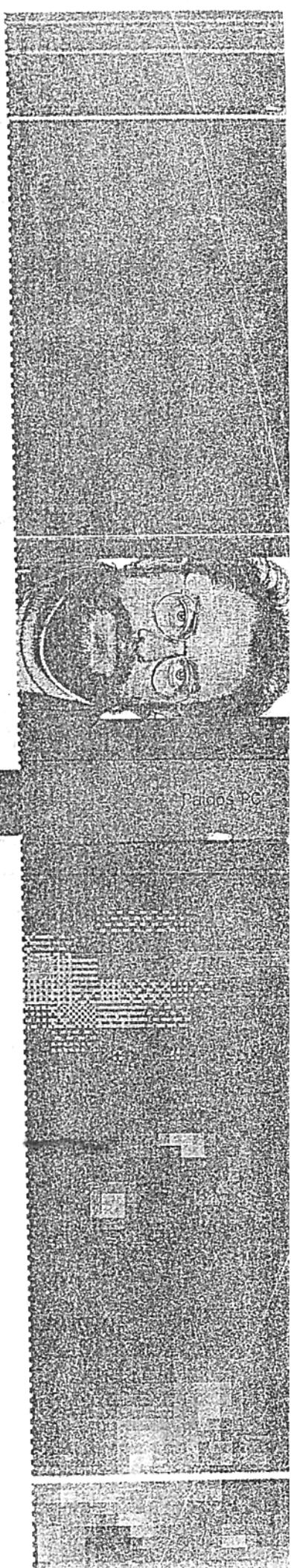


- Audiovisuales - Leng. Audiov. -  
Folio 12

10.90



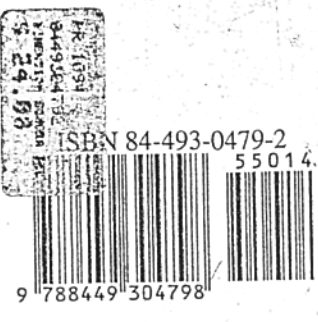
El papel del sonido ha sido, hasta ahora, absolutamente infravalorado en toda la bibliografía sobre el lenguaje audiovisual. Esta obra resitúa al audio frente a la imagen, analizando con un rigor poco habitual su papel expresivo y narrativo.

En el libro, se revisa el concepto de lenguaje audiovisual y sus métodos de estudio a partir de las últimas aportaciones hechas por la psicoacústica y por la psicología de la percepción. Desde esa perspectiva, el autor propone nuevos planteamientos sobre el rol del audio dentro del sistema sonido-imagen, que resultan especialmente fértiles tanto desde una perspectiva analítica como desde la perspectiva productiva y profesional.

A lo largo de la obra se desarrollará un estudio muy didáctico sobre el uso de los métodos de análisis acústico y sus diversas posibilidades de aplicación, y se propone, también, una terminología sistemática para nombrar y describir con eficacia los sonidos.

Finalmente, el estudio desemboca en la aplicación de todos estos instrumentos de trabajo sobre cuestiones tan concretas como la creación de espacios audiovisuales, los usos expresivos y narrativos de la sincronía sonido-imagen o la organización de secuencias de imágenes mediante el sonido.

Àngel Rodríguez Bravo es doctor en Ciencias de la Información y profesor del departamento de Comunicación Audiovisual y Publicidad de la Universidad Autònoma de Barcelona. Ostenta el Premio Extraordinario de Doctorado en la Facultad de Ciencias de la Comunicació en 1991, y fue galardonado en 1992 por la Generalitat de Catalunya con el Premio a la Investigación sobre Comunicación de Masas por su estudio «La construcción de una voz radiofónica». Ha sido director del primer Master de Creación y Producción de Ficción y Entretenimiento Audiovisual. En la actualidad, dirige el Laboratorio de Anàlisis Instrumental en Comunicació de Masas (LAICOM).



La dimensión sonora del lenguaje audiovisual Àngel Rodríguez

Àngel Rodríguez  
La dimensión  
sonora del  
lenguaje  
audiovisual

Ángel Rodríguez Bravo

## La dimensión sonora del lenguaje audiovisual

1. M. Fontcuberta - *La noticia*
2. I. Tubau - *Periodismo oral*
3. F. F. Díez y J. Martínez Abadía - *La dirección de producción para cine y televisión*
4. Á. G. Meseguer - *¿Es sexista la lengua española?*
5. M. Á. Ortiz y J. Marchamalo - *Técnicas de comunicación en radio*
6. J. M. Pérez Tornero - *El desafío educativo de la televisión*
7. P. Rodríguez - *Periodismo de investigación: técnicas y estrategia*
8. L. Arfuch - *La entrevista, una invención dialógica*
9. M. López - *Cómo se fabrican las noticias*
10. J. M.ª Perceval - *Nacionalismos, xenofobia y racismo en la comunicación*
11. M. Á. Ortiz y F. Folpini - *Diseño de programas de radio*
12. P. Soler - *La investigación cualitativa*
13. V. Llorens - *Fundamentos tecnológicos en vídeo y televisión*
14. A. Rodríguez - *La dimensión sonora del lenguaje audiovisual*
15. S. Zunzunegui - *La mirada cercana*
16. M. Onaindia - *El guión clásico de Hollywood*
18. T. Álvarez y M. Caballero - *Vendedores de imagen*
19. J. Rey - *Palabras para vender. Palabras para soñar*

## SUMARIO

|                                                                                                                              |    |
|------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Presentación, Armand Balsebre</b> .....                                                                                   | 11 |
| <b>Al lector</b> .....                                                                                                       | 13 |
| <b>Introducción</b> .....                                                                                                    | 15 |
| <br>                                                                                                                         |    |
| <b>1. Sobre los planteamientos metodológicos que orientan esta obra</b> .....                                                | 17 |
| 1.1. La subjetividad como problema metodológico en las corrientes clásicas de investigación sobre lenguaje audiovisual ..... | 19 |
| 1.2. La alternativa de las metodologías cualitativas .....                                                                   | 20 |
| 1.3. La percepción humana como punto de partida .....                                                                        | 22 |
| 1.4. ¿Existe un lenguaje audiovisual distinto para cada medio? .....                                                         | 22 |
| 1.5. El lenguaje audiovisual como objeto de estudio .....                                                                    | 25 |
| 1.5.1. Definición y características esenciales .....                                                                         | 25 |
| 1.5.2. Sobre la dimensión naturalista del lenguaje audiovisual .....                                                         | 27 |
| 1.5.3. Sobre la naturaleza incompleta del lenguaje audiovisual .....                                                         | 28 |
| 1.6. De la física a la percepción de los mensajes audiovisuales .....                                                        | 31 |
| <br>                                                                                                                         |    |
| <b>2. La acústica y la comunicación audiovisual</b> .....                                                                    | 33 |
| 2.1. La acusmatización .....                                                                                                 | 35 |
| 2.1.1. Desarrollo de un nuevo concepto .....                                                                                 | 35 |
| 2.1.2. La acusmatización en la comunicación de masas .....                                                                   | 36 |
| 2.2. Los conocimientos necesarios para dominar el audio .....                                                                | 38 |
| 2.2.1. La tecnología y la experiencia productiva .....                                                                       | 38 |
| 2.2.2. La acústica .....                                                                                                     | 39 |
| 2.2.3. La psicoacústica .....                                                                                                | 40 |
| 2.2.4. El sentido de aplicar la acústica al lenguaje audiovisual .....                                                       | 41 |
| 2.3. Del fenómeno acústico al sentido audiovisual .....                                                                      | 45 |
| 2.3.1. El sonido .....                                                                                                       | 45 |
| 2.3.2. Fuente sonora .....                                                                                                   | 46 |
| 2.3.3. Objeto sonoro .....                                                                                                   | 47 |
| 2.3.4. Ente acústico .....                                                                                                   | 48 |
| <br>                                                                                                                         |    |
| <b>3. Del análisis acústico a la sensación</b> .....                                                                         | 51 |
| 3.1. Entre el fenómeno vibratorio y el perceptivo .....                                                                      | 53 |
| 3.1.1. Sonidos puros y sonidos compuestos .....                                                                              | 53 |
| 3.1.2. La amplitud y la frecuencia en los sonidos puros .....                                                                | 53 |
| 3.2. Los instrumentos para analizar sonidos compuestos .....                                                                 | 63 |

Cubierta de Mario Eskenazi

1ª edición, 1998

Quedan rigurosamente prohibidas, sin la autorización escrita de los titulares del "Copyright", bajo las sanciones establecidas en las leyes, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamo públicos.

