que se trataba de una característica "de segundo grado", puesto que necesitaba de una abstracción previa para poder aplicarse (la abstracción "segmento"): es decir, en ningún caso la cantidad podría constituir un conjunto solidario con el tono y la intensidad (como "rasgos suprasegmentales") frente al timbre (cuya descomposición da lugar a los rasgos segmentales), pues eso supondría subvertir los términos de la relación acústica que tales características mantienen entre sí.

En general, y resumiendo, podemos definir el acento como el fenómeno lingüístico que consiste en poner de relieve una parte de la cadena hablada sobre el resto. Sin embargo, tampoco esta definición resulta completamente satisfactoria: por una parte, la expresión "una parte de la cadena hablada" es demasiado inespecífica y, por otra, no se determina por qué medios se pone de relieve sobre el resto (aunque, en principio, es de suponer que intervienen el tono, la intensidad y la cantidad).

3.1.1.2. Fenómeno silábico o vocálico

Cuando Quilis habla del acento como de un "rasgo prosódico" con el que poner de relieve "una unidad lingüística superior al fonema (sílabas, morfemas, palabras, sintagma, frase...)" tal vez está dejándose llevar por el entusiasmo, o tal vez está empleando un sentido no estrictamente fonológico del término "accento", y se refiere al concepto más amplio de "prominencia", "énfasis" o "relevancia". En cualquier caso, parece claro que el *accento* no consiste en poner de relieve un sintagma o una frase sobre otros sintagmas u otras frases, y que la opción de Martínez Celdrán de referirse a "una parte de la cadena hablada" es mucho más ponderada.

Generalmente, suele considerarse que el acento actúa en un nivel silábico, pues la sílaba es la unidad de abstracción intermedia entre el fonema y la palabra y, por tanto, es el constituyente fónico inmediato de la palabra (en cuyo nivel cumple el acento, como veremos, una "función distintiva", distinguiendo palabras como "llevo / llevó").

Una sílaba está constituida, necesariamente, por una vocal, que será su núcleo, y por un número indeterminado de consonantes y semiconsonantes (márgenes, semimárgenes y seminúcleos silábicos), que pueden no aparecer. Al tratarse de una unidad de abstracción fonológica, sin embargo, los límites fonéticos de la sílaba tampoco aparecen claramente: no hay ningún índice fonético que permita segmentar una sílaba de otra (Malmberg, 1974: 187, explica que las transiciones consonánticas pueden dar algunas pistas sobre los límites de la sílaba, según unas pruebas realizadas mediante síntesis de voz; sin embargo, reconoce que el simple análisis acústico no permite reconocer límite alguno). Es decir, si bien el continuum sonoro puede segmentarse en segmentos tímbricos con una base acústica innegable (los rasgos de timbre), no se ha

encontrado ninguna base acústica que permita segmentar claramente el continuum sonoro en sílabas.

Distintos autores han pretendido ofrecer una definición satisfactoria de la unidad sílaba, con desigual fortuna (v. un resumen en Martínez Celdrán, 1984): en todos, la idea central es que hay una vocal que ejerce de núcleo (máxima abertura) que puede estar flanqueada, o no, por consonantes; el límite entre tales consonantes, precisamente, es el objeto de la discusión.

En realidad, el concepto de "sílaba" es más intuitivo que científico, y se basa en la vocal como núcleo de la misma. Lo de menos, entonces, es establecer límites entre las sílabas, pues lo único que cuenta, a todos los efectos, es el carácter nuclear de la vocal y el carácter marginal de las consonantes. Por otra parte, una vez segmentado el continuum sonoro en segmentos tímbricos, base de la segmentación fonológica, no tendría sentido añadir a esta una segmentación posterior que requiera una abstracción fonológica más allá del fonema. Como dicen Canellada & Kuhlmann (1987: 43): "la frontera silábica no es fonológicamente pertinente".

Entendiendo, pues, la sílaba como una unidad definida exclusivamente en función de la vocal, queda claro que la marginalidad de las consonantes lo es a todos los efectos. Así, decir que el acento actúa a un nivel silábico equivale a decir que actúa a un nivel vocálico, y decir que tal o cual sílaba está acentuada equivale a decir que lo está la vocal de la sílaba, etc.

Trubetzkoy afirma, en un principio, que "en la mayoría de las lenguas del mundo, las particularidades prosódicas existen sólo en las vocales. Se podría, pues, tender a considerar estas particularidades entre las particularidades vocálicas y a tratarlas al mismo tiempo que los grados de apertura y las series de localización" (1939: 166). Sin embargo, más adelante niega esta posibilidad:

"pero esto se basaba en un error. Las particularidades prosódicas no conciernen a las vocales en cuanto a tales, sino a las sílabas. Una parte de los fonemas que componen la sílaba pueden ser prosódicamente no pertinentes. Por lo general, estos fonemas son consonantes. Pero también pueden ser vocales que, en este caso, son «asilábicas». Por otra parte, en ciertas lenguas aparecen sílabas que no contienen ningún fonema vocálico, de modo que la parte prosódicamente pertinente está ocupada por un fonema consonántico: en este caso se trata de consonantes «silábicas» ... Por consiguiente, las particularidades prosódicas no deben ser consideradas como particularidades de las vocales ... sino como particularidades de una determinada parte de la sílaba; y esta parte de la sílaba debe ser definida en forma distinta según la lengua de que se trate" (íbid.).

Es decir, para Trubetzkoy hay una parte de la sílaba que es relevante, su núcleo, que en algunas lenguas no es necesariamente una vocal. A esta parte prosódicamente relevante de la sílaba la llama "centro silábico": "será designada como centro silábico la parte de la sílaba que según las leyes de la lengua de que se trata detenta particularidades prosódicas distintivas" (íbid.). En nuestra lengua, obviamente, ninguna consonante puede ser centro silábico, y sólo pueden serlo las vocales y los diptongos (las semiconsonantes, en todo caso, sería discutible si forman parte o no del centro silábico: Martínez Celdrán, por ejemplo, las considera "seminúcleo silábico" -1984: 366- y fonemas independientes, a los que llama "glides" -1989: 93-).

Por tanto, concebir el acento en castellano como un fenómeno estrictamente vocálico no sólo no es contradictorio con los postulados de Trubetzkoy, sino que supone una asunción completa de los mismos: para él, sólo el "centro silábico" es relevante prosódicamente, y esto equivale, en castellano, a decir que sólo las vocales son relevantes en los fenómenos del acento y la

entonación.

3.1.1.3. Naturaleza tonal del acento

Decíamos que, en principio, para poner de relieve una vocal frente a las demás vocales de la cadena intervienen los parámetros del tono, la intensidad y la cantidad. Como vimos en el capítulo anterior, sin embargo, para la mayoría de los estudiosos de la entonación el acento debía considerarse fundamentalmente de intensidad, lo que les permitía distinguir netamente entre los fenómenos del acento y la entonación. Sólo en la tradición británica se considera la existencia de un acento "melódico" (*accent* - el núcleo de la entonación) que, con todo, no niega el carácter "dinámico" de los demás acentos de la frase (*stress*).

Sin embargo, en todos los trabajos experimentales llevados a cabo en diversas lenguas (v. un resumen en Quilis, 1981: 320 y ss.) se llega a la conclusión de que la intensidad es un parámetro despreciable en la percepción del acento, y de que el parámetro más relevante es el tono, seguido de la cantidad: "como se desprende de las investigaciones realizadas hasta ahora, hay acuerdo casi general en destacar la frecuencia fundamental como índice del acento, seguido por la duración; la intensidad queda como índice muy secundario" (op.cit.: 326).

Esta contradicción entre la tradición y los análisis experimentales tiene su origen en el carácter coarticulatorio de la intensidad, que ya vimos en el capítulo de introducción: cualquier movimiento en la producción tonal implica un cambio en la intensidad y, más específicamente, cualquier elevación de tono (producida, como dijimos, por un aumento de la tensión de los músculos laríngeos) está asociada, fisiológicamente, a una elevación de la intensidad. Por tanto, no es de

extrañar que los fonetistas tradicionalmente consideraran el acento como un fenómeno de intensidad pues, dado que la vocal acentuada siempre es más larga, más fuerte y coincide con las inflexiones melódicas del habla, esta era la mejor manera de distinguirlo claramente de la entonación, cuya filiación tonal era innegable. Sin embargo, en los trabajos experimentales llevados a cabo desde los años 50 se ha adoptado un punto de vista perceptivo, de modo que lo que interesa averiguar no es qué parámetro ocurre, sino qué parámetro de los que ocurren es más relevante para que el oyente detecte la vocal acentuada.

En español, las primeras investigaciones de este tipo las realizan Bolinger & Hodapp (1961) y, posteriormente, H. Contreras (1963): en ambos trabajos se concluye el carácter tonal del acento, (frente al tradicional "accento dinámico", defendido por Navarro Tomás), se considera la inflexión tonal como un parámetro a tener en cuenta, la duración como un parámetro secundario y la intensidad como un elemento despreciable. Quilis (1971) corrobora estos datos, "pero dando a la duración más importancia de lo que los mencionados lingüistas parecen concederle" (1981: 329). Finalmente, el trabajo de M. J. Solé (1984), basado en la síntesis de palabras, en las que se variaba cada parámetro (tono, inflexión tonal, duración e intensidad) y se sometía a una prueba perceptiva, es definitivo: "el tono es la variable más importante en la percepción del acento, tanto aisladamente como en combinación; de tal manera que cualquier combinación que lo incluya será la que indique la presencia del acento. El tono incluso marca el acento cuando se opone en solitario a las otras tres variables ... A gran distancia le sigue en importancia la duración y por último la intensidad, de muy escasa relevancia en todos los casos" (Solé, 1984: 205).

Por tanto, ahora sí que podríamos concluir nuestra primera definición del fenómeno, diciendo que el acento es el fenómeno lingüístico que pone de relieve

una vocal sobre otras mediante un contraste tonal.

Así, las lenguas no tonales y las lenguas tonales estarían mucho más cerca, en realidad, de lo que siempre se ha creído: al fin y al cabo, una lengua como el español tendría un acento de un tono (vocal tónica / vocal átona), una lengua tonal como igbo tendría un acento de dos tonos, y una lengua también tonal como el chino mandarín tendría un acento cuatro tonos. O, lo que es lo mismo, decimos que una lengua es "tonal" cuando tiene más de un tono para distinguir vocales o sílabas.

3.1.1.4. Funciones del acento

Los manuales suelen distinguir, siguiendo a N. S. Trubetzkoy (1939), dos tipos posibles de acento entre las lenguas no tonales, según la posición que suela ocupar la sílaba tónica dentro de la palabra: el "accento libre" (como el del castellano, el inglés o el ruso) y el "accento fijo" (como el del francés, el polaco o el checo).

En las lenguas de "accento fijo", el acento recae siempre en la misma posición de la palabra: en francés, por ejemplo, todas las sílabas finales de palabra son tónicas; en polaco, todas las penúltimas; en checo, todas las primeras sílabas. En tales lenguas se dice que el acento cumple una "función demarcativa", esto es, que indica el final, o el principio, etc., de las palabras, y que por tanto constituye un poderoso medio en la segmentación léxica del discurso. Como veremos en el apartado 3.2.1., la comprensión auditiva del discurso requiere que la cadena hablada se le ofrezca fragmentada en bloques al oyente, *integrada*: sin embargo, dudamos que esta función pueda cumplirla el acento directamente a un nivel léxico; por otra parte, si la "función demarcativa", en ese nivel léxico, tuviera algún valor real en la comprensión del

discurso, difícilmente podrían prescindir de ella las lenguas de "acento libre". Así pues, creemos que esta función está un poco tomada por los pelos, tal vez para procurar una bella simetría.

En las lenguas de "acento libre", por su parte, el acento puede realizar cualquiera de las sílabas de cualquier palabra. La función que cumple, por tanto, será una "función distintiva": servirá para distinguir palabras constituidas por los mismos segmentos tímbricos, pero con distinta sílaba acentuada (como "TÉRmino, terMIno, termiNÓ"). De este modo, el acento se constituye en un elemento lingüístico de primer orden, pero sólo en las lenguas de "acento libre": de ahí el interés por encontrar una función también lingüística en las lenguas de "acento fijo".

En realidad, esta clasificación, como todas las clasificaciones binarias, resulta demasiado cerrada: por ejemplo, Cruttenden considera al español lengua de acento fijo (1986: 16), teniendo en cuenta la acusada tendencia de nuestra lengua a hacer recaer el acento en la penúltima sílaba de la palabra; para nosotros, sin embargo, se trata claramente de una lengua de acento libre pues, según datos de Quilis (1983), el 20% de las palabras que empleamos no son llanas sino agudas (17,68%) o esdrújulas (2,76%).

En cualquier caso, el rendimiento lingüístico del acento como elemento léxico distintivo (la "función distintiva" de las lenguas de "acento libre") es muy escaso: en nuestra lengua, por ejemplo, se reduce a la diferencia morfológica de algunas formas verbales (del tipo: "miro / miró" o "cobre / cobré") y a unos pocos pares léxicos más (como "abras / habrás", "término / termino", etc.).

Por ejemplo, las parejas léxicas del tipo "término / termino" o "andén / anden" (un sustantivo y una forma verbal distinguidos sólo por el acento) son semejantes a la homófonas "giro / giro", "paso / paso", "vino / vino", etc., que

a pesar de ser palabras homófonas difícilmente se confunden: el valor distintivo del acento en tales parejas, por tanto, es despreciable.

Del mismo modo, la pareja "miro / miró" es equivalente a la pareja "siento / sentó"; en esta última, sin embargo, la diferencia reside en la cualidad vocálica del diptongo "ie" frente a la vocal "e": aunque es cierto que en la forma "siento" el diptongo se formó históricamente a consecuencia del carácter tónico de la "e" abierta latina, para el hablante actual la diferencia entre la estructura acentual de ambas formas ("siento": llana / "sentó": aguda) no es el único dato relevante para su identificación, y podría bastar con su identidad tímbrica. Por otra parte, también encontramos en el paradigma verbal pares homófonos, como "él *sigue*" / "*sigue* tú": es decir, el valor distintivo del acento en los pocos pares morfológicos diferenciados por él es muy escaso, y tales pares podrían cumplir su función igualmente si no estuvieran diferenciados acentualmente.

En definitiva, la distinción acentual entre tales pares léxicos o morfológicos más bien es azarosa (aunque obedece a un origen histórico, sincrónicamente es "casual") y lingüísticamente apenas es significativa: es decir, se trata de una distinción que podría considerarse prácticamente redundante.

Ambos tipos de acentos, fijos y libres, cumplen por igual la llamada "función contrastiva", que consiste en distinguir las palabras que tienen alguna sílaba acentuada -las palabras tónicas- y las que no tienen ninguna sílaba acentuada -las palabras átonas-: sobre el hecho, muy llamativo, de que no todas las palabras contengan un acento hacemos referencia en el siguiente apartado.

Finalmente, se llama "función culminativa" al fenómeno por el cual las sílabas inacentuadas que rodean a la sílaba tónica constituyen con ella una unidad fónica: es lo que en el apartado 3.1.3. llamaremos "palabra fónica".

3.1.2. El acento paradigmático

3.1.2.1. El acento como elemento superpuesto

Como vimos en el capítulo anterior, la fonología generativa se ha aplicado a elaborar una serie de reglas con las cuales predecir los acentos en un discurso dado. Según esta perspectiva, en el "lexicón" ideal de cada lengua tiene que figurar, en forma de rasgos segmentales, información sobre el timbre de los sonidos que componen la palabra, pero no información sobre el acento de la misma: basta con conocer el timbre de las vocales (si son vocales plenas o reducidas), la categoría gramatical de la palabra y su posición en el discurso para determinar la posición del acento en la palabra, y su valor rítmico.

"Considérese la palabra inglesa *theatricality* «teatralidad». Dado que este nombre tiene estrechas conexiones sintácticas y semánticas con el adjetivo *theatrical* «teatral» y con el sustantivo *theatre*, su estructura sintáctica y fonológica podría representarse por la parentización: [# [# [#_N #_N] + ik + æl #_A] + i + ti #_N]. ¿Cómo actúan, entonces, las reglas del acento del inglés en esta palabra?" (Sommerstein, 1977: 263).

En esta cita puede comprobarse cómo el autor considera una forma fónica abstracta, sin acento, en la cual la categoría gramatical es determinante (los subíndices _N y _A se refieren a las categorías "nombre" y "adjetivo"), y sobre la cual el acento recaerá posteriormente, según "las reglas del acento del inglés".

O, dicho en otros términos, el acento no forma parte de la palabra, sino que es un elemento superpuesto. Es lo que podríamos llamar "carácter lecto-

escritor de las reglas de acentuación": puesto que podemos escribir una retahíla de letras (o de símbolos fonéticos, o cualquier otra fórmula de representación *escrita*) para referirnos exclusivamente a los segmentos tímbricos de la palabra, consideramos que sólo ellos constituyen su naturaleza fónica, y que sólo después el acento "recae" sobre una vocal determinada. Así, para la fonología generativa la forma fónica abstracta es previa a la forma sonora concreta: y dicha forma fónica abstracta sólo puede ser concebida como una transcripción *escrita* de los rasgos tímbricos abstractos. Por eso puede concebirse un "discurso dado" sin acentos: porque dicha concepción fonológica de la lengua es, propiamente, una concepción "afónica" y lecto-escritora.

