

# I. INTRODUCCIÓN

## 1. El concepto de entonación

### *1.1. Definición y parámetros de la entonación*

La entonación es el fenómeno lingüístico que constituyen las variaciones de tono relevantes en el discurso oral. Como veremos a lo largo de este trabajo, se trata del principal elemento cohesionador del habla, que además cumple distintas funciones lingüísticas y expresivas en la comunicación oral.

La naturaleza de la entonación, con todo, es múltiple, y en el fenómeno intervienen factores muy diversos que hacen de su estudio un ámbito particularmente resbaladizo y polémico. Tanto es así que conviene, en un capítulo introductorio como este, hacer una breve referencia a algunos de los aspectos básicos que nos ayudarán a enfocar debidamente nuestro objeto de estudio. Así, en estas páginas el lector especializado encontrará explicaciones que puede considerar triviales, pero que sin duda ayudarán al lector que se introduce en el tema por primera vez. A menudo, en nuestra ciencia damos por sabidas cosas que no todo el mundo sabe, y que además algunos autores han tratado con cierto descuido.

Por ejemplo, no pocos autores definen la entonación como un fenómeno en el que interviene el tono, pero también otros parámetros acústicos, como la intensidad, la duración o incluso el timbre.

En general, suele decirse que las características fonéticas de los sonidos del habla son cuatro: el tono, la intensidad, el timbre y la duración. Esta afirmación es esencialmente correcta, y muy útil en un manual de introducción a la fonética, aunque para nuestros objetivos conviene hacerle algunos matices:

En primer lugar, hay que tener en cuenta que la voz es un fenómeno exclusivamente tonal: en el proceso de “fonación” (o producción de la voz) el aire vibra a su paso por la laringe, y genera una serie de frecuencias: la frecuencia fundamental ( $F_0$ ) o tono, y las frecuencias secundarias o armónicos. El timbre, por tanto, es un fenómeno tonal derivado, producido después de la fonación, durante la “articulación” de la voz: los resonadores supraglóticos (las cavidades faríngea, bucal y nasal)

potencian unos armónicos u otros, y moldean el timbre de la voz, incluyendo el timbre de los sonidos vocales. Algunos sonidos, desde luego, no tienen tono: son producidos directamente en las cavidades de resonancia, sin voz (como ocurre con las fricativas sordas).

Lo que llamamos “sonido” es, en realidad, una abstracción de sus rasgos tímbricos: la estructura de armónicos, en el caso de los sonidos vocales, y otros rasgos (silencio, explosión o ruido) en las consonantes. La duración de un sonido, entonces, es una abstracción de segundo grado: medimos la duración de un segmento sonoro, definido por los rasgos de timbre que lo constituyen, midiendo uno de esos rasgos tímbricos. Por ejemplo, la duración de una vocal podemos establecerla midiendo su segundo formante.

La intensidad de la voz depende de la presión infraglótica del aire que la produce, y es directamente proporcional al tono: para hacer vibrar más o menos los pliegues vocales y producir, así, un cambio de tono, es necesaria mayor o menor presión infraglótica. (Cuando en la articulación, además, hay una amortiguación en la salida del aire, o cuando se produce un ruido, también se modifica la intensidad del sonido).

Es decir, el timbre es un fenómeno tonal, la duración es una característica que requiere del timbre, y la intensidad no constituye un parámetro independiente en los sonidos vocales (aunque sí es un parámetro esencial en la caracterización acústica de algunas consonantes).

Así, podríamos decir que el habla se halla constituida, 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, la misma terminología empleada para referirnos a estos 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 (una concepción que tiene un origen abiertamente lecto-escritor y, aún más, alfabético: v. Cantero, 1995); ni siquiera la evidencia fonética de que cada sonido no es más que un haz de rasgos tímbricos fluctuantes, cuyo límite no siempre está muy claro, ha sido capaz de eliminar el mito de que el habla está formada por 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, como hemos dicho, 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

ra 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 (más adelante, cuando tratemos el tema con detenimiento, veremos que es una definición que conviene completar). 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, estos fenómenos suprasegmentales tienen un cierto anclaje en el ámbito segmental, aunque, de nuevo, se trata de fenómenos tonales.

El parámetro principal del acento, en efecto, es el tono (v. 4.1.3.): una inflexión de tono convierte el segmento en acentuado, aunque también una mayor duración o intensidad del sonido en ocasiones pueden ser determinantes.

Si los parámetros físicos del fenómeno suprasegmental acento son, por tanto, la  $F_0$ , la duración y la intensidad, tal vez los parámetros del otro fenómeno suprasegmental, la entonación, puedan ser los mismos. Así, nos encontramos con dos definiciones distintas y, en principio, 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”.

Hemos visto que hay una relación 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: v. 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  $F_0$ .

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; por tanto, 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.

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; 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.

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, definimos, 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 supra-segmentales”, etc. Sin embargo, el segundo tema, el carácter motivado y emocional de la entonación, 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.

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 (1981: 376 y ss.), 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).

Nuestra perspectiva no difiere esencialmente de este enfoque: veremos cómo la entonación actúa en un nivel prelingüístico (cohesionando el discurso); en un nivel lingüístico (distinguiendo unidades); y en un nivel paralingüístico (añadiendo información emocional).

## 1.2. $F_0$ , 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. 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 Hertzios (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; con todo, hace entrar en juego conceptos claramente fonológicos, como “grupo espiratorio” (al que hace coincidir, además, con el “sintagma gramatical”) o como “prominencia” (un concepto similar 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).

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 hemos dicho, interrumpida sólo por los segmentos sordos: una melodía constituida por tonos más o menos estables y por tonos que se deslizan, que fluctúan, etc. Esa melodía, sin embargo, 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 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 un 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 por dos hablantes muy distintos, seguramente un hombre y 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: en frecuencias más bajas, una diferencia menor equivale a una diferencia mayor en 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 intuitivos 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).

## 2. El análisis tradicional de la entonación: precedentes teóricos

### 2.1. 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 (llamado *análisis de niveles*).

En un principio, ambos métodos tienen una vida paralela y, en cierto modo, independiente, hasta que algunos autores intentan conjugarlos en sus trabajos, siguiendo la propuesta conciliadora de Daneš (1960) de analizar la entonación por “configuraciones de niveles”. Durante algún tiempo, en efecto, los autores suponían 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 (como documenta Ladd, 1980: 10).

Las diferencias entre los dos métodos, desde luego, son ciertas y afectan a cuestiones de fondo: 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, considerando la entonación más bien como un fenómeno motivado, activo sólo en niveles expresivos, no lingüísticamente codificados.

El análisis de niveles, así, considera que la entonación está formada por una serie de “fonemas entonativos” cuya sintaxis da lugar a los “morfemas entonativos”, que serían ya unidades significativas y que afectarían a una cláusula o sintagma gramatical. Sin embargo, los análisis de este método rara vez han sobrepasado el nivel