© de todas las ediciones en castellano,  
Ediciones Paidós Ibérica, S.A.,  
Mariano Cubí, 92 - 08021 Barcelona  
y Editorial Paidós, SAICF,  
Defensa, 599 - Buenos Aires

ISBN: 84-493-0479-2

Depósito legal: B-238/1998

Impreso en Gràfiques 92, S.A.,  
Av. Sucarrats, 91 - 08190 Rubí (Barcelona)

Impreso en España - Printed in Spain

|                                                                                    |     |
|------------------------------------------------------------------------------------|-----|
| 3.2.1. Sobre la concepción analítica de los sonidos compuestos...                  | 63  |
| 3.2.2. Los métodos gráficos de análisis acústico .....                             | 64  |
| 3.2.3. La interrelación entre los tres métodos clásicos de análisis acústico ..... | 73  |
| 3.3. El tono y el timbre en los sonidos compuestos .....                           | 79  |
| 3.3.1. La sensación de tono en los sonidos compuestos .....                        | 79  |
| 3.3.2. El concepto de timbre .....                                                 | 83  |
| 3.3.3. La sensación tímbrica .....                                                 | 85  |
| 3.3.4. Sobre la dimensión dinámica del timbre .....                                | 97  |
| 3.4. Las magnitudes físicas para la medición del sonido .....                      | 97  |
| 3.4.1. La medición del tono .....                                                  | 98  |
| 3.4.2. La medición de la intensidad .....                                          | 100 |
| <b>4. Fundamentos de la percepción sonora.</b> .....                               | 109 |
| 4.1. La subjetividad perceptiva como objeto científico .....                       | 111 |
| 4.1.1. En busca de mecanismos expresivos universales .....                         | 112 |
| 4.2. La percepción de las dimensiones sonoras simples .....                        | 112 |
| 4.2.1. El concepto de umbral .....                                                 | 112 |
| 4.2.2. Umbrales de intensidad .....                                                | 113 |
| 4.2.3. La sensación de intensidad en los sonidos compuestos .....                  | 114 |
| 4.2.4. Influencia de la distancia en la sensación de intensidad .....              | 116 |
| 4.2.5. Influencia de la intensidad en la sensación de distancia .....              | 117 |
| 4.2.6. Umbrales de tono .....                                                      | 120 |
| 4.2.7. Sensibilidad absoluta y sensibilidad relativa respecto al tono .....        | 121 |
| 4.2.8. Umbrales temporales .....                                                   | 122 |
| 4.3. La percepción de la complejidad sonora interna .....                          | 125 |
| 4.3.1. La percepción en bandas críticas .....                                      | 126 |
| 4.3.2. El enmascaramiento .....                                                    | 127 |
| 4.3.3. Los diagramas de Zwicker .....                                              | 128 |
| <b>5. Las formas del sonido</b> .....                                              | 137 |
| 5.1. La percepción de formas sonoras .....                                         | 139 |
| 5.1.1. El ruido .....                                                              | 141 |
| 5.1.2. El silencio .....                                                           | 148 |
| 5.1.3. El uso expresivo del efecto-silencio .....                                  | 152 |
| 5.1.4. Formas estacionarias y formas dinámicas .....                               | 155 |
| 5.2. Inercia acústica y discriminación entre formas sonoras .....                  | 158 |
| 5.2.1. Principio de la coherencia espectral .....                                  | 161 |
| 5.2.2. Principio de la estabilidad espectral .....                                 | 163 |
| 5.2.3. Principio de la estabilidad tonal .....                                     | 165 |
| 5.2.4. Principio de la regularidad .....                                           | 165 |
| 5.2.5. Principios de la sincronía y de la asincronía .....                         | 169 |
| 5.3. Taxonomía de las formas sonoras simples .....                                 | 171 |
| 5.3.1. Contorno y textura .....                                                    | 173 |
| 5.3.2. Clasificación según el inicio del contorno sonoro .....                     | 175 |

|                                                                         |     |
|-------------------------------------------------------------------------|-----|
| 5.3.3. Clasificación según el cuerpo del contorno sonoro .....          | 177 |
| 5.3.4. Clasificación según el final del contorno sonoro .....           | 183 |
| 5.3.5. La duración .....                                                | 184 |
| 5.3.6. Clasificación según la textura .....                             | 185 |
| 5.3.7. Cuadro global .....                                              | 189 |
| 5.4. Sobre las formas sonoras complejas .....                           | 190 |
| <b>6. De la forma sonora al sentido</b> .....                           | 193 |
| 6.1. El modelo semiótico .....                                          | 195 |
| 6.1.1. Signos sonoros motivados y signos sonoros arbitrarios .....      | 195 |
| 6.1.2. La construcción de sentido sonoro en el ámbito audiovisual ..... | 196 |
| 6.2. El modelo de los mecanismos de escucha .....                       | 198 |
| 6.2.1. Oír .....                                                        | 199 |
| 6.2.2. Escuchar .....                                                   | 200 |
| 6.2.3. Reconocer .....                                                  | 201 |
| 6.2.4. Comprender .....                                                 | 202 |
| 6.2.5. El aprendizaje en la construcción del sentido sonoro .....       | 203 |
| 6.2.6. Un modelo sobre dos ejes: el perceptivo y el cultural .....      | 205 |
| 6.3. Teoría de la coherencia perceptiva .....                           | 207 |
| 6.3.1. El sentido sonoro como construcción multidimensional .....       | 208 |
| 6.3.2. El sentido audiovisual y la coherencia perceptiva .....          | 210 |
| <b>7. El sonido en la narración audiovisual</b> .....                   | 217 |
| 7.1. Sobre la superioridad de la visión .....                           | 219 |
| 7.2. ¿Que aporta el sonido a lo audiovisual? .....                      | 221 |
| 7.3. El espacio sonoro .....                                            | 223 |
| 7.3.1. La necesidad de una teoría del espacio sonoro .....              | 225 |
| 7.3.2. El concepto de espacio sonoro .....                              | 227 |
| 7.3.3. El espacio sonoro en el contexto audiovisual .....               | 229 |
| 7.3.4. Acústica de la sensación de distancia .....                      | 230 |
| 7.3.5. Sobre el concepto de plano sonoro .....                          | 237 |
| 7.3.6. Movimientos en el espacio sonoro .....                           | 239 |
| 7.3.7. El efecto de profundidad o perspectiva .....                     | 242 |
| 7.3.8. Acústica del volumen espacial .....                              | 243 |
| 7.3.9. El punto de audición .....                                       | 247 |
| 7.4. El papel narrativo de la sincronía imagen-sonido .....             | 251 |
| 7.4.1. El concepto de sincronía .....                                   | 252 |
| 7.4.2. La sincronía como recurso narrativo .....                        | 254 |
| 7.4.3. Ritmo musical y movimiento visual .....                          | 257 |
| 7.4.4. Otras formas de relación sonido-imagen .....                     | 259 |
| 7.5. El sonido como instrumento organizador de la narración .....       | 260 |
| 7.6. Sobre el papel del habla en el lenguaje audiovisual .....          | 263 |



Conclusiones ..... 267

Referencias bibliográficas ..... 268

Bibliografía básica sobre el tema ..... 271

Para ampliar ..... 272

PRESENTACIÓN

El sonido es una vibración en el aire, un fenómeno físico...  
El sonido es un signo: suministra al oyente una información; agita su sistema nervioso y crea una emoción...  
El sonido es algo más que una voz encadenando signos lingüísticos...  
El sonido puede llegar a estimular nuestro sistema perceptivo sensorial con la misma fuerza y presencia que la imagen...  
¿El sonido es imagen...?

Los medios tradicionales de la comunicación audiovisual (radio, cine, televisión) y sus nuevos soportes informáticos (multimedia, internet) fundamentan una buena parte de su impacto comunicativo en la fuerza expresiva del sonido de sus mensajes audiovisuales. El sonido es todavía un manantial inagotable de significados... aunque muchas veces los creadores de la comunicación audiovisual a lo juzguen a *lo sonoro* una importancia relativa.