Obsérvese cómo Chomsky & Halle describen el proceso de asignación de acento a un sintagma:

"Para empezar, está claro que en inglés hay por lo menos dos procesos de asignación de acento. Así, ¹*black*³*board* [pizarra], con un contorno acentual descendente, se debe distinguir de ²*black* ¹*board* [tabla negra], con un contorno ascendente. Los constituyentes elementales, *black*, un adjetivo, y *board*, un nombre, son los mismos para los dos casos; la diferencia está en el modo en que se combinan estos constituyentes, tal y como se refleja en sus diferentes estructuras superficiales, que ofrecemos a continuación ...:

(1a) [_N # [_A #black#]_A [_N #board#]_N #]_N

(1b) [_{SN} # [_A #black#]_A [_N #board#]_N #]_{SN}

En el caso (1a), donde el sintagma completo pertenece a la categoría «nombre», las reglas fonológicas deben dar el contorno 1-3 (acento primario seguido de acento terciario); en el caso (1b), donde pertenece a la categoría «sintagma nominal», las reglas deben dar el contorno 2-1 (acento secundario seguido del

acento primario) (Chomsky & Halle, 1968: 60-61).

Para estos autores, pues, la estructura superficial de ambas expresiones (*blackboard*: nombre compuesto; *black board*: sintagma nominal) es idéntica fónicamente, y la única diferencia es categorial, esto es, gramatical; la forma fonética, finalmente, será distinta porque el acento primario habrá recaído en un elemento distinto (el adjetivo *black* o el sustantivo *board*) según unas reglas de acentuación aplicadas *a posteriori*.

Es decir, para esta perspectiva teórica no supone ningún problema concebir niveles abstractos en los que las palabras no tengan acento, y en los que ni siquiera suenen. Es más, "hay pocas razones para suponer que el contorno acentual que se percibe deba representar alguna propiedad física del enunciado punto por punto; el hablante que utiliza el principio del ciclo transformacional y la regla de los compuestos y del acento nuclear debería «oír» el contorno acentual del enunciado que percibe y entiende, tanto si éste está físicamente presente en alguna forma como si no" (op.cit.: 80). Obviamente, el camino de la fonología métrica ya estaba trazado en esta frase, en la que el carácter "afónico" de la fonología generativa aparece con toda su fuerza: el oyente "oír" el acento aunque no "oiga" nada, aunque no haya parámetro físico alguno que lo marque físicamente, simplemente aplicando las reglas abstractas de asignación categorial (o métrica) del acento.

Sin embargo, concebir el acento como un fenómeno superpuesto a la palabra no es patrimonio exclusivo de la fonología generativa, en la que se manifiesta más llamativamente simplemente porque lleva a sus últimos extremos algunas tendencias algebraicas que ya se apuntaban en la fonología clásica (a fin de cuentas, M. Halle, fundador de la fonología generativa junto a N. Chomsky,

era uno de las principales figuras de la fonología clásica, y fue fundador también de la fonología binarista junto a R. Jakobson).

En general, en todas las tradiciones fonológicas subyace la idea de que el acento se superpone a la palabra. En este sentido, la fonología prosódica de Firth (1948) es un buen ejemplo de hacia dónde podía haber discurrido la fonología clásica de no haberse centrado tan exclusivamente en los fenómenos segmentales.

Para Firth, el sistema fonológico se compone de una serie de "fonemas" (definidos con rasgos segmentales) y de "prosodias", nombre con el que se designa a los fenómenos que afectan simultáneamente a varios segmentos: intensidad, tono y cantidad, pero también rasgos de timbre que pueden extenderse a varios segmentos, como en los hechos de armonía vocálica, de asimilación o de juntura. Los "fonemas", desde esta perspectiva, no coinciden exactamente con los fonemas de la fonología clásica, porque algunos rasgos "segmentales" pueden considerarse "prosodias": por ejemplo, el rasgo segmental de "abertura vocálica" sería una prosodia en un sistema vocálico andaluz de cinco fonemas, según la aproximación que desde la fonología prosódica llevó a cabo A. Contreras Jurado (1974) para dar cuenta de la armonía vocálica (la prosodia de abertura se superpondría, así, a la palabra en su conjunto); desde una perspectiva clásica, sin embargo, otros autores hablan de un sistema vocálico de 9 ó 10 fonemas (Salvador, 1977), en el que la abertura vocálica es un rasgo segmental más.

Así, para la fonología prosódica el acento es una prosodia; sin embargo, no queda claro si debe considerarse como un constituyente previo de la palabra o si actúa superponiéndose a ella. En cualquier caso, el hecho de considerar unas unidades tímbricas (los fonemas) sobre las que "recaen" las prosodias ilustra

muy bien hasta qué punto todo lo suprasegmental se considera siempre "posterior" a lo segmental, incluso desde una perspectiva teórica como la fonología prosódica.

Finalmente, y desde otro punto de vista, incluso las mismas "reglas de acentuación" del español sancionadas por la Real Academia obedecen a dicha concepción del acento *a posteriori*, a pesar de que son exclusivamente ortográficas y de que sólo sirven para determinar la presencia o no de tilde gráfica en la palabra escrita.

Obsérvese, por ejemplo, cómo se explica, desde una posición normativa (y en cierto modo "académica") lo que arriba hemos llamado "función distintiva" del acento:

"La situación de la sílaba tónica en el seno de la palabra puede determinar su significado. Así, la palabra *a-ni-mo* puede tener distintos valores significativos, según la posición del acento de intensidad. Con acento en la primera sílaba, *ánimo* equivale a «valor, esfuerzo, energía, intención»; con acento en la segunda, *animo* es la forma verbal correspondiente a la primera persona del singular del presente de indicativo del verbo *animar*, con acento en la última, *animó* corresponde a la tercera persona del singular del pretérito indefinido del mismo verbo" (Marsá, 1986:20).

El autor se refiere a una suerte de "abstracción tímbrica" de la "palabra *a-ni-mo*", sin acento, que engloba sus tres posibles versiones "acentuales". Sin embargo, sabemos que la "palabra *a-ni-mo*", sin acento, no puede considerarse como un único signo lingüístico, pues el significado de cada una de las tres formas acentuadas está asociado exclusivamente a tales formas, es decir, son tres

signos lingüísticos; por tanto, no puede ser una "palabra", sino tres. Lo que Marsá llama "palabra *a-ni-mo*", sin acento, apenas es una cadena de segmentos tímbricos abstracta, cuya única realidad efectiva es la forma escrita.

Es decir, de nuevo el carácter superpuesto del acento tiene que ver con una perspectiva lecto-escritora de la palabra, que permite concebir la unidad "palabra" antes de que recaiga sobre ella el acento, y antes, por tanto, de que existan los propios sonidos que la componen.

3.1.2.2. Falsos homónimos y falsos homófonos

Llámanse "homónimos" las palabras iguales (igual pronunciación e igual ortografía) que han dado en serlo desde etimologías distintas: como "haz" (de leña, del lat. *fascis*) y "haz" (y envés, del lat. *facies*); o como el ejemplo ya referido "vino" (sustancia) y "vino" (de "venir"). Por su parte, son "homófonas" las palabras que tienen la misma pronunciación, aunque su ortografía sea distinta: como "baca" y "vaca", "huso" y "uso", "barón" y "varón", etc.

Así, si considerásemos que el acento es un elemento superpuesto a la palabra, los pares "miro" / "miró", "llevo" / "llevó", etc., deberían considerarse en cierto modo homónimos (aunque, curiosamente, no homófonos): así se presentan, al menos, en los ejercicios escolares y en los dictados, en forma escrita, para que los alumnos procedan a distinguirlos colocando la tilde correspondiente.

Por otra parte, las palabras "sino" (sustantivo, del lat. *signum*) y "sino" (conjunción, contracción del grupo *si no es*, datada ya en 1140) podrían considerarse homónimas, y las palabras "asta" (sustantivo, del lat. *hasta*) y "hasta" (preposición, del ár. *háttà*) serían homófonas: entre ellas, sin embargo, la diferencia no es de posición acentual, sino de tonicidad. Separadamente, en

efecto, la conjunción "sino" y la preposición "hasta", palabras átonas en el discurso hablado, han de pronunciarse igual que "sino" y "asta", sustantivos y palabras tónicas: ya que es imposible pronunciar una palabra átona aisladamente. Por tanto, todo parece indicar que ambas parejas son homófonas y que sólo su inclusión en una cadena hablada confiere a los términos átonos su atonicidad.

Sin embargo, si adoptamos una perspectiva no lecto-escritora sino exclusivamente auditiva, difícilmente podrían confundirse las palabras tónicas y las átonas: no sólo ya por su categoría gramatical distinta, sino por su propia estructura fónica. Las palabras tónicas, como veremos en el próximo apartado, constituyen los núcleos de los grupos rítmicos, mientras que las palabras átonas se agrupan necesariamente alrededor de otras palabras tónicas, y no tienen ninguna independencia fónica. En realidad, pues, no se pueden pronunciar aisladamente y, cuando lo intentamos, estamos falseando su naturaleza.

Es decir, las palabras tónicas lo son por naturaleza: "sino" y "sino" no constituyen una unidad de ningún tipo, y el hecho de que la secuencia de segmentos tímbricos coincida en las dos palabras no las convierte en homónimas. Coinciden los segmentos, pero tampoco completamente: en el "sino" sustantivo, la vocal "i" es tónica; en el "sino" conjunción, la vocal "i" es átona. La transcripción fonética de ambas cadenas, por tanto, también es distinta: [sino] vs. [sino]. Son falsos homónimos, como "hasta" y "asta" son falsos homófonos.

En la lengua escrita podrían confundirse, ciertamente, los pares "el" / "él", "de" / "dé", "mi" / "mí", "tu" / "tú", "si" / "sí", etc., razón por la cual se arbitró colocar la tilde sobre la vocal en los términos tónicos, incluso contraviniendo las reglas generales; pero se trata de una norma ortográfica, válida sólo en la lengua escrita: tales parejas nunca se confundirían en la lengua hablada. En la lengua escrita, hacemos que "recaiga" la tilde sobre una vocal u otra; en la lengua

hablada, en cambio, no hacemos "recaer" nada sobre la vocal. En resumen, sólo una perspectiva lecto-escritora de la lengua permite considerar el acento propio de las palabras como un elemento superpuesto.

3.1.2.3. El acento propio de la palabra

Hemos dicho que las palabras átonas no gozan de ninguna independencia fónica en el habla y, por tanto, sólo podemos referirnos a ellas como "palabras" desde un punto de vista estrictamente gramatical y muy abstracto. Por ejemplo, las personas no alfabetizadas no las conciben independientemente, y los analfabetos funcionales las escriben sistemáticamente pegadas a la siguiente palabras tónica: de ahí vulgarismos como "amoto" o "arradio", en los que el artículo ha acabado por integrarse parcialmente en el sustantivo.

Y es que el artículo, palabra átona por excelencia, en algunas lenguas se escribe adosado al sustantivo al que acompaña (como en euskera, que se pospone al sustantivo como un sufijo); en nuestra lengua también ocurre algo parecido, y algunos autores, entre ellos Alarcos Llorach, cuestionan su independencia gramatical. Este autor postula, en la última gramática (*quasi*)normativa de la Real Academia, su carácter estrictamente morfológico, semejante al género o al número: "aunque el artículo precede en la secuencia al sustantivo y en la escritura se mantiene separado por un blanco, al revés de los signos que manifiestan el género y el número, que se posponen y se juntan, es también como estos un accidente del sustantivo" (Alarcos, 1994: 66). Es decir, la independencia gráfica que se le concede en la escritura no es suficiente para negar su carácter realmente dependiente de una palabra tónica.

Desde este punto de vista, el paso histórico de la flexión de caso al empleo de preposiciones como indicadores categoriales no podía ser menos

traumático: ocurrió en todas las lenguas romances desde el latín, y ocurre hoy en las lenguas germánicas y eslavas que conservan dicha flexión sólo en la lengua culta. Dicho proceso consiste, en términos estrictamente fonéticos, en el cambio de varios sufijos (las desinencias de cada flexión) por un prefijo (la preposición). Puesto que antes de la pérdida definitiva de la flexión hay un largo periodo de convivencia de la preposición y la desinencia, para los hablantes de la lengua en cuestión la supresión de la desinencia redundante se resuelve con una mera indistinción fónica de la parte final de la palabra. Como es bien sabido, la posición postónica siempre ha sido considerada menos informativa que la posición pretónica (por ejemplo, y volviendo al tema del vocalismo andaluz, H López Morales, 1984, considera que la pérdida de la -s- final de los sustantivos no supone ninguna ambigüedad morfológica, porque ya el artículo marca suficientemente su pluralidad).

Las palabras átonas, pues, muy bien podrían considerarse no verdaderas palabras sino elementos dependientes de las palabras tónicas: elementos gramaticales como los artículos, las preposiciones, las conjunciones, los relativos y las formas átonas de los pronombres; o elementos léxicos como las fórmulas de tratamiento ("don", "fray", "san") y la primera parte de los compuestos ("trescientos", "Juan José"), además de los prefijos y los sufijos derivativos (aunque, en este caso, escritos sin espacio en blanco).

Por el contrario, las palabras tónicas son palabras auténticas, independientes fónicamente. El acento, pues, es un constituyente básico de la palabra, del significante de la palabra.

Entendida esta como signo lingüístico, como la unión de un significado conceptual (el contenido léxico) y un significante fónico (los sonidos que la

constituyen), diríamos que la característica fundamental del significado de la palabra es que se trata de un concepto independiente con el cual puede operar el pensamiento; y que la característica principal del significante, por su parte, sería una característica fónica: los sonidos que constituyen la palabra han de estar agrupados en torno a un acento que les confiera independencia fónica en el discurso, y que les permita llevar asociados elementos gramaticales o léxicos átonos (morfemas, conectores y elementos derivativos).

De este modo, el acento no sólo no puede concebirse como un elemento "superpuesto" a la palabra, sino que ha de entenderse como una parte fundamental de la naturaleza fónica de la palabra, sin el cual el concepto léxico no podría tener un referente definido e independiente en el enunciado, y sin el cual, por tanto, no podría actuar de núcleo informativo al que se asocian otros elementos gramaticales o derivativos. Canellada & Kulhmann lo explican del siguiente modo: "el vocablo, igual que lleva su distribución de fonemas, tiene su estructura acentual propia, y para ello lleva una sílaba marcada frente a una átona o no marcada. Esas son sus señas de identidad" (1987: 75).

Al "acento propio de la palabra" lo llamamos *acento paradigmático*, pues forma parte de la misma de igual modo que los segmentos de timbre, y no sólo dentro del discurso: el *acento paradigmático* no es determinado por el contexto discursivo, ni por la categoría gramatical de la palabra, ni por reglas *a posteriori*, sino que forma parte de la palabra *ab initio*, y es una información no "predecible" mediante reglas sino "dada".

Cada palabra léxica, por tanto, está constituida por una serie de segmentos tímbricos, uno de los cuales (un segmento vocálico) es tónico: llamamos *acento paradigmático* a la tonicidad de dicha vocal. Dentro de la estructura fónica de la palabra, por tanto, la cualidad tímbrica de sus segmentos no es separable de la

tonicidad de la vocal tónica.

3.1.3. *Constituyentes del ritmo en español*

3.1.3.1. La palabra fónica

Alrededor del acento paradigmático, pues, no sólo se agrupan los sonidos que constituyen el significante de la palabra léxica, sino aquellos otros que constituyen las palabras átonas asociadas a esta en el discurso. Al conjunto de los sonidos agrupados en torno a un acento paradigmático lo llamamos *palabra fónica* o *grupo rítmico*.

En ocasiones, la palabra léxica y la *palabra fónica* pueden coincidir plenamente, como en la cadena: "faltan dependientes experimentados"; sin embargo, es más común que no coincidan: en la cadena "faltan dependientes con experiencia", la preposición más el sustantivo "con experiencia" constituyen una única *palabra fónica* o *grupo rítmico*, y se transcribiría junto [konesperjéneja].

En cualquier caso, palabra léxica y palabra fónica son dos conceptos que pertenecen a niveles de análisis distintos, no equiparables: la palabra léxica es una unidad léxica y gramatical, se define en función de un significado conceptual asociado sistemáticamente a un significante fonológico (en el que ha de haber un acento paradigmático), también en función de su rendimiento gramatical, y se concibe como una unidad abstracta (que puede ser pronunciada, por ejemplo, pero también escrita); la palabra fónica, en cambio, es una unidad fonética, se define únicamente en función de un acento paradigmático alrededor del cual se agrupan los sonidos que la constituyen, y su existencia, por tanto, es rigurosamente concreta y material (sonora).

Los términos *palabra fónica* o *grupo rítmico* que empleamos aquí equivalen a lo que Navaro Tomás llamaba, distinguiendo entre los acentos de intensidad y de tono, "grupo de intensidad" o "grupo tónico" (1918), que años después llamaría "grupo rítmico-semántico": "parte del discurso que tiene por base prosódica un solo acento espiratorio y por contenido ideológico un núcleo de significación no susceptible de divisiones más pequeñas" (1948, 29). Esta designación, ciertamente, resulta poco afortunada porque se mezclan en ella los dos niveles de análisis a que nos hemos referido antes; con todo, en la práctica la palabra fónica siempre coincide con grupo semánticos o funcionales definidos (artículo + sustantivo o adjetivo, posesivo + sustantivo, pronombre átono + verbo, etc., como en: "la casa", "el último", "nuestra habitación", "te quiero", etc.). Este hecho de que las palabras fónicas coincidan en los discursos concretos con grupos funcionales o semánticos, sin embargo, no debe confundirnos sobre la naturaleza esencialmente distinta de ambos niveles de análisis: uno abstracto y funcional, el otro concreto y sensible.

Como veremos, dicha coincidencia entre las palabras fónicas y los núcleos semánticos del discurso es un mecanismo de integración fónica de los sonidos que componen la cadena sonora, un modo de presentarlos al oyente en bloques integrados, y de posibilitar así la comprensión del discurso. El otro modo de hacerlo es el grupo fónico (v. apartado 3.3.2.).

Siguiendo la idea del "grupo de intensidad" de Navarro Tomás, Canellada & Kuhlmann, más recientemente, hablan de "cláusula": "cada altura en la línea de intensidad marca el comienzo de una unidad que llamamos cláusula y que señalamos entre barras. Esta cláusula se extiende hasta que, tras un mínimo, una nueva altura dé comienzo a otra unidad" (1987: 84).

En su trabajo sobre el ritmo en español, por su parte, G. A. Toledo habla

de "pie acentual" en español, unidad que comprende "desde el inicio de la vocal acentuada, hasta el inicio del próximo segmento vocálico con acento" (1988: 73); sin embargo, una vez establecida la poca eficacia de dicha unidad para medir el ritmo en español, opta por considerar otra unidad, el "grupo de acento", que coincide con lo que aquí hemos llamado "palabra fónica" o "grupo rítmico": "esta unidad...está integrada por una palabra con acento primario -generalmente, las palabras semánticamente llenas- y palabras con acento no primario, es decir, las palabras funcionales, vacías de sentido. Estos últimos segmentos integran el grupo de acento en posiciones tanto enclítica como proclítica con respecto a la palabra, con el foco semántico señalado en el plano fónico por la relevancia acentual" (op.cit.: 155). Al "grupo de acento", finalmente, lo llama "grupo rítmico" en trabajos posteriores (Toledo, 1994).

Dado su carácter fónico y concreto, no puede predecirse con precisión el límite de los grupos rítmicos o palabras fónicas de un texto escrito: es decir, se trata de una unidad *de facto*, no de una unidad abstracta. Esto es así porque la palabra fónica no es una unidad significativa *en el paradigma*, esto es, no es un signo lingüístico. Dado un texto escrito, en realidad, en él no hay ninguna palabra fónica, porque en la escritura no hay sonidos, ni acentos, etc. Dado un texto oral, en cambio, sí que hay palabras fónicas: cuyos límites no son necesariamente los que nosotros pudiéramos predecir siguiendo ciertas reglas, sino los que hay porque el hablante lo ha pronunciado así. La palabra fónica o grupo rítmico, por tanto, es una "información dada", por naturaleza.

Así, en el enunciado "la casa estaba vacía y triste", los grupos rítmicos podrían ser: [lakása] - [estáβa] - [βaθía] - [itríste]; o bien: [lakásaes] - [táβa] - [βaθíaaj] - [tríste]. Toledo (1944: 193) ofrece el siguiente ejemplo, muy ilustrativo al respecto: el enunciado "bueno, ese es mi primo" fue pronunciado

por su informante venezolana como [bwenwé] - [sés] - [miprímo]. (Podía haberse pronunciado, también, como [bwéno] - [éses] - [miprímo] o como [bwenwéses] - [miprímo], etc.).

Como en el caso de la sílaba, lo de menos es la limitación precisa de los márgenes, sino saber cuál es su núcleo: en cualquier caso, pues, el núcleo de la palabra fónica es un acento paradigmático. Ciertamente, en las lenguas de acento fijo (de acento paradigmático fijo), la palabra fónica estará siempre mejor definida: en francés, por ejemplo, sabemos que los grupos rítmicos serán preferentemente oxítonos.

En nuestra lengua, la tendencia a tener palabras llanas también puede orientarnos: las palabras fónicas, pues, serán preferentemente llanas, esto es, seguramente el acento paradigmático será seguido, como mucho, de una vocal átona; esto implica, finalmente, que los elementos átonos del discurso tenderán a juntarse con el acento paradigmático siguiente, y no con el anterior, para formar palabras llanas y no esdrújulas (una excepción serían los pronombres personales integrados en el verbo como sufijos y que dan lugar a palabras proparoxítonas o superproparoxítonas: "arreglárselas", "abotónatelo", etc., siempre que no estén seguidas por otras palabras).

3.1.3.2. Isocronía *léxica*

Tradicionalmente, se ha dicho que hay dos tipos de esquemas rítmicos: la "isocronía acentual" (*stress-timing*) y la "isocronía silábica" (*syllable-timing*). La isocronía acentual se define como la tendencia a que la distancia temporal entre acento y acento (el pie acentual) sea similar, lo cual conlleva un efecto de "compresión silábica" en los casos de pies acentuales muy desiguales: los más largos, al durar aproximadamente lo mismo que otros más cortos, estarían

compuestos por sílabas más breves.

La isocronía silábica, en cambio, sería la tendencia a mantener una duración similar en cada sílaba, despreciando el pie acentual como unidad de medida: un grupo rítmico más largo, pues, durará proporcionalmente más que uno corto, en función del número de sílabas de cada uno.

De este modo, fenómenos como el de la reducción vocálica en inglés se explicarían por la isocronía acentual característica de esta lengua: el timbre de las vocales inglesas sería, debido al condicionamiento rítmico, más inestable que en las lenguas de isocronía silábica, como las lenguas romances. Así, ha sido una constante de la fonología en lengua inglesa establecer reglas de acentuación fundamentadas en el timbre de las vocales: las vocales plenas pueden ser tónicas, las reducidas no, etc. (Bolinger, 1964; Halliday, 1967, Lieberman, 1967). En realidad, y como veremos, la reducción vocálica es un fenómeno muy común también entre las lenguas romances de isocronía silábica, como el catalán o el portugués.

Con todo, los dos esquemas rítmicos (isocronía acentual y silábica) han servido de base para la elaboración de reglas métricas, tanto reglas fonológicas como reglas poético-preceptivas: las reglas métricas de Liberman & Prince (1977), así, sólo tienen sentido en una lengua de isocronía acentual, como el inglés, y serían difícilmente exportables; del mismo modo, las reglas de la métrica española sólo sirven a una lengua de isocronía silábica, como el español, y no pueden dar cuenta, por ejemplo, de los esquemas de versificación ingleses.

En el verso, en efecto, el ritmo es el fenómeno tal vez más significativo, hasta el punto de convertirse, en este siglo, en la única marca formal de la poesía moderna (en los versos *blancos*, sin rima, y sobre todo en los versos *libres*, sin un "metro" tradicional, pero sí con un ritmo recurrente muy marcado,

que viene a ser como un esquema métrico cada vez nuevo y distinto). Por su parte, en la prosa literaria el ritmo también juega un papel significativo (Navarro Tomás, 1944, hace un análisis muy detallado del fenómeno), aunque no tanto como para dar pie a una disciplina específica, como la métrica poética.

Sin embargo, en el habla espontánea el ritmo no parece cumplir ninguna función significativa, y no pasa de ser un fenómeno coarticulatorio. Es decir, el ritmo es un fenómeno que tiene como función exclusiva la integración de los sonidos en bloques (los grupos rítmicos) para posibilitar, como veremos en el apartado 3.3.1., la comprensión auditiva del discurso. A esta función se le añade, tal vez, la de caracterizar el registro poético de la lengua.

Curiosamente, la métrica tradicional puede dar respuesta a algunos fenómenos rítmicos que podrían resultar sorprendentes. Así, Toledo (1988 y 1994) explica que diversos estudios experimentales han puesto en duda la integridad de la isocronía silábica en español: por ejemplo, no siempre se observa una duración similar entre sílabas situadas en posiciones diferentes; por otra parte, en ocasiones se observan fenómenos de compresión silábica característicos de la isocronía acentual. Para Toledo, la manera de afrontar estos hechos es considerar un "grupo rítmico" en español susceptible de tres formas distintas (grupos oxítonos, grupos paroxítonos y grupos proparoxítonos) en lugar de considerar como unidad rítmica el pie acentual. De este modo, los fenómenos de compresión silábica observados en sílabas postónicas, por ejemplo, coinciden con el conteo de la métrica tradicional: si el verso acaba en palabra aguda (la sílaba tónica es la última), se cuentan no una sino dos sílabas; si acaba en palabra esdrújula (la sílaba tónica seguida de otras dos), se cuentan no tres sino dos sílabas.

Estas reglas abonan la tendencia paroxítona de la lengua española, la consagran como un fenómeno rítmico (de palabras fónicas) y coinciden con los mismos datos experimentales que en ocasiones han sido interpretados como una cierta tendencia a la isocronía acentual (y no silábica) en español.

En cualquier caso, y una vez determinado que "la sílaba en español ... no sería estrictamente isocrónica" (Toledo, 1994), parece obvio que la sílaba, en sí misma, no puede considerarse una unidad rítmica independiente, sino incluida en el grupo rítmico o palabra fónica, que sería la verdadera unidad rítmica en español: así, una cierta isocronía silábica en realidad sería compatible con una cierta isocronía acentual, entendiendo la palabra fónica en su conjunto como la unidad que tendería a igualarse temporalmente.

Es decir, hablar de isocronía acentual o silábica no es suficiente en español, pues la unidad rítmica no es ni el pie acentual (entendido como el trecho que va de acento a acento) ni la sílaba: la unidad rítmica en la que la isocronía actúa en español es la palabra fónica, y por tanto hay que considerar que en nuestra lengua actúa un esquema rítmico distinto, que podríamos llamar *isocronía de palabra fónica o isocronía léxica*.

3.1.3.3. Conclusión: constituyentes del ritmo

En resumen, el ritmo en español está constituido por la sucesión de palabras fónicas, en las que se engloban los sonidos agrupados alrededor (antes y después, con una acusada tendencia al esquema paroxítono) de un acento paradigmático.

En la figura 39 se muestra un esquema de los constituyentes del ritmo, comenzando por el juego de vocales y consonantes.

Las vocales son el centro de la fonación del sonido y, por tanto, los

constituyentes fónicos fundamentales del discurso: la unidad sílaba no tiene más entidad que la otorgada por la vocal como núcleo silábico. Las consonantes, así, constituyen una suerte de "zona marginal", en términos silábicos, en la que no importan los límites entre ellas, sino la proximidad de los núcleos (las vocales), en los que reside la F_0 , esto es, toda la información tonal relevante: el acento paradigmático y el valor de la entonación en ese punto de la cadena.

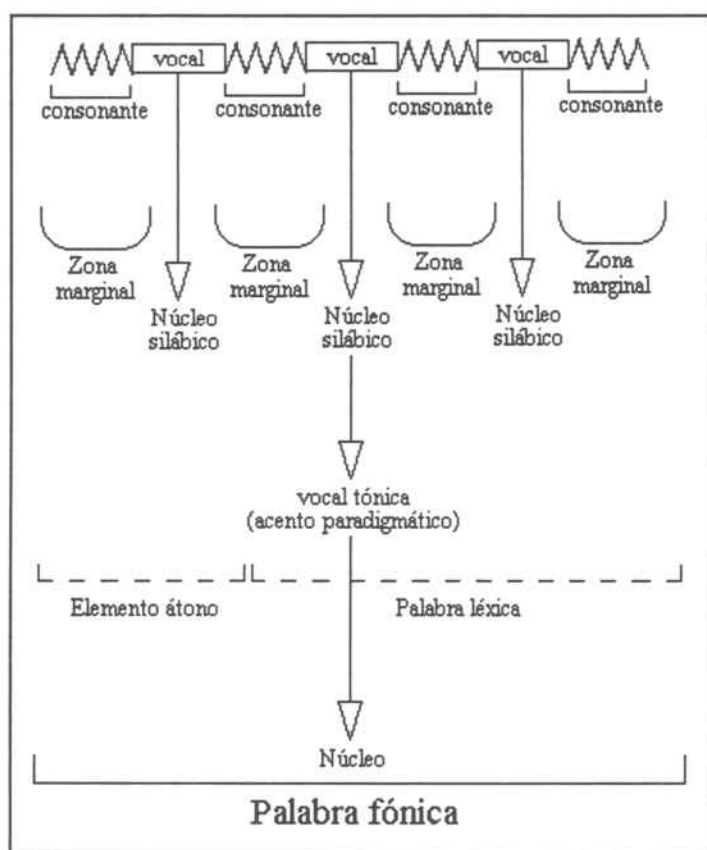


Figura 39 - Constituyentes del ritmo.

La vocal tónica, el acento paradigmático, es el núcleo fónico de la palabra léxica, que no es una unidad fónica sino léxico-gramatical cuyo significado ha de contar con dicho acento para tener una entidad independiente, es decir, para que se le puedan asociar elementos gramaticales y léxicos átonos; el elemento átono, a su vez, puede ser un artículo, una preposición, un prefijo, etc., de una o de dos sílabas, o varios elementos seguidos

(como en el grupo "la de nuestra casa": [laðenwestrakása]).

Finalmente, la unidad rítmica es la palabra fónica o grupo rítmico, cuyo núcleo es la vocal tónica, el acento paradigmático propio de la palabra léxica; los sonidos se agrupan alrededor de dicha vocal, esto es, se perciben formando

un bloque del cual la vocal tónica es el centro. La palabra fónica es la única realidad fonética del concepto "palabra": las palabras léxicas sólo existen, realmente, o bien formando independientemente una palabra fónica, o bien formándola junto a algún elemento átono.

Los límites fonéticos precisos de la palabra fónica no son determinantes, como no lo son para la sílaba: el oyente percibe el bloque tónico y lo identifica con las unidades léxico-gramaticales que conoce; por tanto, no hace falta ninguna precisión acústica, pues ya los fenómenos de redundancia fónica y la propia proyección psicológica del oyente suplen la precisión en la segmentación de las unidades percibidas, las posibles ambigüedades y, sobre todo, las condiciones de ruido del contexto situacional. La falta de precisión en la percepción auditiva del discurso constituye, indudablemente, una garantía de éxito frente a las condiciones auditivas adversas de todo tipo: si fuera necesaria una mayor precisión acústica para descifrar los mensajes percibidos auditivamente, las situaciones en las que un proceso comunicativo oral podría tener éxito se verían drásticamente reducidas.

Así, ya en la propia naturaleza fónica de la lengua se observa esta imprecisión acústica, y se organizan los sonidos no en bloques perfectamente delimitados por algún índice acústico (o juntura), que necesariamente debería figurar en la zona marginal de la fonación (entre las consonantes, que no son producto de la fonación y, por tanto, la información tonal que ofrecen es muy escasa), sino en bloques perfectamente definidos por su núcleo: en bloques culminativos.

El esquema de la figura 39 podría corresponder a cualquiera de las siguientes palabras fónicas: "la casa", "las casas", "mi tierra", "montañas", "por siempre", etc.

La recurrencia de palabras fónicas en el discurso (entre las que puede observarse una cierta isocronía "léxica") constituye su esquema rítmico. El ritmo, en definitiva, no constituye ningún tipo de significante, ni de signo, excepto, tal vez, en el registro poético.

3.2. El acento como fenómeno segmental

Hemos visto que el acento actúa directamente sobre las vocales, núcleos silábicos y, en realidad, núcleos de toda la fonación. Hemos visto, también, que el acento paradigmático es el acento propio de la palabra, constituyente de su naturaleza fónica igual que los rasgos tímbricos de cada uno de los sonidos que la forman. El acento, pues, no es un fenómeno que actúe *a posteriori* sobre las palabras, ni es un fenómeno que "recaiga" sobre las vocales.

No es de extrañar, pues, que nos planteemos ahora el carácter segmental del acento: en primer lugar, revisaremos el parámetro de la duración para determinar su naturaleza fonológica en nuestra lengua y examinar cómo un fenómeno tradicionalmente considerado prosódico y contrastivo puede concebirse también como un fenómeno segmental y opositivo; luego revisaremos el fenómeno de la reducción vocálica, como un medio para relacionar la estructura tímbrica del sonido con su carácter tónico o átono; finalmente, consideraremos las implicaciones fonológicas que tales fenómenos pueden tener en nuestra lengua.

3.2.1. El parámetro de la duración

3.2.1.1. Parámetro de segundo grado

En el apartado 1.2.3. definimos la cantidad como un parámetro *de segundo grado*, que no está al mismo nivel que las características primarias del sonido (tono, intensidad y timbre), sino que sirve justamente para medirlas: es una característica de las características del sonido.

La duración se emplea en el análisis fonético para medir la señal sonora

en su conjunto, o bien para medir un segmento tímbrico. En este último caso, la duración del segmento se establece, generalmente, midiendo uno de sus rasgos tímbricos, el más estable o el más definitorio: por ejemplo, para medir una vocal emitida entre consonantes sordas podemos medir la barra de sonoridad; si está emitida entre consonantes sonoras, podemos tomar el segundo formante; entre vocales y formando un diptongo, la energía de los formantes superiores puede ser un buen índice, o el momento en el que la transición se hace evidente; etc.

Este proceso de medición, como vimos, no es comparable a la medición de los demás parámetros del sonido, porque requiere una abstracción previa: una abstracción intermedia entre el propio sonido concreto y el fonema, abstracción a la que llamamos "sonido del lenguaje" o "fono"

En realidad, no siempre somos conscientes de efectuar dicha abstracción: los tratados de fonética articulatoria, por ejemplo, tratan la producción de los sonidos (los alófonos más comunes de los fonemas) como si se tratara de unidades aisladas e independientes. A menudo, incluso se habla de la "articulación" de los fonemas (Quilis & Fernández, 1964): no en vano, los rasgos fonológicos que emplea la mayoría de los autores (entre ellos, Alarcos, 1950, y Quilis, 1993) para caracterizar los fonemas del castellano son rasgos fonéticos, que además pretenden ser descriptivos, como si los fonemas no fueran abstracciones y pudieran "sonar". Por lo demás, incluso en los enfoques teóricos en los que no se usa el concepto de "fonema" sigue empleándose la abstracción "sonido del lenguaje".

Los modernos instrumentos de análisis digital del sonido (v. 1.3.3.) pueden extraer todos los parámetros acústicos del mismo y suelen ofrecer los datos de un modo gráfico (oscilograma, sonograma, espectro, etc.), pero también numérico: en la pantalla aparecen, por ejemplo, los valores seleccionados por el

cursor en el gráfico. En realidad, y frente a los instrumentos analógicos, en los que la señal sonora se "impresionaba" directamente en el papel, en los instrumentos digitales el gráfico mismo es una forma de representar visualmente los valores numéricos extraídos de la señal sonora, y por eso pueden modificarse los colores, los formatos de presentación, etc.