Porque la *intuición* no ha de ser la única fuerza motriz en la creación audiovisual, Ángel Rodríguez Bravo nos propone con esta obra un estudio sincero sobre la trascendencia expresiva y narrativa del sonido en el lenguaje audiovisual, descubriéndonos toda su enorme complejidad. Huyendo del *vococentrismo* y del reduccionismo lingüístico, el autor defiende un enfoque transdisciplinar para el estudio del sonido y apuesta por un inevitable esfuerzo colectivo: jactúricos, psicólogos, ingenieros... uníos, en el estudio de la comunicación audiovisual! Y sin complejos: quienes ocupan su actividad investigadora de forma exclusiva en el estudio de *lo sonoro* tienen los mismos derechos a la dignidad científica que aquellos que delimitan su objeto de estudio a la imagen icónica o *lo visual*.

El autor define y analiza la dimensión acústica del lenguaje audiovisual a partir de la percepción, desde la perspectiva del sujeto *oyente*, pero al mismo tiempo con un enfoque nada *comunicativista*, pues se aparta de las investigaciones determinadas por un «contexto comunicativo» particular para intentar establecer los límites entre *lo objetivo* y *lo subjetivo* en la formulación del *sentido* en el universo sonoro... a la búsqueda de esos «códigos universales» de la comunicación audiovisual: el estudio de los mecanismos de la interpretación del universo sonoro intentando superar los límites de los factores de percepción específicos de cada medio de comunicación. O dicho de otra forma: el objetivo de la presente obra es la formalización de una *gramática* de la expresión sonora que supere las fronteras comunicativas específicas de los distintos sujetos-oyentes de la comunicación audiovisual (el sujeto-oyente del cine, el sujeto-oyente de la televisión o el sujeto-oyente de la radio).

Estudiantes de la comunicación audiovisual y creadores de mensajes en los medios audiovisuales encontrarán en esta obra un sistema útil de comprensión del universo sonoro. Ahora bien, el conocimiento es sólo el primer paso y necesita fender algo mucho más importante: ¡apreciemos el sonido!... Es el alimento de nuestros oídos y también de nuestra sensibilidad.

ARMAND BALSEBRE  
Catedrático de Comunicación Audiovisual  
de la Universidad Autónoma de Barcelona

## AL LECTOR

### Estimado lector:

La obra que usted tiene en sus manos recoge y ordena un trabajo desarrollado a lo largo de unos doce años de estudio, centrados en la narración y la comprensión del mundo desde el sonido. Esta larga labor es el fruto de una primera etapa de puro interés y voluntad personal y, luego, del enorme estímulo que han supuesto para mí las ayudas recibidas de la Dirección General de Universidades, el Centre d'Investigació de la Comunicació de la Generalitat de Catalunya, y la Dirección General de Investigación Científica y Técnica (DGICYT: PB 94-0732), instituciones que consideraron que mi trabajo en este campo podía ser interesante y fértil. Gracias a estas ayudas he podido llevar a cabo la labor de estudio y de investigación que se presenta en este libro.

Esta obra es, de algún modo, el resultado de haber descubierto que la comunicación audiovisual es algo bastante más complicado que la pura intuición aplicada al manejo de la tecnología de las telecomunicaciones. La comunicación audiovisual es la técnica de engaño más compleja, más extraordinaria y más verosímil que se ha conseguido a lo largo de la historia de la humanidad. Su lenguaje trabaja con la propia esencia perceptiva de la realidad, capturando las informaciones sensoriales que emanan de los objetos, para componer con ellas narraciones que nos hacen oír y ver cosas que en ese momento y en ese lugar no existen, o que quizá no han existido ni existirán jamás, pero que percibimos como si fuesen la realidad misma. Ese fenómeno es de tal alcance humano, social y científico que rebasa violentamente los problemas de la ingeniería que les da el soporte tecnológico.

Este libro es un esforzado intento para dar algunos pasos firmes en la comprensión científica de los mecanismos que estructuran ese fenómeno de alucinación consciente que sufre el ser humano cuando oye y ve sonidos e imágenes artificiales; y responde a la clara convicción de que su estudio no debe quedar centrado solamente en la tecnología que apoya la producción audiovisual.

Es un trabajo que articula la *psicología de la percepción*, la *física acústica* y la *narrativa audiovisual*, procurando que actúen como un todo unívoco, capaz de proporcionar respuestas a aquellos que quieren comprender cómo funciona la interpretación del mundo desde el sonido. Y es, sobre todo, una obra dedicada a proporcionar instrumentos de trabajo útiles y eficaces a quienes pretenden narrar con el sonido y la imagen.

## INTRODUCCIÓN

Creo que es importante aclarar algunas premisas sobre la línea de pensamiento que estructura esta obra:

En primer lugar señalaré que entiendo el sonido como una parte integral e inseparable de la globalidad del **/lenguaje audiovisual/** y que ese criterio es el que guiará todo el estudio que el lector tiene en estos momentos en sus manos. No obstante, el nivel de desarrollo actual del conocimiento sobre lenguaje audiovisual todavía no permite una aproximación multidimensional y simultánea sonico-visual si ésta se desarrolla desde una metodología rigurosamente empírica.

Estas dos razones me han llevado a acotar el objeto de estudio **/lenguaje audiovisual/** centrándome en su dimensión auditiva, y, en consecuencia, a circunscribir esta obra al universo sonoro. Así, nos ocuparemos del sonido, de su percepción, y de la expresión a través de él. No obstante, el criterio de trabajo será siempre el de entender el audio como parte integrante del sistema global sonido-imagen; y las sensaciones sonoras como parte integrante del sistema global de la percepción.

En cualquier caso, el enfoque propuesto no responde exclusivamente a una reducción funcional y pragmática para resolver un objeto de estudio demasiado complejo, sino que es también fruto de una concepción reivindicativa del universo acústico como instrumento de la comunicación audiovisual.

Tradicionalmente, en el entorno de la comunicación de masas, el sonido ha sido relegado a un segundo plano frente a la imagen. Tanto en el ámbito productivo como en el universo de la investigación, se ha arrastrado al sonido como algo inevitable pero de segunda categoría frente a la imagen. Y, salvo en honrosas excepciones, se le ha prestado muy poco interés. Pienso que es importante intentar la inversión de esta tendencia. El universo sonoro es el ámbito en el que se produce la comunicación de las sensaciones más primarias, esenciales y difícilmente racionalizables que es capaz de expresar y percibir el ser humano. Pensemos, por ejemplo, en el papel fundamental del sonido en la comunicación emocional.

La segunda cuestión que considero importante señalar *a priori* es que este trabajo debe ser entendido sólo como un primer paso hacia el conocimiento global de los mecanismos que organizan la interpretación de los mensajes audiovisuales. Un primer paso consciente de los cientos de kilómetros que quedan por recorrer. El enfoque que presentaremos a lo largo de este libro será, en muchas ocasiones, extremadamente reduccionista. Pero la única forma de iniciar una labor sistemática de investigación empírica en el campo del lenguaje audiovisual es empezando el trabajo por el principio. Y creo que la base esencial del funcionamiento de ese

lenguaje depende de la relación que existe entre las dimensiones físicas de los mensajes y las formas percibidas por el ser humano que los interpreta. Sólo cuando podamos disponer de algunos conocimientos contrastados sobre la relación entre estos dos universos estaremos en condiciones de abrir una imprescindible nueva etapa de investigación; una nueva etapa cuyo objetivo sea ya gramaticalizar la influencia de las dimensiones cognoscitivas y contextuales del ser humano en la interpretación del discurso audiovisual.

Finalmente, quiero señalar que *La dimensión sonora del lenguaje audiovisual* no es un estudio sobre radio, sobre la banda sonora del cine o sobre el audio en televisión, sino que es una obra sobre el sonido como forma de expresión. Una investigación sobre las posibilidades expresivas del sonido en su sentido más amplio, tomando como referencia esencial los mecanismos perceptivos con independencia del medio que pueda vehicular el sonido en cada momento. Considero que éste es el planteamiento metodológico más fértil para avanzar en el conocimiento del lenguaje audiovisual en su acepción más amplia, compleja y profunda.