Uno de los valores que se ofrecen es el número de "trama" (*frame*) en el que se ha fijado el cursor, información que puede traducirse en términos temporales: duración desde el inicio de la señal hasta esa trama (el inicio de la señal, además, es el inicio de la grabación, no el inicio del sonido). Sin embargo, para determinar la duración de una vocal, por ejemplo, el analista ha de seleccionar manualmente la porción de señal que corresponde a la vocal, pues el instrumento es incapaz de detectar la unidad "vocal" (que es una unidad abstracta: un fono).

Es decir, un instrumento puede medir el tono, la intensidad, y la estructura tímbrica del sonido, que son sus características acústicas, pero es incapaz de medir la cantidad porque no puede separar cada segmento. Como hemos dicho, el sonido es un continuum de valores fluctuantes; los "sonidos del lenguaje", por el contrario, son abstracciones que concebimos como porciones de sonido estable, basadas en los rasgos del timbre (y ni siquiera en todos, sólo en los pocos que son relevantes fonológicamente: los rasgos inherentes, distintivos o redundantes; la realidad tímbrica del sonido, sin embargo, es mucho más rica y compleja).

En resumen, la duración, como parámetro acústico, ha de ser referida a un segmento, es decir, ha de mediar un proceso de abstracción que permita concebir el segmento como una unidad independiente; como decíamos en el apartado 1.2.3., "la duración debería considerarse, directamente, como un rasgo

fonológico y dependiente sólo del nivel fonológico".

3.2.1.2. La cantidad como rasgo inherente

Puesto que la cantidad es un rasgo referido al segmento ya abstraído como tal, quizá haya que considerarla, pues, como un rasgo no suprasegmental sino intrínseco del segmento. Es decir, como un rasgo opositivo y no contrastivo. Efecto, en algunas lenguas se oponen las vocales largas a las vocales breves, o las consonantes largas a las consonantes breves. Por ejemplo, en latín clásico el sistema vocálico se desdoblaba en vocales largas y breves, también en holandés moderno hay desdoblamiento de vocales largas y breves, en inglés y en alemán la duración de las vocales es relevante fonológicamente, etc. Del mismo modo, en otras lenguas la duración es un rasgo distintivo en el sistema consonántico.

Ciertamente, el rendimiento lingüístico del rasgo de duración puede considerarse muy escaso: en general, las vocales largas son también vocales cerradas, y las vocales breves son vocales abiertas; en la mayoría de los casos, la abertura podrá considerarse como el rasgo distintivo y la duración como el redundante. Eso es lo que ocurrió, por ejemplo, en latín vulgar: el rasgo de duración se hizo redundante frente al rasgo de abertura, que se convirtió en el rasgo distintivo, lo cual provocó la confluencia entre la /i/ breve (abierta) y la /e/ larga (cerrada), entre la /a/ larga y la /a/ breve, y entre la /o/ larga (cerrada) y la /u/ breve (abierta). Así, frente al sistema desdoblado del latín clásico, con cinco grados de abertura y la duración como rasgo distintivo, el resultado fue el de un sistema único con siete grados de abertura y en el que la duración se convirtió en un rasgo redundante.

También en inglés coinciden la brevedad de la vocal con su abertura y el rasgo de la duración podría considerarse, pues, redundante. En otras lenguas, sin

embargo, la duración aparece junto a otros rasgos: en alemán, por ejemplo, las vocales largas son tensas y las breves son laxas.

Puede distinguirse entre una duración intrínseca, propia de la naturaleza del segmento, y una duración extrínseca, dependiente del contacto con otros segmentos o de la posición silábica. Así en latín, cuyas reglas métricas establecían la duración de la vocal según su posición (Martínez Celadrán, 1984: 373), o en alemán, donde la posición de la vocal, en sílaba abierta o trabada, determina su longitud relativa. En castellano, según Monroy (1980: 37), la duración de la vocal no está en relación directa con la consonante que sigue, como afirmaba Navarro Tomás (1916); por otra parte, Monroy niega que la naturaleza silábica (sílaba libre / sílaba trabada) afecte en castellano a la duración de las vocales.

Un caso aparte lo constituyen los sonidos "geminados": dos sonidos iguales seguidos, pero pertenecientes a sílabas distintas, como las vocales de "proveer", "está aquí", etc.; y las consonantes de "innato", "el lago", etc. En principio, habría que distinguir entre estos casos de geminación contextual y aquellos otros en los que las unidades alargadas puedan considerarse fonemáticas. Así, por ejemplo, Recasens habla de la geminación en catalán como de un fenómeno fonético frecuente, pero no considera que pueda hablarse de fonemas consonánticos largos, ni de un "proceso fonológico de alargamiento" (1993: 61), sino de la suma de dos realizaciones simples.

Por su parte, Malmberg (1974) aduce que las consonantes largas no deben confundirse con las geminadas, pues estas últimas están separadas por una frontera silábica. Sin embargo, para Martínez Celadrán no hay tal frontera, según se desprende del análisis acústico; comentando el sonograma en el que aparecen las palabras italianas *aprendo* / *apprendo*, dice el autor: "la primera (p) posee

una duración de 83 mls y la segunda de 151, casi el doble; no obstante, si hubiera dos *p*, como sugiere de alguna manera la geminación, existirían dos barras de explosión y sólo vemos una, por lo que se debe hablar de cantidad y no de geminación. Lo que suele ocurrir es que muchos autores a las vocales las llaman largas y a las consonantes geminadas" (1984: 248).

Y es que, como vimos en el apartado anterior, los límites silábicos no tienen correlato acústico, la sílaba es definida únicamente en función de la vocal (que en español es su núcleo o "centro silábico") y la zona consonántica es una zona marginal; por tanto, a efectos silábicos una geminación de dos consonantes equivale al alargamiento de una sola. Sin embargo, si a efectos silábicos no importa la diferencia, y si en términos acústicos en realidad se trata de una misma cosa, para una interpretación fonológica la diferencia es notable, pues hablar de geminación o alargamiento puede equivaler a concebir una unidad fonológica (repetida) o dos (pues se opone una larga a otra breve).

La perspectiva de considerar la duración como un rasgo inherente del segmento, por lo demás, cuenta con numerosos antecedentes: en las descripciones fonológicas del holandés o del alemán, por ejemplo, la duración suele considerarse como un rasgo segmental inherente (Cohen, 1972; Philipp, 1970); y los mismos Chomsky & Halle (1968) proponen el rasgo de longitud [+/- largo] como rasgo universal, incluido en un principio en el bloque de "rasgos prosódicos", como el acento, aunque empleado inequívocamente como rasgo segmental en todas las reglas fonológicas en las que aparece posteriormente: por ejemplo, en la "regla de acortamiento de grupo", que reza $V \rightarrow [-\text{largo}] / ______ C_0 \{CC\}$ (Durand, 1990: 127). En nuestra lengua, finalmente, hay el precedente de E. Lorenzo (1972), que propone la duración

como rasgo distintivo en español. Canellada & Kuhlmann, por su parte, insisten en que "es la duración lo que marca el acento" (1987: 68), por encima de los demás parámetros.

3.2.1.3. El rasgo inherente de tensión

En fonología segmental tenemos otro ejemplo muy llamativo de fenómeno prosódico convertido, al menos parcialmente, en un rasgo segmental inherente: el rasgo *tenso / laxo*.

Ya Trubetzkoy tomaba en consideración dos correlaciones fonológicas que afectaban al sistema consonántico, y en concreto al obstáculo articulatorio que da lugar a la consonante: la "correlación de tensión" y la "correlación de intensidad". La primera se refiere a "la oposición entre «fuertes» y «suaves», en la que la fuerza del obstáculo está en proporción con la del medio empleado para franquearlo (presión del aire): si el obstáculo está reforzado por la tensión de la musculatura de la boca, la presión del aire es más fuerte; si, por el contrario, los músculos de los órganos bucales se relajan, también se debilita la presión del aire" (1939: 140-141).

En cuanto a la "correlación de intensidad" (o "de presión"): "presenta una relación algo diferente entre la solidez del obstáculo y la fuerza de la presión del aire: cuando los músculos de los órganos bucales están relajados la presión resulta demasiado fuerte, de allí el acortamiento y la eventual aspiración de los miembros «débiles» de la oposición; cuando la musculatura bucal está tensa la presión de aire resulta ser exactamente la requerida para desarrollar su labor, de allí el largo relativo, la falta de aspiración y el franqueamiento dificultoso del obstáculo en los miembros «fuertes» de la oposición" (op.cit.: 141).

En realidad, las consideraciones de Trubetzkoy no tienen que ver con la

intensidad como fenómeno suprasegmental, pues si bien hace referencia a la presión del aire y a la tensión muscular, lo cierto es que se centra en la relación entre estos fenómenos y la tensión muscular de la boca: por tanto, parece hacer referencia, más bien, a la "precisión" en la articulación del sonido consonántico y, sobre todo, a la tensión de la propia articulación bucal, no de la fonación (lo cual afecta por igual a las consonantes sonoras, producto de fonación y articulación, como a las sordas, en las que no hay fonación).

Jakobson & Halle proponen el rasgo "tenso / flojo" como un rasgo inherente de sonoridad y lo definen del siguiente modo: "acústicamente - cantidad total de energía más elevada (o en cambio más baja) junto con mayor (o menor) difusión de la energía en el espectrograma y en el tiempo; genéticamente - mayor (o menor) deformación del sistema de formación con respecto a su posición de reposo" y añaden que "el papel de la tensión muscular de la lengua, las paredes del canal bucal y la glotis requiere un estudio más detenido" (1956: 54). En estos autores, por tanto, se ha eliminado por completo la referencia que hacía Trubetzkoy a la "presión del aire", que en su definición era un elemento fundamental.

Por su parte, Chomsky & Halle definen así el rasgo: "el rasgo «tensión» especifica cómo la musculatura supraglotal ejecuta el movimiento articulatorio completo de un sonido dado. Los sonidos tensos se producen con un movimiento deliberado, preciso, con una distinción máxima, y necesitan un considerable esfuerzo muscular; los sonidos no tensos se producen rápidamente y con cierta falta de distinción. En los sonidos tensos, tanto vocales como consonantes, el periodo durante el cual los órganos articulatorios mantienen la configuración adecuada es relativamente largo, mientras que en los sonidos tensos se ejecuta la totalidad del movimiento de una forma bastante superficial"

(1968: 217).

Sin duda, esta última definición es lo suficientemente explícita, y queda claro cómo el rasgo tiene que ver con la precisión y con la intensidad de la propia articulación (no de la fonación). Los autores, además, hacen hincapié en la duración relativa de los sonidos tensos. Pero tampoco hacen ninguna referencia a la "presión del aire".

Podríamos pensar que Trubetzkoy fue, simplemente, inexacto en sus apreciaciones, y que la intensidad de la presión del aire (directamente proporcional a la intensidad de la fonación del sonido, como sabemos) no tiene nada que ver con la precisión y la intensidad de la propia articulación: de este modo, Jakobson & Halle y Chomsky & Halle ponen las cosas en su sitio y evitan toda posible confusión entre el fenómeno prosódico de intensidad y el rasgo tímbrico de tensión. En cualquier caso, esto implica que no aparezca con claridad a qué fenómeno concreto hace referencia el rasgo: por ejemplo, Recasens afirma que "algunos fonólogos postulan la existencia de rasgos fonológicos sin correlatos fonéticos claros -por ejemplo, «tensión»-" (1993: 48).

Sin embargo, Martínez Celdrán nos aporta nuevos datos: "las vocales tensas se distancian más de la posición neutral que las relajadas, además de ser más distintivas y de duración mayor. Lo mismo sucede con las consonantes. En las obstruyentes, la sonorización aparece mucho más fácilmente en las no tensas que en las tensas, al estar las paredes del aparato fonador más relajadas y las cuerdas vocales en la posición adecuada para ponerse en vibración" (1984: 240). Es decir, el rasgo de tensión también afecta a la fonación, al aparato fonador en su conjunto y a las propias cuerdas vocales, hasta el punto de facilitar la sonoridad (o, lo que es lo mismo, la vibración de las cuerdas vocales: la fonación de sonido genuino). Esa es la razón por la que en castellano, por

ejemplo, "tensión" y "sonoridad" sean rasgos paralelos, y uno de los dos deba considerarse redundante.

En definitiva, si bien no puede decirse que no haya "correlatos fonéticos claros" (como apunta Recasens), sí vemos que no es fácil distinguir claramente el rasgo de "tensión" del fenómeno de "intensidad": desde un punto de vista estrictamente acústico, la "tensión" de un sonido sonoro depende de la "intensidad" de sus formantes, que es la misma "intensidad" considerada fenómeno suprasegmental, pero referida a un único segmento. Así, podríamos concluir que tal vez se trate del mismo fenómeno entendido desde una perspectiva segmental (tensión) o suprasegmental (intensidad): referido, pues, bien a la articulación/fonación del sonido (tensión), bien a la fonación de la cadena sonora en su conjunto (intensidad). Desde la perspectiva suprasegmental, como sabemos, la intensidad apenas tiene relevancia lingüística.

Lo mismo podría decirse de la duración: hay una duración que puede entenderse como un rasgo inherente al segmento, y hay una duración total de la cadena hablada, que apenas aporta información lingüística. En cualquier caso, la duración como parámetro acentual deberá entenderse, desde esta perspectiva, como un fenómeno segmental.

3.2.1.4. La mora

Si consideramos la duración como un rasgo segmental inherente, por ejemplo en los sistemas vocálicos desdoblados, la cantidad vocálica no puede ser un valor relativo, de contraste vocal larga / vocal breve, sino que ha de tener un valor absoluto determinado que forme parte de la propia naturaleza segmental del sonido.

El concepto de *mora* nos permite asignar un valor relativo a la duración

del sonido cuando se trata de un rasgo distintivo: definimos la vocal larga como la que tiene dos moras, y la vocal breve como la que tiene una sola mora. Así, para Trubetzkoy las lenguas en las que la duración es un rasgo fonológico y, por tanto, se distinguen sonidos largos y breves son "lenguas que cuentan las moras", frente a las "lenguas que cuentan las sílabas" (1939: 175).

La mora, efectivamente, es una unidad de medición inferior a la sílaba con la que puede identificarse la duración de un sonido, generalmente una vocal. Una vocal sólo puede estar constituida por una o dos moras, puesto que la diferencia de duración sólo puede ser *breve / larga*.

Como hemos dicho, cuando la cantidad es un rasgo inherente y, por tanto, opositivo, no puede ser a la vez contrastivo; es decir, su valor no se establecerá por contraste entre unos segmentos y otros, sino que cada segmento tendrá una duración normal comprendida entre unos valores máximos y unos mínimos: el campo de dispersión del rasgo.

Por lo demás, podría establecerse la duración normal de una mora en cada lengua en la que la duración sea fonológica. Por ejemplo, Delattre describe la relación entre la duración de las vocales breves y largas en alemán: de 2 a 5; en inglés: de 2 a 3; y en francés: de 3 a 4 (1965: 63). Es decir, según Delattre las vocales largas en alemán son un 150% más largas que las vocales breves, en inglés un 50% más largas, y en francés un 33%.

En castellano no hay diferencias cuantitativas relevantes entre los sonidos, aparte del fenómeno de la geminación y, por tanto, la cantidad no es, en principio, un rasgo relevante.

Sin embargo, hemos visto que en la percepción del acento paradigmático la duración constituye un indicador importante. Puesto que el acento

paradigmático forma parte de la naturaleza de la palabra, y puesto que hemos visto que la duración puede considerarse también como un rasgo segmental, tal vez pudiéramos determinar si la diferencia de cantidad entre las vocales tónicas y las vocales átonas son los suficientemente relevantes y estables como para poder hablar de *moras* en nuestra lengua, referidas a la duración de las vocales tónicas o átonas.

Los datos que ofrece Navarro Tomás establecen una duración media de las vocales iniciales de 64 mls; de las protónicas de 58 mls; de las tónicas de 106 mls; de las postónicas de 5 mls; y de las vocales finales de 114 mls (1918: 203). Esto es, unas diferencias que rondan el 50%, semejantes a las que en otras lenguas se consideran fonológicas.

Más recientemente, sin embargo, los valores ofrecidos por otros autores son muy distintos. Según Monroy (1980: 134), en castellano la duración de las vocales tónicas oscila entre un 23% y un 33% más que la de las vocales átonas, lo cual no puede interpretarse, a su entender, como una diferencia relevante. En Quilis & Esgueva (1983), y según los datos que proporcionan, esta diferencia se rebaja a un 14%.

Es decir, y en palabras de Martínez Celdrán (1984: 246): "los fonetistas españoles están de acuerdo en afirmar que el concepto de vocal larga no es muy adecuado para las vocales del castellano, pues éstas son en general breves, y las llamadas largas manifiestan una ligera duración mayor respecto de la breve, pero comparadas con las del alemán o inglés suelen ser mucho más breves". Aunque añade: "la posición acentuada favorece la duración y, entre las tónicas, la posición aguda es la más larga".

No parece que pueda hablarse, por tanto, de *mora* referida a la duración estándar de las vocales en castellano, pues ni siquiera las vocales tónicas son lo

suficientemente largas en comparación con las átonas como para que la duración constituya un rasgo relevante en la producción del rasgo acentual. A pesar de todo, y en abierta contradicción con los autores citados, Canellada & Kulhmann citan como principal parámetro de tonicidad la mayor duración de la vocal tónica, que en ocasiones, y según los datos que ofrecen, puede doblar a la duración de la vocal átona en el mismo contexto (19875: 77 y ss.). En cualquier caso, el concepto de "desdoblamiento vocálico" que emplean estos autores (op.cit.: 80) equivale a un desdoblamiento no de la vocal propiamente, sino de la unidad de medición temporal: de la *mora*.

Como veremos más adelante (v. 3.4.1.1.), el concepto de *mora* lo emplearemos para referirnos a este alargamiento de algunas vocales, generalmente tónicas, con un efecto entonativo.

3.2.1.5. Conclusión

En este apartado hemos revisado el carácter secundario del parámetro de la duración, que no puede compararse a las otras tres características del sonido (tono, intensidad y timbre) porque requiere una segmentación previa: también hemos visto cómo los "sonidos del lenguaje", por tanto, no son una realidad directamente sensible, sino que constituyen una suerte de abstracciones que median entre el sonido real y concreto y el concepto de fonema.

Una vez determinado el carácter segmental de la duración, hemos examinado las situaciones en las que la cantidad puede considerarse un rasgo segmental inherente. Paralelamente, hemos examinado cómo el rasgo de "tensión" está relacionado con el fenómeno de la "intensidad", desde una perspectiva segmental. Hemos apuntado la posibilidad de que con la duración ocurra otro tanto: que haya un rasgo segmental inherente referido a la duración

del segmento, junto al fenómeno global de la duración suprasegmental.

Hemos visto también que la cantidad, como rasgo inherente, no puede considerarse contrastiva, sino que el segmento marcado con dicho rasgo ha de contener en sí mismo toda la información relevante para ser [+/-largo]: es decir, en alemán, por ejemplo, las vocales largas o cortas no se establecen en el sintagma y en contraste con las demás vocales presentes, sino que ya la propia constitución del sintagma cuenta con la duración de las vocales como un rasgo más de su naturaleza, *ab initio*. Así, hemos definido "mora" como el concepto con el que identificar la duración de los sonidos [+/-largos]: la vocal larga tiene dos moras, la vocal breve tiene una mora. Tales diferencias no son relativas, sino absolutas y estables, dentro de un margen de dispersión y en cada lengua.

En castellano, la cantidad no es un rasgo relevante. Sin embargo, en la percepción del acento la duración intrínseca de la vocal tónica parece tener una cierta relevancia. Este dato, junto a la observación de que las vocales tónicas suelen ser más largas que las átonas, dejaba entrever la posibilidad de que en nuestra lengua también pudiera hablarse de "mora" para referirse a la duración de las vocales tónicas frente a las átonas. Con todo, los datos de que disponemos parecen rechazar esta idea. Esto quiere decir, finalmente, que la duración no es realmente un parámetro significativo en la producción del acento.

Queda por determinar, sin embargo, si la cantidad puede considerarse un rasgo inherente de las vocales tónicas, aunque redundante: como un rasgo de apoyo en la producción y, sobre todo, en la percepción del acento.

3.2.2. El fenómeno de la reducción vocálica

3.2.2.1. Definición del fenómeno

Según apunta Delattre (1965: 66), "parece ser que en todas las lenguas las vocales cambian de timbre cuando el acento es débil o cuando es fuerte". Es decir, la cualidad de vocal tónica o átona se corresponde no sólo con un cambio de tono (y duración) sino también con un cambio en la estructura tímbrica: las vocales átonas presentan, por lo general, una estructura de formantes más inestable, con una fuerte tendencia a la centralización. "Las sílabas acentuadas ... registran un desplazamiento más importante de los órganos articulatorios que las sílabas inacentuadas; efectivamente, las diferencias de duración entre segmentos acentuados e inacentuados suelen ir acompañadas de diferencias de grado de centralización" (Recasens, 1993: 64). A este fenómeno se le ha llamado "reducción vocálica".

En ocasiones, la reducción vocálica tiene que ver, además, con el fenómeno rítmico de la compresión silábica que se asocia a la isocronía acentual: puesto que cada pie acentual tiende a tener una duración similar, en los pies acentuales con mayor número de sílabas estas se comprimen proporcionalmente. Esta compresión afecta, fundamentalmente, a la vocal, pues es en ella en la que recae la mayor información fónica, como sabemos. Así, la compresión silábica se manifiesta en forma de reducción vocálica de las vocales átonas, y especialmente de las vocales postónicas, que en el habla coloquial llegan incluso a sincoparse (como en *chocolate* → *choc'late*, *every* → *ev'ry*, etc.; Shane, 1973: 94).

En inglés, la reducción vocálica de vocales "plenas" en posición átona da lugar a la unidad llamada *schwa* (transcrita como [ə] o como [ɚ]): "desde el

punto de vista fonológico, el *schwa* es un segmento especificado categorialmente como vocal, [+sil, -cons], pero con «articulación cero» ... desde el punto de vista fonético, la *schwa* debe tener alguna articulación, pero cabe suponer que ésta viene dada por reglas de omisión, ya sean universales o específicas de una lengua" (Durand, 1990: 122).

La mayoría de las veces, sin embargo, la reducción vocálica no tiene nada que ver con el ritmo sino que se debe exclusivamente a la propia tonicidad de la vocal: es el caso del sistema vocálico del catalán.

3.2.2.2. La reducción vocálica en catalán

En catalán, el sistema vocálico está desdoblado en un subsistema de siete fonemas tónicos: / i e ε a ɔ o u /; y otro de tres fonemas átonos: / i ə u /.

Para Recasens (1993: 76-78), este fenómeno constituye un "proceso fonológico" que se hace efectivo sólo después de aplicarse las "reglas de acentuación": es decir, sólo habría que considerar como vocales del catalán los siete fonemas tónicos, pues las formas átonas serían el resultado de un proceso de neutralización en posición átona. De este modo se explicarían las alternancias fónicas del tipo "posar" (con la realización neutralizada [u]) - "poso" (con la realización tónica [ɔ]); así como las producciones dialectales sin reducción vocálica o las excepciones a la norma de la reducción.

Es decir, desde este punto de vista el sistema vocálico del catalán contaría con siete fonemas vocálicos plenos en posición tónica, que en posición átona se neutralizan del siguiente modo: / i / se realiza [i]; / e ε a / se realizan [ə]; / ɔ o u / se realizan [u].

Sin embargo, Recasens habla también de los "fonemas subyacentes" /ə/ y /U/ para explicar formas léxicas en los que las realizaciones [ə] y [u] no alternan

con las formas acentuadas "correspondientes", de las que, por tanto, no pueden "provenir": "[ə] de las palabras «arribar», «rafal», «bossa»; [u] de las palabras «homilia», «encomanar», «carro»" (1993: 78). Este postulado, obviamente, sólo tiene sentido desde una perspectiva morfológica cuya discusión, aquí, no nos interesa.

Sí nos interesa, en cambio, apuntar las implicaciones fonológicas que supone concebir la reducción vocálica en catalán como una neutralización posicional (posición tónica / átona). En primer lugar, habría que considerar, necesariamente, la vocal neutra [ə] como un alófono compartido por los fonemas /e/, /ɛ/ y /a/; del mismo modo, el sonido [u] habría que considerarlo un alófono compartido por los fonemas /ɔ/, /o/ y /u/; el fonema /i/ sería el único que no compartiría alófonos con otros fonemas. En segundo lugar, si bien [ə] podría considerarse un *schwa*, semejante fonética y fonológicamente al *schwa* [ə] del inglés, no podría decirse lo mismo del sonido [u], pues se trata de una realización bien marcada y definida: la asimetría entre ambas realizaciones resulta desconcertante, y habría que postular, en definitiva, un *schwa* redondeado. En tercer lugar, los fonemas que comparten alófonos constituirían dos correlaciones vocálicas neutralizables en dos archifonemas: un archifonema "redondeado" (que podría transcribirse como /O/) y un archifonema "no redondeado" (que podría transcribirse como /A/).

El sistema fonológico resultante, pues, sería:

<u>fon.</u>	<u>alóf.</u>	<u>arch. fon.</u>	<u>alóf.</u>	<u>arch. fon.</u>	<u>alóf.</u>
/i/	→ [i]	/A/	→/e/ →[e]	/O/	→/ɔ/ →[ɔ]
			↘ [ə]		↘ [u]
		→/ɛ/	→ [ɛ]	→/o/	→[o]
			↘ [ə]		↘ [u]
		→/a/	→[a]	→/u/	→[u]
			↘ [ə]		

La explicación tradicional del sistema vocálico catalán (Badia, 1988) considera que existen dos subsistemas, como decíamos en un principio: un subsistema tónico y otro átono. Desde esta perspectiva, la vocal neutra constituye un fonema de pleno derecho /ə/ (lo cual posibilita su existencia como fonema tónico en las hablas mallorquinas, por ejemplo), no hay que postular alófonos compartidos, ni *schwa* en catalán, ni archifonemas vocálicos, ni "fonemas subyacentes".

Sin embargo, contemplar un desdoblamiento fonológico en función de la tonicidad o no de las vocales también supone algunas implicaciones fonológicas de peso, que Martínez Celdrán expone de este modo: "(en algunas lenguas) el acento no tiene una función fonológica o su rendimiento es muy escaso. El catalán oriental central es una modalidad lingüística que neutraliza sus vocales en posición átona, de modo que pasa de las siete vocales en la posición acentuada ... a las tres de la posición átona ... por tanto, hay un cambio de timbre simultáneo en el paso de una posición a otra. Desde el punto de vista fonético, el catalán es una lengua de acento libre, pero el cambio de timbre resta importancia a su carácter fonológico. En esta modalidad lingüística, el

rendimiento funcional del acento es muy escaso" (1989: 65).

Es decir, la existencia de cualidades tímbricas intrínsecas al carácter tónico o átono de las vocales resta, en cierto modo, funcionalidad al acento como fenómeno suprasegmental: "es evidente que en todos estos casos el acento carece de valor fonológico, pues la presencia de una de esas vocales (no reducidas) es suficiente para poner de relieve esa sílaba frente a las demás, por cuanto en ellas predomina el timbre sobre los rasgos prosódicos" (íbid.: 66).

En realidad, múltiples realizaciones dialectales y no pocas excepciones normativas a la "regla de la reducción vocálica" aconsejarían más bien una posición en la que los rasgos propios del acento se consideren prioritarios frente al cambio tímbrico, que es mucho menos frecuente en el conjunto del dominio lingüístico del catalán de lo que cabría esperar. Por eso Martínez Celdrán circunscribe bien su aguda interpretación fonológica al dialecto "catalán oriental central".

Considerar las vocales átonas como fonemas, con todo, conlleva otras complicaciones teóricas: y es que no pueden hallarse parejas léxicas en las que el método de la conmutación funcione entre fonemas tónicos y átonos. Ejemplos como "posa / posar", que ortográficamente parecen una pareja opuesta por el acento, se transcriben como [pósə] / [pusá]: es decir, no constituyen una pareja homófona distinguida sólo por un segmento, sino que todos los segmentos vocálicos son distintos (la alternancia [ó] / [u] y la alternancia [ə] / [á]). Es decir, excepto en casos como "bulli / bullí" (en los que, justamente, no hay reducción vocálica) los fonemas tónicos / átonos no pueden oponerse en el mismo contexto fónico. O, dicho en otras palabras, no pueden considerarse, desde una perspectiva trubetzkyana, fonemas distintos. O bien las formas reducidas (átonas) son variantes de los siete fonemas tímbricamente plenos

(tónicos), como hemos visto, y puede hablarse de neutralización y archifonemas, o bien las alternancias / e ε a / - / ə / y / ɔ o u / - / u / constituyen fonemas de distribución defectiva (que no pueden oponerse porque no aparecen nunca en las mismas posiciones: posición tónica / posición átona). Considerar la primera opción permitiría, además, una explicación muy sencilla de las variedades dialectales en las que no hay reducción vocálica, así como de las excepciones normativas: la reducción vocálica de las átonas sería un fenómeno fonético y contextual, pero no fonológico; es decir, no podría hablarse realmente de archifonemas.

De cualquier modo, el fenómeno en sí permite extraer jugosas conclusiones sobre el carácter del acento en relación con las vocales.

Frente a la idea de que el acento "recae" sobre las vocales en el sintagma, la evidencia de que en algunas lenguas ya en el paradigma las vocales son "tónicas" o "átonas", incluso con un cambio de timbre, refuerza la idea contraria, según la cual el acento es consustancial a la vocal *ab initio*: en catalán, por ejemplo, la cadena hablada se elabora a partir de vocales tónicas y átonas, que lo son *por naturaleza* y que incluso puede considerarse que constituyan unidades fonológicas distintas.

Esta evidencia de timbre también es coherente con la naturaleza breve de las vocales átonas, frente a las tónicas: "hay correlación entre las diferencias de tonicidad y duración vocálicas: así, las vocales [i] y [u] átonas presentan una duración menor que [i] y [u] tónicas, y una duración similar a [ə]. De hecho, las diferencias entre todas las vocales tónicas y todas las vocales átonas del sistema vocálico del catalán resultaron significativas según el test Newman-Keuls ... por el contrario, ninguna de las diferencias entre las duraciones de las vocales átonas [i], [u] y [ə] da resultados significativos. Puede concluirse, pues, que la duración

vocálica permite oponer vocales átonas y cualquier vocal tónica pero no vocales átonas entre sí" (Recasens, 1986: 150).

3.2.2.3. Tendencias en el vocalismo del castellano

El fenómeno de la reducción de las vocales átonas ocurre en catalán y en otras lenguas de nuestro entorno, como el portugués, cuyo sistema vocálico está compuesto por tres subsistemas, según la posición pretónica, tónica y postónica. En castellano, sin embargo, no se observa dicha reducción, al menos en términos fonológicos.

En términos fonéticos, en cambio, tal vez sí pueda hablarse de una cierta tendencia a la reducción vocálica, teniendo en cuenta la centralización obvia a la que tienden las realizaciones vocálicas átonas.

Examinaremos brevemente esta cuestión. En principio, los valores formánticos de las vocales del castellano suelen ofrecerse sin discriminar entre vocales tónicas y vocales átonas: unas veces sólo se tienen en cuenta las vocales tónicas y otras se promedian los valores de ambos grupos. Así, por ejemplo, Delattre (1965: 49) ofrece los siguientes valores:

	i	e	a	o	u
F1	275 hz	450 hz	725 hz	450 hz	275 hz
F2	2300 hz	1900 hz	1300 hz	900 hz	800 hz

Quilis (1981: 171-174), por su parte, ofrece los siguientes:

	i	e	a	o	u
F1	364 hz	405 hz	607 hz	486 hz	324 hz
F2	1863 hz	1701 hz	1012 hz	931 hz	729 hz

En Quilis & Esqueva (1983) se distingue entre los informantes masculinos y los femeninos, y entre la posición tónica y la posición átona. Los valores medios ofrecidos de los informantes masculinos son (1983: 244):

	i	e	a	o	u
F1	264,5 hz	453,8 hz	657,28 hz	474,5 hz	293,5 hz
F2	2317,5 hz	1995,01 hz	1215 hz	888,4 hz	669,08 hz

Martínez Celdrán (1984), por último, ofrece los siguientes:

	i	e	a	o	u
F1	345,5 hz	461 hz	680 hz	499,5 hz	373 hz
F2	2197,5 hz	1884 hz	1265 hz	1022 hz	981 hz

Todos estos datos corresponden a los valores medios obtenidos entre vocales tónicas y vocales átonas, y generalmente de análisis acústicos realizados sobre frases leídas por los informantes. Esto quiere decir que en el habla espontánea pueden verse muy acentuadas las tendencias que aparezcan en tales análisis de laboratorio.

Por ejemplo, si, según Quilis & Esqueva, en las realizaciones de sus informantes masculinos "se manifiesta un pequenísima tendencia hacia realizaciones más abiertas de las vocales tónicas, produciéndose muy pocas diferencias entre las tónicas y las átonas" (1983: 243), posiblemente en el habla espontánea dicha tendencia no sea tan "pequeñísima". Así lo han demostrado Poch & Harmegnies (1992), en cuyos datos aparece una tendencia muy acusada a la centralización de los valores formánticos, característicos del habla espontánea.

Según Iruela (1993: 82): "este fenómeno, llamado «reducción» o

«centralización» está causado por la diferencia de estilos de habla. Las vocales descontextualizadas o en palabras aisladas, tienden a realizarse en estilo cuidado, mientras que las extraídas de un discurso espontáneo o una narración pertenecen al estilo vernacular, por lo tanto, se pronuncian de una forma más relajada y tienden a desplazarse hacia el centro, acortándose las distancias entre ellas". En el trabajo de Iruela (centrado en la descripción acústica de las vocales españolas pronunciadas por alumnos hablantes de holandés), en efecto, entre las vocales tónicas y las vocales átonas hay diferencias de timbre considerables.

En resumen, podríamos tomar la carta de formantes ofrecida por Martínez Celdrán (1984: 299) como paradigma de tales diferencias de timbre entre las vocales tónicas y las átonas: en la figura 40 puede verse cómo las vocales átonas están claramente contenidas en el triángulo de las vocales tónicas, es decir, son realizaciones más centralizadas, más igualadas entre ellas ([i] y [e], por una parte, y [o] y [u], por otra, están más cerca en el triángulo átono que en el tónico).

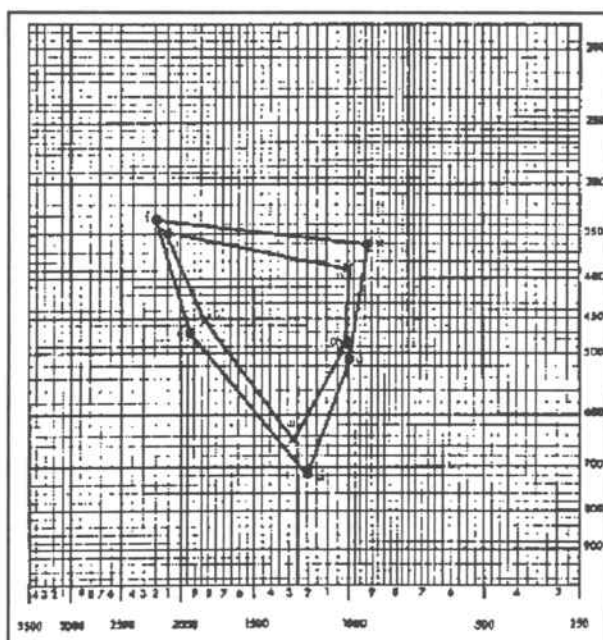


Figura 40 - Carta de formantes de las vocales tónicas y átonas del castellano.