## Capítulo 1

### SOBRE LOS PLANTEAMIENTOS METODOLÓGICOS QUE ORIENTAN ESTA OBRA

*En este primer capítulo, antes de enfrentarnos directamente con los mecanismos de la expresión acústica, expondré los criterios epistemológicos y metodológicos desde los que abordaremos la comunicación audiovisual como objeto de estudio y como marco conceptual de trabajo.*

## 1.1. La subjetividad como problema metodológico en las corrientes clásicas de investigación sobre lenguaje audiovisual

La investigación sobre el lenguaje audiovisual en el ámbito de la comunicación de masas se ha desarrollado hasta ahora siguiendo dos grandes corrientes paralelas:

1. La primera parte de una perspectiva inspirada en el estructuralismo y apoyada básicamente en la lingüística y la semiótica. Esta corriente ensaya un modelo teórico que intenta dar respuesta global a los problemas narratológicos de la comunicación de masas y, consecuentemente, también de la comunicación audiovisual, desde todos los puntos de vista posibles. Roland Barthes ha sido probablemente el autor más paradigmático en los comienzos de esta línea de trabajo que se desarrolló en los años cuarenta y cincuenta.

2. La segunda establece unos planteamientos mucho más ligados a los problemas productivos y se apoya básicamente en el análisis de las posibilidades narrativas de cada medio partiendo del conocimiento de su tecnología. Esta segunda línea no trabaja con un paradigma globalizador sino que fragmenta su análisis en función de los distintos medios de comunicación (cine, radio, televisión y últimamente los llamados multimedia). Uno de los fundadores de esta corriente es sin duda S. M Eisenstein. Este autor llega al análisis conceptual de la comunicación audiovisual desde su experiencia como cineasta, y desarrolla sus reflexiones mas fértiles también en los años cuarenta.

Desde el punto de vista metodológico existen aspectos comunes entre las dos corrientes. Ambas líneas desarrollan su estudio sobre la eficacia comunicativa (el discurso audiovisual observando los productos audiovisuales y observando al emisor mismo: sus criterios, su contexto, sus orígenes, etc.; pero son escasísimas las investigaciones que observen las respuestas de grupos independientes de receptores expuestos a los mensajes. Y, también, cuando se enfrentan al problema de la descodificación de los mensajes, las dos líneas de trabajo lo han hecho utilizando sistemáticamente el método introspectivo.

El argumento tradicional y recurrente en contra de los resultados obtenidos a partir de estas dos grandes líneas de estudio ha sido el de *subjetividad*. Revisemos, pues, en qué se sustenta esta crítica histórica sobre la investigación en torno al lenguaje y a la comunicación audiovisual.

Al utilizar el método introspectivo, el investigador se toma a sí mismo como receptor-tipo. Si el estudioso es también autor y creador de productos audiovisuales

las, recurriendo a este método es el propio narrador de las historias el que se autoselecciona como receptor-tipo. Esto genera un bucle metodológico que retorna siempre al emisor, manteniéndose sistemáticamente muy lejos de la localización de los elementos generalizables.

Realmente, los modelos sobre el funcionamiento de los códigos audiovisuales están determinados por el punto de vista del emisor (director, narrador, realizador, escritor, guionista, compositor...) impregnando la mayor parte de la investigación sobre lenguajes. Y, puesto que estos modelos teóricos no intentan contrastar si los mecanismos comunicativos que el estudioso ha encontrado actúan igual sobre todos los receptores posibles, la investigación queda limitada metodológicamente al estudio de una sola parte del modelo clásico de la comunicación: el emisor.

Es cierto que la comunicación de masas es unidireccional y que el criterio de codificación narrativa y estética de un solo sujeto emisor se impone a un gran número de sujetos receptores, en consecuencia, esta perspectiva de estudio de lo audiovisual no deja de ser fértil y, por tanto, valiosa; no obstante, los resultados obtenidos en una investigación marcada sólo por el punto de vista del receptor son, efectivamente, *subjetivos* en tanto que no pueden hacerse extensibles a todo el universo posible de receptores, sino a un solo sujeto emisor.

En coherencia con esta reflexión, considero que es necesario que la investigación sobre el lenguaje audiovisual dirija su punto de vista hacia el receptor. Una forma de conseguir esto es que las investigaciones sobre lenguajes intenten responder sistemáticamente a la pregunta:

¿Es esta interpretación del mensaje extensible a todos los receptores posibles? ¿Por qué?

No es nuevo que frente a cualquier estímulo estético o narrativo los distintos receptores perciban sensaciones distintas, pero tampoco lo es que una parte de estas sensaciones son también coincidentes. Sin coincidencias de decodificación de todos los receptores frente a un mismo estímulo, la narración audiovisual no sería posible. Y cuando lo que nos preocupa son las coincidencias estamos dirigiendo nuestra mirada hacia los códigos universales.

Sin duda todo cineasta cuenta sus propios fantasmas en sus películas, pero no todos esos fantasmas se ven en la sala de proyección, cada espectador descubre unos u otros en la oscuridad del cine, pero siempre hay algunos de ellos que son vistos por todo el público, por todos los públicos de todas las salas, éstos son los universales. Ésos son los que han sido narrados utilizando los elementos esenciales del lenguaje audiovisual. A ese público no le preocupa cuáles son los fantasmas del autor, sólo le preocupan los que ha visto y ha sentido en la pantalla.

## 1.2. La alternativa de las metodologías cualitativas

La tendencia actual para resolver ese efecto de subjetividad de la investigación centrada en el emisor es la de estudiar toda la complejidad contextual que en-

vuelve la producción y la recepción de los mensajes (cultura, situación económica, hábitos de uso de los medios de comunicación, características sociales, historia...). Desde luego, en tanto que la información contextual sea rica estamos en mejores condiciones para intuir algunos de los elementos que influyen en la interpretación de los mensajes de determinados grupos. No obstante, entiendo que esta orientación de las llamadas metodologías cualitativas (véanse Jensen y Janowski, 1993) es, en cierto modo, contradictoria en sí misma. Si aceptamos que uno de los problemas esenciales de la investigación en comunicación de masas es la subjetividad de los resultados, pienso que esta nueva orientación metodológica vuelve a incurrir en el mismo tipo de error. Es cierto que supone un avance en tanto que con estos criterios no estudiamos ya emisores aislados sino grupos que comparten contextos, hábitos y culturas; no obstante las mismas características agrupadoras son, a la vez, elementos diferenciadores de gran trascendencia.

La introducción sistemática del contexto comunicacional en la investigación sobre lenguajes no nos lleva a códigos universales sino a una nueva compartimentación en grupos mayores pero también con diferencias mucho más acen- tuadas.

El conocimiento profundo de los hábitos de consumo comunicativo en la cultura japonesa nos puede ayudar a comprender los mecanismos que hacen que un ciudadano de Tokio interprete de distinto modo que uno de París determinada secuencia fílmica, pero seguimos sin saber cuáles son los elementos comunes a ambos, seguimos sin dominar las variables esenciales que ambos entenderían de manera idéntica al ver la misma secuencia cinematográfica. Entiendo que desde esta perspectiva contextualizadora que promueven las metodologías cualitativas, nuestra mirada contempla mucho mejor los elementos diferenciadores que los elementos comunes y, por tanto, sólo observa con eficacia a determinados grupos de receptores y no a todos los receptores posibles.

Probablemente la dispersión que existe hoy respecto a los métodos de estudio de la comunicación audiovisual en el ámbito de la comunicación de masas provenga de haberse iniciado su estudio desde la perspectiva de las ciencias sociales en lugar de hacerlo desde las ciencias físicas y humanas. Pienso que la sociología de la comunicación debería ser posterior al estudio de la psicofísica comunicativa. Uno de los objetivos esenciales del conocimiento científico es la previsión eficaz de los fenómenos a partir de informaciones previas, y difícilmente podemos prever las respuestas de todos los espectadores posibles de una serie televisiva norteamericana si no conocemos todavía, por ejemplo, cómo actúa la percepción del color en la comprensión de una narración; o cómo influyen los cambios de intensidad sonora en la construcción del sentido global de una secuencia, o cuáles son exactamente los mecanismos visuales y sonoros que determinan la sensación de segmentación y de coherencia narrativa.

No habría sido eficaz hacer sociología médica sin haber desarrollado antes la medicina, ni sociolingüística antes de que la lingüística se consolidase como ciencia de partida.



### 1.3. La percepción humana como punto de partida

Partamos de la idea de que dentro de los objetivos esenciales de la investigación sobre comunicación audiovisual ha de estar la localización de códigos universales.

Ciertamente, los criterios del emisor que codifica los mensajes no van a coincidir siempre con los del receptor que los interpreta. Tampoco el contexto cultural del equipo de producción que guioniza, interpreta y filma, pongamos por caso, una telenovela sudamericana va a corresponderse con el de sus espectadores europeos. ¿Dónde está entonces el territorio común? ¿Qué es lo que permite que determinados productos mediáticos den la vuelta al mundo sin problemas de descodificación? ¿Y qué diferencia a éstos de otros que no lo consiguen?

Una breve reflexión antropogenética puede orientarnos en la búsqueda de este tipo de respuestas.

En tanto que las características biológicas de todos los miembros de una misma especie de seres vivos determinan necesidades muy similares para su supervivencia, su fisiología establece también una forma homogénea de percibir e interpretar su entorno inmediato para poder relacionarse con él de acuerdo con esas necesidades. Consecuentemente, si entendemos que todos los seres humanos perciben los mismos estímulos físicos mediante los mismos mecanismos fisiológicos, y que estos mecanismos determinan la interpretación de cualquier variación del entorno próximo a nuestro cuerpo de acuerdo con nuestras características biológicas, uno de los puntos de partida esenciales en la investigación sobre comunicación ha de ser necesariamente la percepción humana.

Si la gama de frecuencias sonoras que percibe cualquier ser humano es prácticamente idéntica a la que perciben todos los demás y la gama de frecuencias lumbínicas también lo es, si el sistema receptor de información externa responde a las mismas necesidades biológicas, espaciales, cinéticas, alimenticias, táctiles, de temperatura, etc., no creo que incurramos en ningún error afirmando que es posible localizar variables universales en los mecanismos humanos de interpretación de los mensajes sonoros y visuales. De acuerdo con esta reflexión, esta obra toma como punto de partida la percepción del ser humano.

Es necesario destacar también que en este trabajo nos aproximaremos a la percepción contemplándola como un único sistema global. Siguiendo el mismo planteamiento antropogenético que expuse más arriba, la lógica perceptiva del ser humano no aísla el sonido de la luz, o del tacto; los objetos se ven y se oyen a la vez. Cuando entramos en una catedral, la miramos, la escuchamos, la olemos, nos movemos por ella, y todo ocurre simultáneamente.

### 1.4. ¿Existe un lenguaje audiovisual distinto para cada medio?

Si consideramos que lo esencial en el ámbito del conocimiento de los mecanismos que estructuran la narración audiovisual es la percepción, desde la pers-

pectiva del lenguaje audiovisual la lógica nos lleva inmediatamente a no separar radio, cine y televisión como objetos de estudio diferentes.

Expondré con más detenimiento en qué se fundamenta esta afirmación.

La idea de que las características de cada medio empujan al desarrollo y a la orientación de un lenguaje específico distinto tiene un origen esencialmente productivo. Que exista un lenguaje específico para cada medio audiovisual sólo es parcialmente cierto si nos enfrentamos al medio desde la perspectiva del emisor. O, más concretamente, desde la perspectiva del realizador o del director. Es decir, de aquel que ha de manejar la tecnología del medio para construir físicamente los productos narrativos. Utilizar soporte químico o soporte magnético para una grabación plantea sin ninguna duda dos técnicas de trabajo radicalmente distintas. Pero éste sólo es un problema fundamental para aquel que produce los mensajes, y no lo es en absoluto para el que los recibe.

El receptor se limita a ver y a escuchar, no necesita aprender previamente ningún código complejo. O, en todo caso, este aprendizaje se realiza de un modo inconsciente. En cambio el emisor sí ha de enfrentarse constantemente al problema de cómo hacer inteligible un mensaje para cualquier espectador a partir del conjunto de posibilidades y de limitaciones expresivas del medio con el que está trabajando. En consecuencia, el emisor (director, realizador, guionista) necesita aprender en profundidad el *lenguaje del medio*. Se articulan entonces una serie de convenciones que responden esencialmente a las necesidades de producción narrativa de cada medio de comunicación en concreto. Convenciones que no suelen preocupar para nada al receptor, que sencillamente se expone al mensaje que le gusta o no, y lo entiende o no lo entiende.

Si nos aproximamos al efecto generado por los distintos medios audiovisuales desde el lugar de quien contempla los resultados, es decir, desde la perspectiva del espectador, la necesidad que tiene un realizador de resolver las dificultades narrativas mediante técnicas diferentes se disuelve en los problemas perceptivos hasta desaparecer. Es cierto que la definición de la imagen o el tamaño de la pantalla plantean diferencias importantes entre un medio y otro, pero la misma trasposición del cine a la televisión, o de las bandas sonoras musicales fílmicas al mercado discográfico, acota perfectamente la cuestión de la diferenciación entre los medios como un problema esencialmente productivo.

En suma, pienso que sólo tiene sentido hablar de *un lenguaje específico de cada medio* desde el punto de vista del emisor, pero no desde el punto de vista del receptor.

Si nuestro objetivo es la localización de rasgos esenciales, el norte que ha de orientar la navegación hacia el conocimiento del lenguaje audiovisual, todavía en proceso de elaboración, es el estudio de las respuestas de la percepción humana.

El ser humano usa códigos distintos para interpretar flujos de información de orígenes diferentes. Pero para hacer eso no altera en absoluto su base perceptiva. Provienga la información de la televisión, de la radio o de otro ser humano, sigue escuchando con el mismo sistema auditivo y sigue mirando con el mismo sistema visual. Varía el tipo de empaquetamiento informativo en función del medio utili-

zado, pero jamás cambia el sistema de reconocimiento de formas del público receptor.

Cualquier mensaje se constituye a partir de las perturbaciones organizadas de un determinado medio físico (presión sobre el aire, movimientos de las formas visuales, variaciones lumínicas, etc.) que son percibidas e interpretadas por el ser humano. Así, a determinado orden de perturbaciones físicas percibidas corresponden determinadas sensaciones. A continuación interpretamos esta información procesándola con todos los códigos de que disponemos que sean pertinentes y resulten útiles para hacerlo. El oído actúa para todos los sonidos del mismo modo, sin diferenciar si su origen es radiofónico, televisivo o cinematográfico.

La clave del conocimiento de los lenguajes está, entonces, en qué relaciones se establecen entre las perturbaciones físicas del medio percibidas y la interpretación que hace el hombre de ellas.

Es cierto que desde el punto de vista productivo los conocimientos que precisa un realizador de televisión o un director de cine son distintos, especialmente los tecnológicos, pero también es cierto que estos dos roles profesionales se intercambian con muchísima frecuencia. Pienso, en consecuencia, que es importante no confundir el lenguaje audiovisual, con la técnica y la tecnología de los medios de comunicación.

No confundimos el lenguaje musical, que es la forma de estudio y de codificación de la música, con la música como composición estética de formas sonoras. Ni confundimos tampoco el lenguaje musical con la técnica de uso de cada uno de los instrumentos que la producen. El lenguaje musical es el mismo para todos los instrumentos a pesar de que la técnica para tocar el violín tenga muy poco que ver con la del oboe, la del piano, o la de los timbales. El lenguaje musical es un código estructurado que se ajusta con precisión a los mecanismos perceptivos y fisiológicos del ser humano. Las escalas tonales se organizan en octavas, es decir, lógicamente, de acuerdo con nuestros mecanismos de percepción del tono. El tiempo musical se estructura en impulsos rítmicos que se ajustan perfectamente a la duración de los movimientos más habituales del ser humano. Las estructuras del compás son esencialmente binarias, adaptándose a las combinaciones que es posible realizar golpeando con los pies o las manos. Las pautas de velocidad en la sucesión de los sonidos musicales (alegro, vivace, lento, adagio...) responden perfectamente a las cadencias que pueden generar los pasos de una persona alterando la velocidad de su marcha.

Del mismo modo, no debemos confundir el lenguaje audiovisual, como estructura organizativa conjunta de la imagen y el sonido que persigue la comunicación eficaz, con la técnica de cada medio audiovisual. Mientras la técnica del medio intenta resolver los problemas derivados de las limitaciones tecnológicas para organizar imágenes y sonidos, el lenguaje audiovisual responde a la capacidad del ser humano para entender las composiciones audiovisuales. La radio, el montaje con imágenes estáticas y sonido, el cine, la televisión, la multivisión, el multimedia..., son distintos instrumentos para tocar las posibles composiciones audiovisuales que es capaz de comprender el ser humano. En cambio el lenguaje

audiovisual es el conjunto de los modos de organización artificial de la imagen y el sonido que utilizamos para transmitir ideas o sensaciones, ajustándonos a la capacidad del hombre para percibir las y comprenderlas.