Con todo, las diferencias tímbricas entre tónicas y átonas no son lo suficientemente acusadas en castellano como para que por sí solas pudieran justificar un desdoblamiento fonológico, como en catalán o portugués. Sin embargo, merece la pena tener en cuenta el dato de que existe una cierta tendencia a la centralización de las realizaciones

átonas: sin duda, algún día habrá que investigar dicha tendencia en el habla espontánea, así como en los distintos dialectos. Nuestra hipótesis, de momento, es que se vería claramente confirmada: pues las tendencias actualmente documentadas se hallaron, como hemos visto, en frases leídas, en un estilo muy cuidado y realizadas por informantes cultos. Tal vez, en castellano haya que hablar de posición pretónica, tónica y postónica, como en portugués: así lo hace Navarro Tomás (1918) y así parece establecerse en los fenómenos de inestabilidad vocálica y pérdida de vocales.

Pero hay otro argumento importante: según Mateo (1988: 177), "las vocales tónicas presentan su frecuencia fundamental más alta que las mismas vocales en posición átona". Es decir, el tono intrínseco de las vocales varía significativamente en las vocales tónicas o átonas.

Como decíamos en el apartado 1.4.1., el tono intrínseco de un sonido depende del juego de los formantes que lo caracterizan (la altura de la lengua, su adelantamiento, etc., se manifiesta acústicamente en la altura de los formantes: ya vimos que a cada movimiento articulatorio corresponde un rasgo de timbre y que, por tanto, decir que el tono intrínseco es un producto tímbrico equivale a decir que obedece a las características articulatorias del sonido). En definitiva, si la tonicidad de la vocal influye en su tono intrínseco es que entre las vocales tónicas y átonas hay diferencias de timbre (y, por tanto, diferencias articulatorias) lo suficientemente acusadas. Este dato coincide, de nuevo, con la tendencia a la centralización tímbrica de las vocales átonas.

En definitiva, como dice Solé (1984:149): "las vocales inacentuadas castellanas no tienen la misma cualidad vocálica o frecuencia de formantes que sus homólogas acentuadas". Sin embargo, el cambio de timbre no puede considerarse un rasgo relevante del acento: "la reducción de la frecuencia de

formantes de las vocales inacentuadas es tan pequeña ... que resulta casi negligible ... por tanto consideramos que la estructura formántica vocálica no es una dimensión importante para la percepción del acento en el lenguaje real" (op.cit.: 150).

Aun así, la autora reconoce que la reducción vocálica podría tal vez marcar el acento, mediante la síntesis de voz: "no ponemos en duda que se pudiese conseguir el efecto perceptivo del acento variando exageradamente las frecuencias de formantes de las vocales castellanas; sin embargo, este resultado no se correspondería con el lenguaje real donde no existe una diferencia significativa en este respecto entre las vocales acentuadas y las inacentuadas" (ibid.). Es decir, y como dijimos de la duración intrínseca de las vocales asociada a su tonicidad, tal vez el cambio de timbre entre las vocales tónicas y átonas pueda considerarse un rasgo redundante más, que "podría en principio ayudar a distinguir la vocal que lleva el acento" (op.cit.:149), aunque en principio no constituyera un elemento relevante.

Así, tal vez en el habla cuchicheada la percepción del acento dependa de tales rasgos redundantes: la duración de la vocal tónica y la estructura tímbrica de las vocales átonas (también en el apartado 1.4.1. vimos cómo puede haber una sensación tímbrica sin fonación, y cómo de la estructura tímbrica puede inferirse la F_0 , aunque con un elevado porcentaje de error: los rasgos redundantes de duración y cambio tímbrico vendrían a suplir tal porcentaje de error y a asegurar la percepción del acento).

En cualquier caso, los fenómenos de la duración vocálica y el cambio tímbrico son rasgos segmentales que podrían colaborar en la producción y en la percepción del acento, como simples tendencias fonéticas o como rasgos redundantes (también fonéticos, por tanto).

Esta idea, que no entra en contradicción con el carácter tonal del acento, ni con los experimentos que han descrito dicho carácter tonal, y que además permite explicar mejor algunos de los fenómenos que hemos examinado, podría formularse del siguiente modo: las vocales tónicas y las vocales átonas lo son por naturaleza, esto es, las vocales tónicas son largas, plenas tímbricamente y, sobre todo, tienen un tono intrínseco característico de su tonicidad; las vocales átonas, en cambio, son breves, su timbre está centralizado o es inestable y su tono intrínseco es inherente a su atonicidad.

3.2.3. Implicaciones fonológicas

Una vez apuntada la posibilidad de contemplar el acento como un fenómeno segmental, no contrastivo sino opositivo, conviene repasar muy brevemente qué tipo de implicaciones fonológicas supondría esta interpretación de los hechos en nuestra lengua.

En primer lugar, si el acento paradigmático no es un fenómeno superpuesto a la palabra, sino propio de su naturaleza; y si tampoco es un fenómeno superpuesto a las vocales que forman la palabra sino que es propio también de la naturaleza de tales vocales, habría que considerar que los sistemas vocálicos están compuestos no por vocales indistintas al acento sobre las cuales recaería este, sino por vocales tónicas y átonas, que lo son ya por naturaleza. Así ocurre en catalán, como hemos visto (y al margen de su interpretación fonológica). En castellano, esto implicaría desdoblar el sistema de cinco vocales en dos sistemas: cinco tónicas y cinco átonas.

El desdoblamiento fonológico, sin embargo, tal vez no estaría justificado por las diferencias encontradas entre las vocales tónicas y las átonas: se

impondría, entonces, considerar al menos un desdoblamiento de los alófonos de cada vocal. En realidad, la situación no sería muy distinta a la que encontramos en catalán (aunque sin una reducción vocálica tan acusada): podría considerarse que las vocales tónicas y las átonas (que constituyen sonidos distintos) son variantes del mismo fonema, según el esquema: fonema / a / → alófonos [á] - [a]; fonema /e/ → alófonos [é] - [e] - [´ε] - [ε]; etc. En la descripción fonética de las vocales, finalmente, debería incluirse también la descripción de los rasgos de tonicidad: duración, cualidad vocálica, tensión y tono.

Con todo, las vocales tónicas y átonas no pueden considerarse simplemente alófonos distintos, sino que hay que especificar que entre ellos hay necesariamente una distribución defectiva, pues en cada palabra ha de haber necesariamente una vocal tónica, y solo una. Es otra forma de referirse a la cualidad "contrastiva" de la posición acentual. En "ceno / cenó" [θéno] - [θenó], el alófono [é] no puede aparecer en la segunda palabra, ni el alófono [e] en la primera.

Ya hicimos referencia al escaso rendimiento lingüístico del acento en castellano. Nótese, además, que en nuestra lengua hay muy pocas palabras cuyas vocales sean de "igual" timbre que se opongan únicamente por el acento: "abras / habrás" [áβras] - [aβrás]; normalmente, las palabras opuestas acentualmente están compuestas por vocales "distintas": "llevo / llevó" [léβo] - [λεβó]. El hecho de que casos como "abras / habrás" sean tan raros corrobora el carácter escasamente distintivo, escasamente fonológico, del acento en sí, cuyo rendimiento apenas hace rentable tener unidades fonológicas distintas: pues, según la perspectiva que hemos propuesto, muchos casos como este aconsejarían realmente el desdoblamiento fonológico del sistema vocálico.

En realidad, nuestra perspectiva relacionaría muy de cerca la fonología y

el léxico: pues en la realidad los fonemas no existen aisladamente, sino formando palabras. El acento paradigmático, propio de la naturaleza de las vocales tónicas, es a la vez propio de la naturaleza de las palabras.

Generalmente, las palabras constituyen unidades bien definidas, aunque en ocasiones encontramos palabras cuya naturaleza aún no está bien fijada por el uso popular. Por eso, tal vez, en las palabras nuevas, desconocidas o cultas, en los barbarismos y en los acrósticos, hay notables fluctuaciones acentuales y hasta de timbre: porque en tales casos la cuestión fonológica no está bien establecida socialmente, y por tanto fluctúa. Así, acrósticos como "OTAN" se pronuncian indistintamente [ótan] - [otán]; vocablos cultos como "perito, erudito, cuadriga" se pronuncian a menudo "périto, erúdito, cuádriga"; barbarismos como "consola, *Parkinson*" se pronuncian [kónsola] - [parkinsón]; etc.

También las parejas del tipo "periodo / período" se diferencian únicamente por la pronunciación, es decir, son la misma palabra cuya naturaleza acentual fluctúa; semejante a este fenómeno, hay una fluctuación tímbrica en las vocales de no pocas unidades léxicas, como en "somos / semos", "civil / cevil", "ministro / menistro", "oscuro / escuro", "abuelo / agüelo", "ovular / uvular", e incluso "verter / vertir", "durmiendo / dormiendo", etc., no achacables simplemente a la ignorancia del hablante, sino a la poca relevancia del timbre de tales vocales a la hora de usar y entender la palabra.

En otros casos, en fin, el propio criterio de los especialistas fluctúa, lo cual no puede ser achacable a su ignorancia, sino a la propia naturaleza fluctuante de la palabra a que nos referimos: por ejemplo, Quilis (1993: 397) considera la forma "medula" [medúla] como "acentuación errónea", dando por tanto por buena la forma "médula", mientras que Corominas (1961: 393) documenta la forma "medula" en 1444 y añade: "mal acentuado *médula* ya en

el S. XVII".

En tales fluctuaciones, es la propia naturaleza de la palabra la que no está bien fijada por el uso: fenómeno que a veces afecta a su morfología ("el / la mar", "boda / bodas"), a su función sintáctica ("español": sustantivo / adjetivo), etc.

El acento no puede regirse por reglas posteriores a la selección léxica para formar el discurso: por el contrario, en el lexicón ya está la información acentual de la palabra, puesto que el acento paradigmático es intrínseco a la palabra, que se ha formado combinando vocales tónicas y átonas, que ya lo eran por naturaleza.

El rasgo distintivo de la tonicidad de la vocal es el tono: debería considerarse, pues, un rasgo segmental inherente (como vimos que puede considerarse también la duración y la intensidad), que tiene que ver con el tono intrínseco de la vocal, más o menos alto: [+/- tónica].

El rasgo [+/- tónica] puede manifestarse acústicamente como un tono intrínseco elevado, pero también, dependiendo de la duración del sonido, como una inflexión tonal, con tantas alturas tonales como moras (salvo raras excepciones, dos moras por un sonido [+ largo]). Así, la palabra "bien" [bjén], pronunciada sin énfasis, está nucleada en torno a la vocal [é], de una mora de duración y de un tono elevado; pronunciada con énfasis, "¡bien!" [bjéːn], se nuclea en torno a la vocal [éː], larga, de dos moras y por tanto de dos alturas tonales: es decir, la manifestación de la tonicidad, en el segundo caso, toma la forma de una inflexión tonal. A esta doble duración de la vocal tónica en algunas ocasiones Canellada & Kuhlmann lo llaman "desdoblamiento" de la vocal (1987: 80).

La diferencia entre [bjén] y [bjéːn], obviamente, no es fonológica, sino

estrictamente fonética: así, el rasgo de duración [+/- larga] es redundante en la caracterización fonológica de la vocal tónica o átona, como habíamos apuntado. Del mismo modo, podría emplearse el rasgo [+/- tensa] para describir la producción de una vocal concreta, como un rasgo fonético también redundante y no distintivo en castellano. El rasgo de plenitud o reducción vocálica, por último, y según los datos de que disponemos actualmente, podría considerarse una tendencia debida a un efecto coarticulatorio de los rasgos de tensión y tonicidad; no obstante, sería importante realizar un estudio experimental en nuestra lengua para determinar si se trata efectivamente de una tendencia despreciable o si por el contrario en el habla espontánea la centralización es importante y sistemática. En este caso, el rasgo [+/- centralización] sería un rasgo redundante más.

Obviamente, si consideramos las vocales tónicas y átonas como alófonos de un mismo fonema, el rasgo [+/- tónica] tampoco sería un rasgo distintivo del tipo [+/- densa], [+/- difusa], [+/- grave], pues no distinguiría fonemas y ya hemos visto el escaso rendimiento fonológico del acento; sin embargo, sí sería un rasgo distintivo en la determinación de las variedades tónicas / átonas de cada fonema vocálico, con distribución defectiva.

3.3. Acento sintagmático y grupo fónico

3.3.1. El acento sintagmático

3.3.1.1. Definición

Hasta ahora nos hemos referido al acento identificándolo, genéricamente, con lo que hemos llamado *acento paradigmático*, el acento que distingue vocales tónicas y átonas y que es propio de la naturaleza de las palabras. Cuando se habla de entonación, sin embargo, y como vimos en el capítulo anterior, las distintas escuelas y tradiciones hablan también de "acento de frase" (*sentence stress*) o de "núcleo" de la entonación (*nucleus*).

En efecto, la cadena fónica no constituye un único nivel en el que se suceden indistintamente los sonidos sino que, como hemos ido viendo, se organiza en bloques fónicos alrededor de los acentos, verdaderos núcleos del discurso. Ya hemos visto cómo los sonidos consonánticos se organizan alrededor de los sonidos vocálicos (formando sílabas) y estos alrededor de las vocales tónicas para constituir las *palabras fónicas* o *grupos rítmicos*.

Pero la jerarquía fónica no se queda en el nivel rítmico. Por el contrario, las sucesivas palabras fónicas se organizan también y forman bloques, alrededor de un acento: no de un acento paradigmático, que se circunscribe al nivel léxico, sino de un acento jerárquicamente superior, que actúa a un nivel no léxico sino sintagmático, y al que llamamos, pues, *acento sintagmático*.

Tenemos, así, dos tipos de acento: el propio de las palabras léxicas, sobre el que se forman las palabras fónicas y, por tanto, el ritmo del discurso (acento paradigmático); y el propio de los sintagmas, sobre el que se articulan las sucesivas palabras fónicas (acento sintagmático).

Los bloques de palabras fónicas nucleados en torno a un acento sintagmático son los llamados *grupos fónicos*, que fragmentan el discurso en unidades comprensibles, como veremos, y que constituyen la unidad fónica superior, más allá de la cual habrá unidades gramaticales y discursivas, pero no unidades fónicas (Canellada & Kuhlmann, 1987, por ejemplo, hablan, más allá del grupo fónico, del "período", unidad a la que no atribuyen, sin embargo, ninguna entidad fónica).

El acento sintagmático, núcleo de un grupo fónico, sí es un acento que "recae", un acento "superpuesto" (frente al acento paradigmático, propio de la vocal y de la palabra), aunque no "superpuesto" sobre una vocal o una sílaba, sino sobre una palabra, lo cual equivale a decir que "recae" sobre la vocal tónica de esa palabra (de un modo parecido a como veíamos con respecto a la relación sílaba / vocal, en este caso toda la palabra no puede estar acentuada en su conjunto, sino únicamente su núcleo fónico: su vocal tónica).

En realidad, el acento sintagmático podría definirse como el acento propio de los sintagmas, del mismo modo que el acento paradigmático es el propio de las palabras. Sin embargo, el hecho de que no haya un inventario o un "lexicón" de sintagmas, pues no constituyen significantes sistemáticamente unidos a conceptos estables, y de que su creación sea inmediata y contextual, obliga a concebir el acento sintagmático como un fenómeno realmente suprasegmental, semejante a la propia entonación, y no es posible, como ocurre con el acento paradigmático, interpretarlo desde una perspectiva segmental.

El acento sintagmático, así, consiste en un acento paradigmático puesto de relieve sobre los demás acentos de su mismo grupo fónico: generalmente, el acento sintagmático es el último acento paradigmático del grupo (sobre esta cuestión, sin embargo, cfr. 3.4.2.).

El parámetro mediante el cual se pone de relieve dicho acento paradigmático sobre los demás acentos es fundamentalmente el tono, una vez más: en concreto, una inflexión tonal. Ya dijimos que, en ocasiones, la marca del rasgo [+ tónica] no era simplemente un tono elevado, sino una inflexión tonal, en los casos en los que la duración de la vocal es mayor: así, el acento sintagmático se caracteriza, fonéticamente, por una mayor duración con respecto a los demás acentos del grupo fónico, así como de un movimiento tonal complejo, en forma de inflexión. Cada acento sintagmático, pues, será a la vez un acento paradigmático (con su función léxica y rítmica concreta) y un acento sintagmático (con la función de agrupar en torno a sí el grupo fónico), y sus marcas fonéticas diferenciales (duración + inflexión tonal) deben considerarse formas variantes de las marcas fonéticas habituales del rasgo [+ tónica].

Por tanto, el acento sintagmático podría considerarse un fenómeno de transición entre el acento y la entonación, pues, por una parte, consiste en un acento paradigmático puesto de relieve sobre los demás y, por otra, el modo de ponerse de relieve es mediante una inflexión tonal que, como veremos, constituye la parte más informativa de la entonación.

La posición de la tradición norteamericana es la de distinguir netamente el fenómeno acentual (el "acento de frase") y el fenómeno entonativo (la inflexión final, considerada una juntura), aunque coincidan en la misma sílaba: dicha posición, como vimos, estaba fundamentada en la negación del tono como parámetro acentual, y en la idea del "acento de intensidad". Vimos incluso cómo algunos autores negaban validez científica a los experimentos que probaban el carácter tonal del acento, con tal de no renunciar a la separación tradicional del acento y la entonación (v. 2.3.).

La tradición británica, más sutil, considera el "núcleo" del contorno como

un fenómeno entonativo, pues se trata de una inflexión tonal, aunque sigan identificando los demás acentos de la frase con el parámetro de intensidad. En realidad, de lo que se trata es de distinguir igualmente entre acento y entonación, aunque la línea divisoria entre ambos fenómenos se coloca justamente entre los dos tipos de acentos (v. 2.2.).