En consecuencia con todo lo expuesto un poco más arriba, no definiré mi objeto de estudio como es tradicional en la literatura sobre comunicación audiovisual acotando un medio en concreto, sino que partiré de la perspectiva del receptor para estudiar los mecanismos de la expresión/percepción sonora como un objeto de estudio único que alcance y abarque a cualquier medio de comunicación que utilice el sonido como instrumento expresivo.

## 1.5. El lenguaje audiovisual como objeto de estudio

### 1.5.1. Definición y características esenciales

En tanto que la revisión y las nuevas acotaciones que hemos propuesto sobre la comunicación audiovisual la sitúan muy cerca de los objetos de estudio clásicos de la psicología de la percepción, creo que es necesario desarrollar una reflexión sobre las características específicas del lenguaje audiovisual como objeto de estudio.

Pienso que para avanzar en el conocimiento de la comunicación audiovisual el objeto de estudio central ha de ser el lenguaje audiovisual entendiéndolo como *los modos artificiales de organización de la imagen y el sonido que utilizamos para transmitir ideas o sensaciones, ajustándonos a la capacidad del hombre para percibir las y comprenderlas* y no cada uno de los medios tecnológicos que utilizan ese lenguaje. Para estudiar este lenguaje, uno de los instrumentos científicos imprescindibles es la psicología de la percepción. No obstante, el lenguaje audiovisual configura un objeto de estudio más complejo que la percepción humana en tanto que la incluye parcialmente en algunos aspectos y la rebusa en otros.

La psicología de la percepción se preocupa de cómo el ser humano percibe e interpreta la información del entorno, o del propio organismo (Lieury, 1992, pág. 31), o, más concretamente, de cómo el hombre, de toda la información disponible, recoge únicamente aquella que es importante para su supervivencia dentro de su ecosistema (Guski, 1992, pág. 7).

Evidentemente, si nuestro objeto de estudio es el lenguaje audiovisual nos preocupan los mecanismos de recogida de información del entorno con que actúa el ser humano. Pero no podemos decir que la mayor parte de la información generada por los medios de comunicación audiovisual sea precisamente una información esencial para la supervivencia del hombre. A pesar de que una de sus características principales es su similitud formal con la información natural, es decir, con la información generada por el entorno sin ninguna intervención consciente del hombre; la información transmitida mediante el lenguaje audiovisual solamente simula de forma artificial las perturbaciones físicas del entorno humano que originariamente eran fundamentales para la supervivencia humana. Y esta si-

mulación está hecha artificialmente de una forma consciente y voluntaria por otra u otras personas.

El lenguaje audiovisual es un conjunto sistematizado y gramaticalizado de recursos expresivos que han sido siempre previamente imaginados por un narrador, y que permiten estimular en el público series organizadas de sensaciones y percepciones que se transformarán en mensajes concretos y complejos. Esto supone que el fundamento mismo de este lenguaje es el conocimiento de un sistema de recursos narrativos artificiales que permiten emular a voluntad del narrador muchas de las sensaciones que suele producir el entorno natural del hombre. Por otra parte, el lenguaje audiovisual, desde su base expresiva esencialmente perceptivo-naturalista, configura un complejo entramado en el que convergen la música y la lengua (tanto la oral como la escrita) con toda la cultura iconográfica, literaria y dramática de la civilización actual. Así, dentro del lenguaje audiovisual se articulan perfectamente la lengua y la música como sistemas de códigos completos que se entrelazan con las simulaciones perceptivas naturalistas características del dibujo, la pintura, la fotografía, los montajes de imagen fija y sonido, el cine, la radio, la televisión, etc., transfiriéndole su propia capacidad expresiva.

Pienso, en definitiva, que las tres características esenciales que dan una entidad propia y específica al lenguaje audiovisual como objeto de estudio son las siguientes:

1. El hecho de que exista siempre voluntad previa por parte de unos emisores para estimular en otras personas series organizadas de percepciones naturalistas simuladas.
2. Su capacidad de generar artificialmente mensajes que estimulan sobre el sistema sensorial del hombre percepciones muy similares a las que producen las informaciones de origen natural.
3. Su capacidad de articular dentro de él cualquier otro lenguaje basado en la percepción humana.

Estos atributos del lenguaje audiovisual son los que le dan especificidad y lo diferencian tanto de otros lenguajes, por ejemplo el del habla o el de la música, como de las informaciones indexativas de origen natural.

Y pienso, también, que el conocimiento profundo de la percepción humana es un punto de partida imprescindible para el conocimiento de lo audiovisual; pero hay que insistir en que solamente es el punto de partida. La psicología de la percepción difícilmente puede ser el instrumento exclusivo para dar cuenta del complicado tejido cultural que queda tramado en cualquier producción cinematográfica o televisiva. Entiendo que la psicología de la percepción y la física estrechamente vinculada a ella han de ser las disciplinas básicas de apoyo para una larga etapa de contrastación experimental que permitirá asentar el conocimiento sólido y bien contrastado de las leyes básicas que ordenan la comprensión de los mensajes audiovisuales (véase Rodríguez Bravo, 1994, págs. 150-171).

Y, por fin, creo firmemente que sólo después de haber superado esta etapa estaremos en condiciones de comenzar a deshilar con garantías de éxito en el com-

plejísimo entramado cultural y cognoscitivo que supone la articulación audiovisual de múltiples códigos complejos.

### 1.5.2. *Sobre la dimensión naturalista del lenguaje audiovisual*

El lenguaje audiovisual se construye sobre dos paradojas especialmente interesantes.

Desde el principio de la historia, el hombre está desarrollando formas de expresión como la pintura, la escultura y el teatro, basadas en la imitación del entorno mediante medios artificiales, que han ido evolucionando con el desarrollo del conocimiento técnico y tecnológico. Estas formas de expresión naturalista son especialmente simples desde el punto de vista perceptivo. En tanto que la percepción de los mensajes artificiales se parezca cada vez más a la percepción de la realidad misma es mucho más fácil para el receptor decodificarlos y comprenderlos. A medida que el conocimiento técnico se va haciendo más rico y complejo, la capacidad para desarrollar mensajes cada vez más similares a la realidad misma aumenta y, como consecuencia directa de esto, los mensajes producidos son cada vez más fáciles de interpretar, son más simples y universales.

Puesto que lo que se persigue mediante el lenguaje audiovisual es transmitir mensajes articulando estímulos que el receptor ya conoce porque existen en el medio natural, el pintor-escultor-radiofonista-cineasta... intenta trabajar cada vez con mayor completitud perceptiva en sus mensajes para que la necesidad de articulación de lenguajes complejos sea cada vez menor.

Cuanto más abstracto es un lenguaje mayor es su dificultad de comprensión para el receptor, que necesita aprender a dominar estos códigos complejos. Entonces, la alternativa de lo audiovisual es facilitar la tarea al receptor simplificando los códigos al usar organizaciones expresivas que generan artificialmente estímulos muy parecidos a los naturales. Utilizando distintas técnicas de imitación, el lenguaje audiovisual copia, con mayor o menor fortuna, fragmentos de la realidad misma y los recompone para transmitir información de la forma más simple y universal posible. La primera de las paradojas que mencionábamos más arriba aparece con la propia capacidad del lenguaje audiovisual de engullir esos lenguajes complejos de los que intentaba huir. Los lenguajes apoyados en códigos abstractos generan también formas físicas objetivas que pasan a formar parte del universo real que puede ser percibido por el hombre. Y, consecuentemente, las técnicas de imitación y recomposición de fragmentos de la realidad física, que utiliza el lenguaje audiovisual, le permiten también copiar las formas físicas generadas por los lenguajes abstractos. El resultado es una integración perfecta entre el lenguaje naturalista original de la comunicación audiovisual y otros lenguajes abstractos como la lengua o la música, contruidos a partir de códigos comunicativos mucho más complicados y arbitrarios.

Cuando las técnicas para imitar la realidad en las que se basa lo audiovisual permiten reproducir también las formas físicas de los lenguajes más complejos, la fuerza comunicativa global se multiplica al combinar las posibilidades de la ex-

presión naturalista con la capacidad de expresión abstracta y conceptual. El resultado es un salto cualitativo enorme, y en cierto modo casual, en las posibilidades comunicativas del lenguaje audiovisual, que con la evolución tecnológica adquiere unas posibilidades de combinación expresiva extraordinarias y en muchos aspectos todavía por explorar.

### 5.3. Sobre la naturaleza incompleta del lenguaje audiovisual

Escribía más arriba que una de las características esenciales que definen al lenguaje audiovisual como objeto de estudio es su dimensión naturalista y su complejidad perceptiva. Dicho de otra forma, su capacidad de vehicular narraciones mediante mensajes dirigidos simultáneamente a varios sentidos, de modo que se perciban de forma muy similar a cómo son percibidos los fenómenos naturales. Esta característica es la segunda de las paradojas sobre las que se construye el lenguaje audiovisual. En realidad, la lógica de este lenguaje no está sólo determinada por el hecho mismo de la similitud formal entre el mensaje y el referente, es decir, entre el mensaje y la realidad a la que éste se refiere, sino que se fundamenta también, paralelamente, en las diferencias reales que existen entre ambos.

Intentaré explicar esto con mayor precisión.

La secuencia cinematográfica de un caballo negro galopando por la playa es capaz de transmitirnos sensaciones que tienen una similitud incuestionable con las que habríamos percibido si hubiésemos estado realmente en esa misma playa el día en que se filmó la secuencia. Pero también es incuestionable que existe una distancia importante entre ambas percepciones. La pantalla es bidimensional, la imagen encuadrada por la pantalla sólo nos muestra una pequeña parte del espacio o que hubiésemos visto, el mar no huele, los altavoces en lugar de proyectar el sonido de las olas y el golpeo de las pezuñas del caballo sobre la arena emiten una música sugerente... Las diferencias entre las percepciones que tendríamos frente a la realidad referencial y las que tenemos frente a la realidad virtual cinematográfica exigen al cineasta la invención de una serie de recursos narrativos que le ayuden a suplir las deficiencias que comporta su material de trabajo. Y es, precisamente, sobre estos recursos para suplir diferencias y, a la vez, sobre la similitud objetiva que existe entre realidad virtual y realidad referencial, donde se sustenta la esencia del lenguaje audiovisual. Es el entrelazado de ambos fenómenos: *similitud* y *diferencia* respecto a la realidad lo que da sentido a los códigos narrativos del lenguaje audiovisual.

Veamos ahora este mismo ejemplo desde su perspectiva sonora.

Los micrófonos son instrumentos capaces de realizar la transducción de una señal acústica transformándola en señal eléctrica de un modo relativamente similar a como lo haría el oído para enviar información a nuestro cerebro. No obstante, la sensibilidad de la membrana de un micro suele ser bastante más limitada que la de la cóclea del oído humano. Un micrófono de gran fidelidad, es decir, lo que técnicamente se denomina cómo un micrófono de *respuesta plana*, es capaz de captar una gama de frecuencias muy similar a la del oído humano,

sin potenciar ni atenuar ninguna de las zonas del espectro sonoro. No obstante, cualquier micro es siempre mucho más dependiente de la distancia respecto a la fuente sonora que el oído humano. Cuando la fuente sonora se aleja unos pocos centímetros de la zona de captación óptima del micro, el sonido captado se altera, dando como resultado un cambio importante de su calidad al ser reproducido.

Pero volvamos de nuevo a nuestro caballo galopando por la playa. El oído humano es perfectamente capaz de estar percibiendo, a la vez, el sonido de las olas de la playa, el chapoteo de las patas en el agua, o el golpeo sordo y rítmico de los cascos sobre la arena. Y todo esto lo podemos escuchar sin ningún problema aunque estemos a cien o doscientos metros del caballo y aun cuando éste se acerca y se aleja de nosotros. Sin embargo, la captación electrónica de este tipo de paisaje sonoro es bastante más complicada.

Invitamos ahora al lector a que nos siga en una reflexión sobre los problemas que debió de tener el realizador cuando intentaba contar acústicamente las sensaciones que experimentó al ver y oír el caballo en la playa.

Sin duda, cualquier micrófono de alta fidelidad colocado a 50 centímetros de las pezuñas del corcel captaría perfectamente el ruido de los cascos sobre la arena, pero hacer esto con un caballo al galope no es un tarea sencilla. Si nuestro cineasta optaba por seguir al caballo con un vehículo desde el cual se haría la toma de sonido, esto iba a suponer una situación en la que el vehículo generaría mucho más ruido que el caballo mismo. Por otra parte, en el momento en que, este tipo de micrófono se aleja de la fuente sonora el resultado de la grabación es ya muy deficiente, por lo que tampoco se podría hacer la toma de sonido desde un punto fijo. Otra opción era hacer la toma de sonido desde lejos con un micro *de cañón* (altamente direccional). Esta solución supondría distorsionar el sonido perdiendo todas las frecuencias graves, con lo que el ruido del galope iba a resultar muy poco natural. En cualquier caso, en ninguna de estas dos situaciones era posible conseguir también, a la vez, una grabación satisfactoria del rumor de las olas. Este sonido habría que grabarlo independientemente y mezclar ambos después en el estudio de audio.

Posiblemente, la mejor forma de solucionar el problema iba a ser haciendo directamente una sonorización artificial que mezclase en el estudio algunos de los sonidos pregrabados que hay disponibles en las colecciones de efectos sonoros. Naturalmente, el resultado de esto sólo se parecería remotamente al paisaje sonoro original que envolvía al corcel negro mientras nuestro realizador lo contemplaba en la playa.

No se adoptó ninguna de las tres soluciones anteriores. Más arriba hemos dicho ya que mientras el caballo galopaba por la playa, en lugar del paisaje sonoro natural, sonaba una música sugerente. Obviamente, nuestro amigo realizador resolvió, finalmente, que la asociación entre una música determinada y las imágenes del caballo transmitiría mucho mejor las sensaciones vividas en la playa, que una recomposición artificial hecha con sonidos pregrabados. Las diferencias entre la percepción de la situación real y el resultado que era posible obtener copiando técnicamente la realidad eran demasiado grandes, así que se optó por una forma



de expresión mucho más abstracta para conseguir el efecto narrativo que deseaba nuestro realizador.

En suma, ese juego articulado de recursos expresivos que dan soluciones narrativas a la incompletitud audiovisual para resolver las diferencias que existen entre realidad referencial y realidad reproducida audiovisualmente es el que configura el lenguaje audiovisual. Así, toda narración audiovisual se apoya siempre en un equilibrio constante entre:

- a) la similitud naturalista del mensaje con lo contado;
- b) los recursos expresivos que escamotean su verdadera naturaleza de copia incompleta.

No obstante, una de las características más interesantes del lenguaje audiovisual es la posibilidad de transformar su propia incompletitud y sus propias dificultades técnicas en instrumentos narrativos útiles.

La sensación auditiva que suele producir el mal uso de un micrófono es una fuerte exageración de la sensación de distancia. Pensemos en lo que ocurre, por ejemplo, cuando un locutor de radio poco experimentado se balancea frente al micro siguiendo el ritmo de su discurso, al desplazarse tan sólo unos veinte centímetros de la zona de captación óptima, el oyente, que recibe la voz del locutor ya amplificada de nuevo por su receptor de radio, tiene la sensación de que éste se ha levantado de la mesa y va y viene por el estudio alejándose y acercándose unos dos o tres metros del micro. Y al contrario, si el locutor se acerca tan sólo unos diez centímetros más a la membrana del micrófono, el oyente va a tener en su casa la sensación de que quien habla prácticamente se ha echado encima de él y habla junto a su oreja.

Este fenómeno de percepción natural alterada por el medio técnico puede ser interpretado y utilizado por el emisor de dos formas muy distintas.

La primera consiste en entender que el micro debe manejarse de modo que el sonido que llega a través del receptor de radio a los oyentes sea lo más parecido posible al que se está produciendo en el punto de emisión. Es decir, que lo que suena en el estudio se parezca lo más posible a aquello que escucharán los oyentes en su casa. Este planteamiento estrictamente naturalista supone usar el micrófono exclusivamente como un medio para transportar sonido de un lugar a otro y supone, fundamentalmente, la preocupación por que la separación locutor-micro se mantenga constantemente a la distancia óptima.