Según nuestra perspectiva, este es uno de los elementos teóricos más débiles de ambas tradiciones, pues no aceptan plenamente las evidencias experimentales del carácter tonal del acento y, por tanto, no proceden a un replanteamiento general de sus teorías en el sentido de contemplar cómo en la realidad se relacionan el acento y la entonación (v. apartado 2.6.).

En definitiva, nuestra definición de los dos tipos de acento consiste en considerarlos a los dos como fenómenos tonales, uno anclado en el nivel léxico y rítmico, otro en el nivel sintagmático; uno como elemento propio de las palabras, y por tanto inventariado, y otro como elemento propio del sintagma, y por tanto creativo; uno como elemento segmental, otro como fenómeno suprasegmental; uno, en definitiva, plenamente acentual, otro concebido como puente y unión entre los fenómenos del acento y la entonación.

3.3.1.2. El grupo fónico

El grupo fónico está constituido por una serie de palabras fónicas organizadas alrededor de un acento sintagmático. Ciertamente, el grupo fónico suele coincidir con un sintagma gramatical, aunque su naturaleza acentual y fónica nos obliga a distinguir entre ambas unidades: el grupo fónico es el recipiente en el cual se amolda el sintagma gramatical, pero no puede ser definido en función de la solidaridad gramatical que pueda establecerse entre las palabras que lo componen, porque pertenece a otro nivel. Por el contrario, más

bien entendemos el sintagma gramatical como el conjunto de palabras que constituye un grupo fónico, gramaticalmente solidarias precisamente por constituir un único grupo fónico. Por ejemplo, si emitimos una serie de palabras inventadas, el oyente no puede entender lo que hemos dicho, aunque sí podrá interpretar los grupos fónicos en que hemos segmentado nuestra emisión. Es decir, el grupo fónico es una unidad fónica, sin dependencias gramaticales.

Para Navarro Tomás, el grupo fónico es "la porción de discurso comprendida entre dos pausas o cesuras sucesivas de la articulación" (1918: 30). De este modo, establece el carácter estrictamente fonético de la unidad, pues es definida sólo en función de elementos fonéticos, y no gramaticales o de sentido.

Numerosos autores han tomado directamente la definición de Navarro Tomás y la han incorporado a sus constructos teóricos (Canellada & Kuhlmann, 1987: 103; Quilis, 1993: 418; etc.). Sin embargo, definir esta unidad en función de las pausas, como propone Navarro Tomás, conlleva algunos problemas, pues no siempre pueden encontrarse pausas en los límites de grupo; Canellada & Kuhlmann, por ejemplo, dicen: "las pausas que delimitan los grupos pueden ser: pausas grandes, como las que existen al final de la frase, o de algunos grupos; pausas pequeñas, que delimitan algunos grupos; pausas mínimas, que a veces ni siquiera son interrupciones de voz, sino pausas virtuales" (íbid.). Es decir, puede haber pausas grandes, pequeñas, mínimas o virtuales, esto es, pausas que pueden no ser pausas. O, lo que es lo mismo, cuando estos autores se refieren a la "pausa", en realidad están empleando el concepto de "juntura" propio de la tradición norteamericana. (Sobre la relación entre pausa y entonación, v. el apartado 4.1.3.).

El hecho es que nunca se ha empleado el concepto de "acento de frase" para definir el grupo fónico y, por tanto, había que buscar límites externos al

propio grupo: pausas o junturas. Sin embargo, empleando el "acento de frase" o *acento sintagmático* (como lo hemos llamado) no es necesario acudir a elementos extraños. Esto es así por la tendencia de los fonetistas a segmentarlo todo claramente; pero, como vimos con la delimitación de la sílaba y del grupo rítmico, lo importante tanto en la producción como en la percepción de unidades fónicas no son los límites, que normalmente constituyen zonas marginales y poco informativas que pueden perderse sin menoscabo, sino los núcleos de tales unidades: la vocal, en el caso de la sílaba; la vocal tónica, en el caso del grupo rítmico; el acento sintagmático, en el caso del grupo fónico. La percepción es culminativa, y sólo le interesan los núcleos fónicos.

Es decir, la percepción del grupo fónico no depende de escuchar la pausa final, que puede ser "mínima" o "virtual" (¡y, por tanto, necesariamente irrelevante!), sino que consiste en identificar la vocal tónica en la que recaiga una inflexión tonal: generalmente, esa será la marca de final de grupo fónico, pues generalmente el último acento paradigmático del grupo es el que cumple la función de acento sintagmático. Al fin y al cabo, esto mismo es lo que viene a significar, en el análisis de niveles, la marca de "juntura", que coincide con la inflexión final de la frase entonativa.

La diferencia entre las frases especificativas y explicativas suele cifrarse en la aparición o no de comas y, en el discurso oral, de pausas. Por ejemplo, en las frases (tomadas de Quilis, 1981: 389, que las aduce como un ejemplo de la función delimitadora de la entonación -v. apartado 3.4.1.1.-): "los alumnos que viven lejos llegan tarde" (especificativa) y "los alumnos, que viven lejos, llegan tarde" (explicativa), el sintagma "que viven lejos" puede estar separado o no por pausas (por comas), según el sentido de la frase. Sin embargo, según nuestra perspectiva la obligatoriedad de la coma está determinada por la norma escrita

de separar los grupos, pero no se refiere a la pausa en la pronunciación (como en los casos en los que la norma escrita obliga a colocar coma tras las expresiones del tipo "sin embargo", "con todo", "no obstante", etc., aunque no haya pausa; o como la prohibición de separar por coma sujeto y predicado, aunque suela haber pausa; etc.).

Según nuestra perspectiva, pues, la diferencia entre estas frases no obedece a las pausas (que pueden estar o no) sino a los acentos sintagmáticos con que las formulemos: dos o tres acentos sintagmáticos, es decir, dos o tres grupos fónicos.

En el primer caso, la frase es especificativa porque se divide sólo en dos grupos fónicos:

/losalúmnos kebíben IÉxos/ - /légan tÁrde/

(en mayúscula y negrilla marcamos el acento sintagmático: núcleo del grupo fónico); en el segundo caso, la frase es explicativa porque se divide en tres grupos fónicos:

/losalÚmnos/ - /kebíben IÉxos/ - /légan tÁrde/

Podemos aducir otros ejemplos de la articulación del discurso en grupos fónicos. En la cadena [bámos akomér pólos] ("vamos a comer pollos" - ejemplo brindado por la Dra. V. Salcioli) podemos distinguir hasta cuatro frases distintas:

/bámos akomér pÓlos/

("vamos a comer pollos": con un solo acento sintagmático y, por tanto, constituyendo un único grupo fónico);

/bÁmos/ - /akomér pÓlos/

("vamos, a comer pollos": la primera palabra se convierte en un grupo fónico independiente y, por tanto, el acento paradigmático de "vamos" se convierte aquí en acento sintagmático, a la misma altura jerárquica que el acento de "pollos");

/bámos akomÉR/ - /pÓlos/

("vamos a comer, pollos": en esta ocasión, el acento de "vamos" no se pone de relieve, pero el de "comer" sí, a la misma altura que el de "pollos"; sigue habiendo dos grupos fónicos, pero se trata de grupos distintos, al tratarse de acentos sintagmáticos distintos);

/bÁmos/ - /akomÉR/ - /pÓlos/

("vamos, a comer, pollos": en este último caso, cada palabra fónica constituye por sí sola un grupo fónico, y cada acento paradigmático constituye un acento sintagmático; es decir, los tres acentos están al mismo nivel y no hay ninguno jerárquicamente superior a los demás).

En este ejemplo, la diferencia entre las cuatro frases obedece exclusivamente a la organización en grupos fónicos de la cadena sonora, que es idéntica porque en los cuatro casos no sólo hay los mismos segmentos tímbricos y en el mismo orden, sino que incluso los acentos paradigmáticos son exactamente los mismos: sólo cambian los acentos sintagmáticos. Otros fenómenos distintivos entre cadenas en principio iguales pero con una organización acentual distinta no son comparables a este, sin embargo, pues no obedecen sólo a la organización en grupos fónicos: así, la frase "esta va calentita" no es confundible con "estaba calentita", porque los acentos paradigmáticos no son iguales: /ésta bá kalentÍta/ (tres grupos rítmicos) frente

a /estába kalent**í**ta/ (dos grupos rítmicos). Sí pueden distinguirse sólo en función de los acentos sintagmáticos, en cambio: /ésta bá kalent**í**ta/ ("esta va calentita") y /ésta b**Á**ka lentita/ ("esta vaca lentita").

Según sus datos, Navarro Tomás establece la duración de los grupos fónicos del siguiente modo: de una a cinco sílabas, un 12,55%; de cinco a diez sílabas, un 67,60%; más de diez sílabas, un 26,32%. Los más comunes son los grupos fónicos de siete y ocho sílabas: un 26,32% del total (citado en Canellada & Kuhlmann, 1987: 103, que, según sus propios datos, establecen que "el grupo más corriente tiene entre cinco y diez sílabas").

Estos datos corresponden a la lengua literaria; en cambio, y según Quilis, "en la lengua hablada los hechos son radicalmente distintos: ... el 51% de los grupos fónicos corresponden a los que tienen de una a cinco sílabas ... la media del número de sílabas por grupo fónico es de 9,5" (1993: 418-419).

En efecto, la forma natural de la lengua hablada es el diálogo y en él predominan, sin duda, las emisiones breves de voz, así como los monosílabos.

Finalmente, y según nuestras propias observaciones, al parecer el grupo fónico es la unidad dentro de la cual funciona la isocronía léxica, con una cierta tendencia a la isocronía de grupos fónicos: es decir, el ritmo tiene como base el grupo rítmico pero también tiene como contenedor el grupo fónico; por otra parte, entre grupos fónicos de distinta duración, y especialmente cuando el discurso constituye un monólogo, los grupos fónicos tienden a igualar su duración. Estos fenómenos explicarían, tal vez, la dispersión de los datos que se observan en numerosos trabajos sobre el ritmo en español y algunas contradicciones notables sobre su carácter silábico o acentual (Toledo, 1988).

Es decir, más allá del grupo fónico no hay ningún fenómeno fónico: ni siquiera el ritmo.

3.3.2. El proceso de comprensión del discurso

Como hemos ido viendo, la forma fónica del discurso no consiste en una mera continuidad de segmentos encadenados e indistintos, sino que constituye unidades superiores al segmento: la palabra fónica o grupo rítmico, esto es, los sonidos agrupados en torno a un acento paradigmático; y el grupo fónico, es decir, las palabras fónicas organizadas en torno a un acento sintagmático.

La comprensión del discurso requiere que el oyente discrimine las unidades que lo constituyen. Las unidades que el oyente discrimina, sin embargo, no son directamente las unidades gramaticales o semánticas (palabras léxicas, sintagmas gramaticales, oraciones, etc.), sino las unidades fónicas, que son la forma material de aquellas (palabras fónicas y grupos fónicos).

La estructura acentual del discurso, pues, es el principio organizador de la materia sonora y, por tanto, el filtro mediante el cual el hablante organiza y el oyente identifica bloques fónicos compactos. El oyente no se enfrenta a una mera sucesión de sonidos, sino a una estructura lineal pero bien jerarquizada de segmentos átonos nucleados en torno a segmentos tónicos, a su vez nucleados en torno a acentos sintagmáticos.

Esta organización de la materia sonora permite distinguir dos tipos de unidades: el grupo fónico, identificable dentro de todo el discurso como una unidad, y la palabra fónica, identificable dentro del grupo fónico.

Como hemos dicho, los grupos fónicos suelen contener una estructura léxico-gramatical coherente y suelen coincidir con los núcleos semánticos del

discurso (hecho que, sin embargo, no implica dependencia alguna entre ambos niveles de análisis). Así, y frente a lo que se podría suponer, la comprensión no va "de menor a mayor": de fonema a morfema, a palabra, a frase, a texto; sino que comienza directamente por el grupo fónico y va en dos sentidos (v. figura 41):

a) "de mayor a menor y a mayor" (lo que llamamos *comprensión en zig-zag*): el oyente identifica el grupo fónico, y dentro de él la palabra fónica; de la palabra fónica abstrae la información léxico-semántica e identifica el sintagma gramatical, que constituye un núcleo semántico del discurso; y

b) "de mayor a mayor" (lo que llamamos *comprensión progresiva*): de grupo fónico, identificado como núcleo semántico, a grupo fónico, etc.

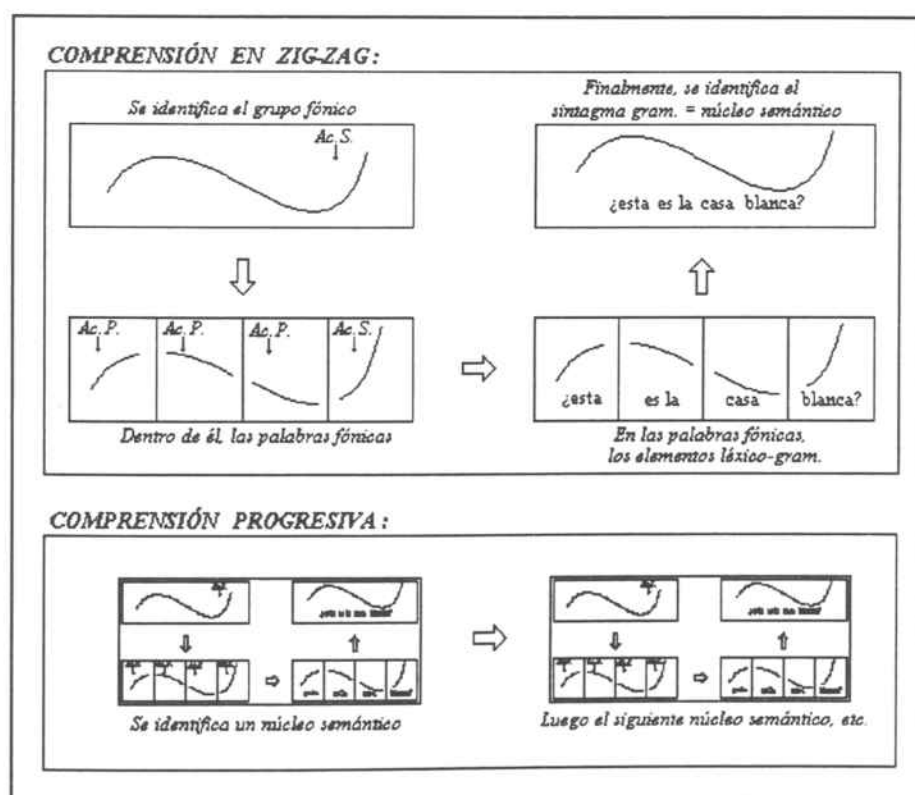


Figura 41 - Proceso de comprensión del discurso.

Es decir, justamente el primer nivel de contacto auditivo entre hablante y oyente (el material fónico del discurso) es el que ofrece las unidades significativas mediante la segmentación en bloques fónicos del discurso. Es lo que más adelante (v. apartado 3.4.1.) llamaremos la *función prelingüística* del acento y la entonación.

En realidad, en el discurso escrito la disposición gráfica y visual de las letras y las palabras cumple buena parte de dicha función prelingüística; aun así, para comprender el texto escrito el lector ha de segmentarlo en bloques del único modo que sabe: oyéndolo, oyéndolo mentalmente y agrupando los sonidos que oye mentalmente en grupos rítmicos y en grupos fónicos.

En ocasiones, el lector está tratando de leer un texto pero también está pensando en otra cosa, o tal vez está oyendo música, o la televisión, o una conversación ajena: en tales casos, el lector tendrá irremisiblemente la sensación de que no entiende nada, y volverá a leer la misma página una y otra vez hasta que se "concentre", es decir, hasta que pueda *oír* su propia voz mental leyendo el texto. Si no puede *oír* el texto e integrarlo fónicamente, será incapaz de comprender su contenido. Es lo que podríamos llamar *carácter auditivo de la lectura*. Muchos casos de fracaso escolar, por ejemplo, están relacionados con este fenómeno: el alumno que no ha elaborado una estrategia de comprensión auditiva del texto escrito se enfrenta a los textos como un profano en solfeo se enfrenta a un pentagrama, obcecado por su carácter engañosamente visual y sin poder *oír*, ni por tanto entender, nada. Esta estrategia errónea, sin duda, puede ser determinante en un ambiente educativo fundamentalmente lecto-escritor.

El proceso de comprensión del discurso, sin embargo, no acaba en la identificación de las unidades fónicas que contienen a las unidades léxico-gramaticales y, por tanto, a los núcleos semánticos del discurso, sino que se ve

completado por otros dos fenómenos paralelos.

El primer fenómeno es que la comprensión de las unidades discretas obedece a varios factores, entre los que sobresalen la *influencia del contexto* (que decide el significado de las palabras, la intención que puede suponerse al hablante, etc.) y, sobre todo, la comparación con los modelos propios del oyente: modelos que los lingüistas suelen imaginar como un "lexicón" o inventario, aunque la realidad parece ser más compleja. Tales modelos del oyente le permiten identificar: los sonidos que componen el discurso como pertinentes a la lengua; las palabras fónicas como unidades en las que puede identificarse una palabra léxica conocida y otros elementos gramaticales o léxicos (artículos, preposiciones, etc.; y prefijos, sufijos, etc.), o tal vez una palabra léxica desconocida aunque reconocida como posible (y cuyo significado, entonces, hay que adivinar o deducir); las estructuras gramaticales aceptables por el oyente y/o comprensibles; etc.

Cuando decimos que el acento sintagmático es un acento que "recae" sobre una vocal tónica, nos referimos exclusivamente al proceso de formación del discurso, en el que suponemos, por tanto, que la selección del léxico (con su acento propio) es previa a la formación de los sintagmas. Cuando, por el contrario, decimos que el acento sintagmático podría considerarse el "propio del sintagma", adoptamos una perspectiva perceptiva y nos referimos al proceso de comprensión del discurso, en el que el oyente se enfrenta con una serie de grupos fónicos *dados*; es decir, para el oyente el acento sintagmático es previo, no está "superpuesto" a la vocal tónica sino que esa vocal tónica tiene la forma de una inflexión tonal, y no puede separarse, en ella, la cualidad de acento paradigmático (propio de la palabra en sí, y de la vocal) del acento sintagmático.

El otro fenómeno que forma parte de la comprensión del discurso es el

de las *expectativas* del oyente, que irá elaborando, a medida que entienda las primeras unidades, una serie de hipótesis sobre el sentido, la intención y la dirección que tomará el discurso del hablante, incluida la dirección rítmica y entonativa. Tales hipótesis serán más o menos acertadas según el grado de "intersubjetividad" (grado de acuerdo entre los interlocutores sobre los presupuestos del discurso, o grado de aceptación del oyente, o grado de conocimiento mutuo) en que se desarrolle la comunicación; en cualquier caso, las hipótesis del oyente condicionarán totalmente la comprensión del discurso, a pesar de que irán reelaborándose a medida que este avance.

Es decir, podría hablarse no sólo de "preplanificación" de la entonación en el hablante (v. 2.4.2.), sino también, en cierto modo, de "preplanificación" tonal en el oyente, durante el proceso de comprensión y en forma de "hipótesis" sucesivas: eso explicaría cómo ya en el mismo comienzo de una entonación enfática, por ejemplo, de sorpresa y alegría, el oyente comparte los gestos de alegría con el hablante, pues en el comienzo de la entonación hay suficiente información entonativa como para que el oyente pueda elaborar su hipótesis y "preplanificar", anticipar su propia audición. (Sobre los procesos de formación y comprensión del discurso, cfr. Luria, 1975a, 1975b, 1979; y Cantero & De Arriba, en prensa).

3.3.3. *Grupo fónico y contorno entonativo*

Dice Navarro Tomás que "los límites de la unidad melódica coinciden en español con los del grupo fónico" (1944: 31). En efecto, el grupo fónico es la unidad fónica mayor en que se segmenta el discurso y la única que tiene entidad propia, no supeditada a otras unidades: el grupo rítmico, por ejemplo, está

integrado y organizado dentro del grupo fónico, como las sílabas y los segmentos lo están en el grupo rítmico. Así, la única unidad fónica que puede contener la línea melódica es el grupo fónico.

Tradicionalmente, la unidad de análisis entonativo se ha identificado con unidades de carácter semántico o gramatical: se habla de "entonación de la oración", de "frase entonativa", de "entonación de la cláusula", etc. Como hemos visto, el grupo fónico debe entenderse como una unidad de carácter estrictamente fónico, definido por un acento sintagmático, aunque suele agrupar un sintagma léxico-gramatical que constituye un núcleo semántico. Esta es la principal relación, tal vez la única (v. apartado 4.1.1.), que podemos establecer entre el nivel fónico (acentual-entonativo) y los niveles abstractos y conceptuales del enunciado (léxico-gramatical y semántico).

A la unidad melódica la llamamos *contorno entonativo*, siguiendo la terminología aceptada usualmente en nuestro ámbito. Con este término nos referimos exclusivamente a la línea melódica que constituye una unidad estructural o significativa en la entonación del discurso. Así, el grupo fónico, unidad fónica mayor del discurso, en la que se organizan jerárquicamente los sonidos del lenguaje y dentro de la cual se establece la estructura rítmica del discurso, también es el contenedor del *contorno entonativo*.

En los ejemplos anteriores "vamos a comer pollos", etc., por tanto, las cuatro frases resultantes según la distinta estructura de grupos fónicos son frases distintas también en cuanto a su estructura entonativa, pues el número de contornos también es distinto. Obsérvese, en la figura 42, cómo las palabras fónicas "vamos" y "a comer", tienen una curva melódica ascendente cuando forman parte del mismo contorno entonativo que la palabra siguiente y descendente cuando constituyen un grupo fónico independiente.

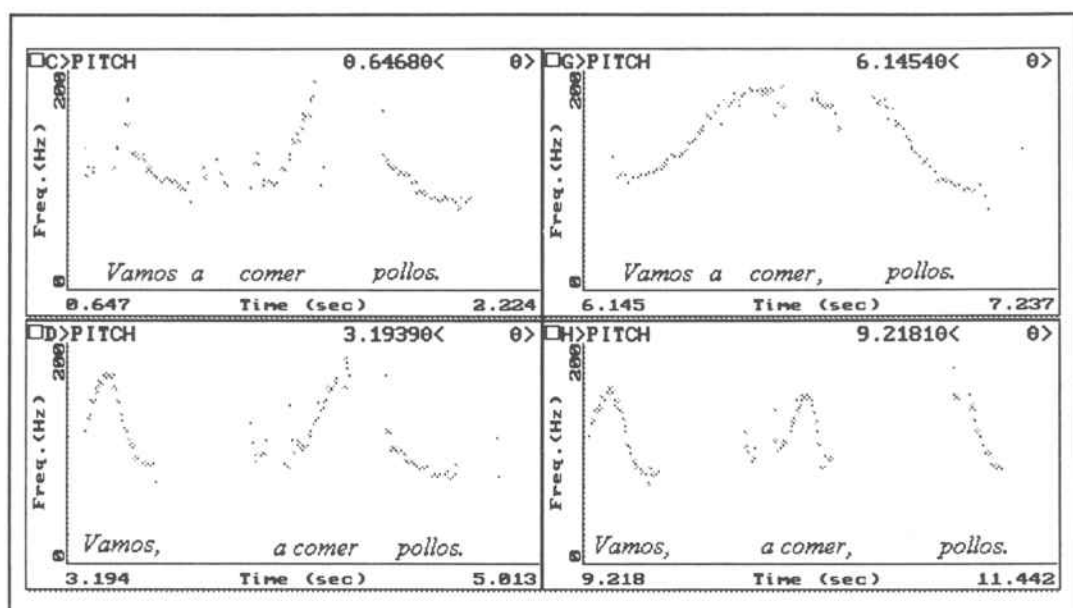


Figura 42 - Las cuatro estructuras entonativas del enunciado: "Vamos a comer pollos".

Es decir, no sólo la estructura de grupos fónicos es distinta en las cuatro frases: igualmente, es distinta la propia estructura entonativa de cada una.

3.4. Acento y entonación

3.4.1. *La función prelingüística*

3.4.1.1. La función integradora-delimitadora

En el apartado 2.5.2. revisamos la división que hace Quilis de las funciones de la entonación en un nivel lingüístico: la función integradora, la función delimitadora y la función distintiva (1981: 377 y ss.).

La primera consiste en que la entonación es el medio de "integrar las palabras para formar una oración" (op.cit.: 384); es decir, no hay ni puede haber enunciados sin entonación, pues es la entonación la encargada de unir e integrar los sonidos en el habla: "pueden aparecer enunciados sin forma gramatical, pero sin entonación, no" (op.cit.: 385, citando a Daneš, 1960).

Por su parte, la función delimitadora consiste en que "la entonación delimita los enunciados y segmenta el continuum de discurso en un determinado número de unidades por razones fisiológicas (necesidad de inspirar el aire necesario para la fonación), por razones de comprensión del mensaje ... o por otros motivos lingüísticos" (op.cit.: 388).

Es decir, las funciones integradora y delimitadora constituirían una suerte de anverso y reverso del mismo fenómeno: por una parte, la entonación une e integra, por otra segmenta y divide. Conviene, por tanto, ser más precisos en su caracterización.

Hemos visto que en el proceso de comprensión del discurso la segmentación del continuum sonoro en bloques fónicos es fundamental para la identificación de unidades lingüísticas por parte del oyente: identificando el grupo fónico y el grupo rítmico el oyente puede proceder a identificar la palabra

léxica y el sintagma gramatical o núcleo semántico. Y hemos visto que tanto el grupo fónico como el grupo rítmico son unidades fónicas nucleadas y definidas en torno a un acento: respectivamente, el acento sintagmático y el acento paradigmático.

Por otra parte, la división del continuum sonoro en bloques fónicos (función delimitadora) es equivalente a la integración de los sonidos en bloques (función integradora), pues no hay una verdadera división entre elementos dados, sino que nosotros detectamos la adscripción de un elemento fónico a un bloque o a otro porque está integrado y nucleado en dicho bloque: delimitación e integración no sólo son fenómenos similares o paralelos, sino que son exactamente el mismo fenómeno.

Ciertamente, el contorno entonativo, unidad melódica de la entonación, está contenido en el grupo fónico; pero el principal elemento responsable de la identidad del grupo fónico-contorno entonativo no es la propia melodía sino el acento sintagmático: es decir, la integración (o nucleación) de los sonidos en unidades fónicas (la palabra fónica o grupo rítmico, y el grupo fónico-contorno entonativo) no es un fenómeno entonativo (en el sentido restrictivo de que no es un fenómeno melódico), sino un fenómeno acentual, en el que toman parte tanto el acento paradigmático, propio de las palabras (nucleando los grupos rítmicos) como el acento sintagmático (elemento común del acento y la entonación, cuya función primordial es la de nuclear el grupo fónico).

En realidad, los ejemplos que Quilis aduce en su trabajo son propiamente de naturaleza acentual (1981: 388): por ejemplo, las frases especificativas o explicativas, cuya diferencia vimos en el apartado 3.3.1.2. que obedece a una estructura distinta de acentos sintagmáticos y, por tanto, de grupos fónicos; o ejemplos como "es la villanita" frente a "es la villa, Anita", cuya diferencia,

obviamente, nada tiene que ver con la entonación, pues la propia estructura acentual de ambas frases es distinta: en "es la villanita" /ÉslabiλanÍta/ hay dos acentos paradigmáticos, mientras que en "es la villa, Anita" /ÉslabÍλanÍta/ hay tres.

En este ejemplo, por tanto, las frases "es la villanita" y "es la villa, Anita" no sólo no puede decirse que sean distinguidas por la entonación, como hemos visto, sino que ni siquiera son "distinguidas" por su estructura acentual: simplemente, nunca podrían confundirse, porque constituyen dos frases cuya materia fónica está organizada, está integrada de una forma distinta, alrededor de sus propios acentos paradigmáticos (que obedecen, como sabemos, a la propia selección léxica de los constituyentes del discurso).

Así, la función integradora y la función delimitadora constituyen una única función indivisible, que podríamos llamar "función integradora-delimitadora", siguiendo la nomenclatura tradicional, pero que es preferible llamar *función prelingüística*: pues el fenómeno se refiere a la forma fónica previa del habla, a su propia naturaleza fónica antes de cualquier otra consideración. Es decir, no es que no haya ni pueda haber habla sin entonación, sino que el fenómeno va más lejos aún: y es que los sonidos del lenguaje sólo pueden existir integrados en unidades fónicas merced a la acción nucleadora de los dos tipos de acento.

El carácter fundamentalmente acentual de la *función prelingüística* de la entonación viene dado por la solidaridad que mantienen el grupo fónico y el contorno entonativo, y especialmente por el carácter tonal del acento sintagmático: como veremos, la misma inflexión que es relevante en el contorno entonativo es la marca del núcleo del grupo fónico, en torno al cual se organizan los grupos rítmicos, etc. Con todo, aún veremos otros elementos de relación

entre acento y entonación.

Podríamos decir, por ahora, que la entonación cumple su función prelingüística (la integración de los sonidos en bloques fónicos) a un nivel básicamente acentual, en el que la melodía apenas juega un papel secundario.

3.4.1.2. La entonación idiomática

Tradicionalmente, los dialectólogos se han ocupado de describir las diferencias fonéticas y morfonéticas del dialecto estudiado con respecto a la lengua estándar o normativa: diferencias de pronunciación de los sonidos aislados, tales como el seseo, el ceceo, el yeísmo, el rehilamiento de las palatales, la abertura de las vocales, la aspiración de [x], la aspiración de {-s}, la reducción de grupos consonánticos, la diptongación, la monoptongación, etc.; y diferencias morfonéticas, como las desinencias verbales características, los sufijos, los pronombres, etc. (la obra de Zamora Vicente, 1960, constituye un buen ejemplo de este procedimiento de caracterización dialectal, completado por una caracterización léxica y gramatical no siempre bien desarrollada, precisamente por falta de datos suficientes).

Sin embargo, los rasgos de pronunciación en realidad apenas caracterizan la forma fónica propia del dialecto en cuestión. Por el contrario, es fácil darse cuenta de que no la caracterizan en absoluto: cuando un hablante de un dialecto intenta imitar el habla de otro dialecto únicamente con alguno de estos rasgos de pronunciación (por ejemplo, un humorista aficionado tratando de imitar a un andaluz sirviéndose sólo del ceceo) el resultado es lamentable y muy revelador. Por otra parte, en ocasiones tenemos la oportunidad de conversar con algún extranjero cuya pronunciación resulta impecable: a pesar de lo cual no puede evitar el característico "acento extranjero". Del mismo modo, a un hablante

andaluz difícilmente dejará de notársele el "acento dialectal" aunque pronuncie cada sonido aisladamente conforme a la norma: aunque no sesee, ni cecee, ni aspire, etc.

Es decir, el "acento" característico de cada dialecto no es una cuestión de pronunciación de los sonidos aislados, sino que es una cuestión de entonación. Navarro Tomás llamó a este fenómeno "entonación idiomática" y lo definió en estos términos: "la cadencia habitual del habla de cada país, el dejo característico de cada dialecto y hasta el sello que imprimen en el lenguaje de cada individuo las circunstancias especiales de su actividad o profesión habitual y de su propia y particular manera de ser" (1944: 10). Se correspondería, pues, con lo que Quilis llama "nivel sociolingüístico" de la entonación, que informa de las características propias del individuo, así como de su procedencia geográfica, social, etc. (v. 2.5.2.).

En Quilis (1993: 453 y ss.), este autor presenta las diferencias entre distintas entonaciones dialectales atendiendo exclusivamente a la melodía, y llega a la conclusión de que "frente a la multiplicidad de variantes entonativas, los patrones fundamentales de lo que es el funcionamiento de la entonación en el plano de la lengua se mantiene constante y en pleno rendimiento" (op.cit.: 489); es decir, entre las distintas entonaciones dialectales hay un común denominador en un nivel distintivo. Y añade: "este sería un ejemplo más de una hermosa variedad dentro de la unidad de nuestra Lengua Española" (íbid.).

Sin negar validez a la afirmación de Quilis, sin embargo, conviene llamar la atención sobre una realidad bastante menos idílica, que en algunas ocasiones se manifiesta en forma de la más absoluta incompreensión entre hablantes de una misma lengua, pero de dialectos distintos: en efecto, el fenómeno afecta a la manera de integrar los sonidos en el discurso, y por tanto al propio proceso de

comprensión del discurso.

No se trata, pues, y una vez más, de un fenómeno exclusivamente entonativo (melódico) sino que tiene que ver con la función prelingüística de la entonación: se trata de un fenómeno a la vez acentual y entonativo.

El acento y la entonación son los responsables de la integración del habla en bloques fónicos, como hemos visto, que es lo que permite la identificación por parte del oyente de las unidades significativas del discurso. Tales elementos, pues, son los primeros con los que se enfrenta el oyente. También son, por tanto, la primera barrera que separa, por ejemplo, al nativo del no nativo: el hablante extranjero suele tener una gran dificultad en identificar los bloques fónicos emitidos por un nativo, del mismo modo que a este le irrita profundamente la anómala integración fónica del discurso que, por su parte, realiza el extranjero.

Esta situación se debe a un enfoque fundamentalmente lecto-escritor en la enseñanza de lenguas extranjeras (supuestamente "comunicativa"), que no habitúa al alumno a emplear el recipiente fónico propio de la lengua meta. Así, el léxico, las reglas gramaticales, las normas de adecuación pragmática y toda la competencia comunicativa adquirida por el alumno de una lengua extranjera (que puede verter exitosamente en un texto escrito), e incluso todas las normas de pronunciación correcta de los sonidos aislados, en el momento de ponerlos en juego para producir un discurso oral genuino los integra según el recipiente de su propia lengua de origen: el ritmo, la formación de los grupos fónicos y las cadencias tonales no son los de la lengua meta (en la que está codificado el contenido del discurso) sino los de su propia lengua materna. A este fenómeno lo llamamos "acento extranjero".

Paralelamente, la dificultad que le supone la comprensión auditiva en la

lengua aprendida (seguir el ritmo e identificar las unidades fónicas) es un producto del mismo fenómeno: la manifestación auditiva del acento extranjero.

Entre hablantes de la misma lengua, pero de dialectos distintos, el fenómeno es similar: desde el efecto hipnótico que supone escuchar un discurso comprensible pero integrado de una forma distinta y chocante (y que explica el éxito reciente de algunas series televisivas venezolanas en nuestro país, como antes el de las películas mejicanas de *Cantinflas* o el de los filmes protagonizados por tonadilleras andaluzas), o la irritación que en algunos individuos producen *acentos* dialectales distintos al suyo propio (y que no es ajeno, al parecer, a algunos casos de xenofobia *blanda*), hasta la franca incompreensión que se produce entre hablantes con un *acento* dialectal "muy cerrado". En todos estos casos, la integración del discurso obedece a una estructura rítmica diferente (que puede afectar al tempo, o a la compresión silábica, o al mismo esquema de isocronía silábica, acentual, léxica, etc.: cfr. Toledo, 1988, 1994) y a una cadencia entonativa característica.

Es decir, el *acento* dialectal o extranjero tiene que ver con la propia integración de los sonidos en bloques fónicos (grupos rítmicos y grupos fónicos). La *entonación idiomática*, por tanto, es un producto de la función prelingüística; o, en otras palabras, la *entonación idiomática* es una entonación "prelingüística": no distintiva, no lingüística propiamente, tampoco expresiva, sino previa, en íntima relación con el acento y el ritmo, y mediante la cual la materia fónica que compone el discurso se organiza jerárquicamente.