La segunda interpretación, que podríamos denominar expresionista, consiste en considerar las diferencias entre sonido original y sonido captado por el micro como un instrumento expresivo. Es decir, prescindiendo de si lo que suena en el estudio y lo que emite el receptor radiofónico son sonidos distintos o no y preocupándonos exclusivamente de cómo va a interpretar el radioyente aquello que escucha. El locutor puede transformar estas alteraciones sobre la percepción natural que genera el movimiento frente al micrófono en un recurso para reforzar, por ejemplo, la sensación de proximidad física entre él y el locutor. Para conseguirlo, simplemente tendrá que acercarse unos centímetros al micrófono dismi-

nuendo conscientemente la intensidad de su voz, tal como lo haría si se acercase realmente a alguien para hablarle cerca del oído. El resultado de la combinatoria entre el efecto de la distorsión acústica producida al acercarse al micro y su utilización racional por parte del locutor da como resultado una percepción, por parte del oyente, que es muy similar a la que produce el sonido de la voz en una situación natural de relación íntima a distancia de contacto físico entre los interlocutores.

Lo que en principio era un problema de diferencia entre percepción natural y percepción mediática, puede ser transformado por el emisor en un recurso narrativo que permite recrear acústicamente una situación virtual de relación íntima olocutor-oyente.

Todos estos ejemplos muestran que *los códigos narrativos del lenguaje audiovisual se configuran como instrumentos orientados a solucionar los problemas de interpretación, derivados de la diferencia que existe entre la percepción directa de la realidad referencial y la percepción de las formas audiovisuales que la representan*. Así, lo que se desprende de esta ejemplificación es que la dimensión naturalista con la que trabaja el lenguaje audiovisual y las diferencias objetivas que existen entre percepción natural y percepción mediática son cosas facetas inseparables, que es imprescindible tener presentes cuando intentamos descubrir las leyes que estructuran el lenguaje audiovisual y los códigos que lo organizan.

## 1.6. De la física a la percepción de los mensajes audiovisuales

De la reflexión desarrollada en el apartado 1.5.1. se desprende que *la función básica del lenguaje audiovisual es simular de forma artificial las perturbaciones físicas naturales del entorno humano para transmitir voluntariamente ideas o sensaciones*. Y que, necesariamente, estas perturbaciones artificiales tienen que adaptarse a las capacidades de la percepción humana.

Sabemos que la percepción humana es fundamentalmente un proceso de captación-interpretación de las perturbaciones físicas del entorno. Y que la captación audiovisual introduce, como elemento nuevo en este sistema, perturbaciones físicas artificiales que simulan voluntariamente a las naturales.

La psicología de la percepción se ocupa de estudiar los mecanismos de captación y de interpretación de las perturbaciones físicas que envuelven al ser humano prescindiendo de estudiar la organización de las perturbaciones mismas. Simplemente asume que éstas existen e investiga cómo son recogidas y comprendidas por los sujetos.

Puesto que la comunicación audiovisual modifica la esencia misma de la percepción sustituyendo las perturbaciones originales del entorno por otras producidas y manipuladas de forma artificial, un modelo que intente dar cuenta del lenguaje audiovisual tiene necesariamente que investigar también las formas físicas de esos estímulos artificiales para conectarlos con las interpretaciones que son capaces de generar.

El lenguaje audiovisual nos permite manipular las formas físicas sonoras y visuales que percibe el ser humano, a través de sus sentidos, de un modo concreto para estimular unas interpretaciones concretas, desencadenando en el receptor sensaciones realistas controladas por el emisor. En consecuencia, considero que la fenomenología de la comunicación audiovisual debe ser estudiada desde un paradigma que incluya y rebase la psicología de la percepción, añadiendo a ésta el análisis físico de las formas visuales y sonoras artificiales creadas expresamente para la comunicación. Pienso que la clave del lenguaje audiovisual está en el descubrimiento progresivo de las relaciones que existen entre las series de perturbaciones físicas acústicas y lumínicas, creadas artificialmente por un narrador, y la interpretación de estas perturbaciones que hacen los receptores.

En coherencia con esta afirmación, este paradigma físico-perceptivo será el que orientará a partir de ahora todo el desarrollo de la obra que el lector tiene en sus manos.

## Capítulo 2

# LA ACÚSTICA Y LA COMUNICACIÓN AUDIOVISUAL

*En este capítulo haremos una reflexión sobre el sonido como un fenómeno con autonomía y entidad propia dentro de la comunicación de masas. Analizaremos, también, la necesidad de recurrir a la acústica para mejorar la eficacia narrativa de las producciones audiovisuales y para desarrollar el conocimiento sobre las leyes que ordenan el lenguaje audiovisual.*

## 2.1. La acusmatización

### 2.1.1. Desarrollo de un nuevo concepto

El concepto *acusmatización* tiene su origen en una técnica pedagógica utilizada por Platón para incrementar la efectividad de las enseñanzas que impartía a sus discípulos. El ilustre sabio griego hizo que sus alumnos le escucharan tras una cortina mientras hablaba para que así el contenido de sus discursos adquiriera toda la fuerza posible al desvincularse de su propia imagen. A los discípulos que escucharon las lecciones del maestro en esta situación durante cinco años se les denominó *acusmáticos*. Posiblemente el sabio descubrió que, a menudo, sus oyentes prestaban más atención a su aspecto y a su gesticulación que a aquello que intentaba contar, así que decidió atajar de raíz el problema desvinculando lo narrado de la fuente física que lo narraba.

La consecuencia de este origen ha sido que el término *acusmático* haya pasado a ser utilizado para denominar aquello que se oye sin ver la fuente de donde proviene.

Hasta que se desarrollaron los sistemas de grabación y de radiotransmisión de los sonidos, para que un sonido fuese *acusmático* el objeto físico que lo generaba podía estar escondido a la visión del oyente, pero nunca alejado de él. Sólo es posible oír aquello que está suficientemente cerca del receptor. Pero a partir de momento en que T. Edison inventa el primer gramófono en 1877 y que el físico canadiense Reginald A. Fessenden consigue en 1900 realizar la primera transmisión de voz humana a través de la radio (Franquet y Martí, 1985, págs. 20-21), el fenómeno acusmático adquiere una dimensión radicalmente nueva. Con la evolución de la tecnología del audio, el objeto original productor del sonido ya no necesita esconderse de la visión del oyente sino que realmente desaparece; ya no tiene que coincidir con el receptor ni en el espacio ni en el tiempo. El sonido puede ser ahora reproducido por un aparato (un gramófono o un receptor de radio) que nada tiene que ver con aquello que lo generó; la fuente sonora inicial deja de tener el valor físico sustancial y este valor pasa al sonido mismo que se hace independiente de su origen natural.

Esta independencia física que la tecnología del audio otorga al sonido ha pasado también a ser independencia semiótica. Actualmente, escuchar, por ejemplo, el sonido de una melodía de violín no es ningún índice de que en nuestro entorno inmediato haya una persona haciendo sonar un instrumento musical. Ese sonido puede estar expresando que cerca de nosotros alguien escucha su equipo de alta fidelidad; que en esos momentos, en algún lugar del mundo, un violinista da un concierto que nos llega transmitido en directo a través de la radio; que la pareja de protagonistas de una telenovela por fin han llegado a comprenderse mutuamente, etc.



La *acusmatización* aísla los objetos sonoros y los convierte en portadores de conceptos. Ahora, con mucha frecuencia, el sonido prescinde de su fuente y se conecta con un sentido nuevo que ya no tiene nada que ver con su origen directo sino con su forma sonora y con su situación en el contexto audiovisual. Todo esto nos lleva a que en el concepto actual de *acusmatización* debamos incluir también esta nueva vertiente de independencia física entre el sonido y su ente productor original. Pierre Shaeffer (Shaeffer, 1988) y Michel Chion (Chion, 1982) han utilizado con profusión el concepto con este nuevo sentido para investigar los mecanismos expresivos del universo sonoro.

### 2.1.2. La *acusmatización* en la comunicación de masas

Sin duda, uno de los fenómenos más importantes y de más trascendencia social en la evolución de la moderna comunicación de masas ha sido la *acusmatización*. La posibilidad que proporciona la tecnología del audio, de separar el sonido de la fuente sonora original y situarlo a voluntad del narrador en cualquier otro tiempo y lugar espacial, ha abierto una larga serie de nuevas posibilidades. Pensemos por ejemplo en la radiodifusión, en el cine sonoro, en la industria discográfica y en su complementaria: la del audio doméstico de alta fidelidad; en la mal llamada televisión (que en justicia debiera llamarse *tele-audio-visión*), en la industria de la música ambiental, en las reconstrucciones escenográfico-audio-visuales de los museos modernos; en la sonorización electrónica de las salas de fiestas, las salas de conferencias, los teatros... Todas estas formas de comunicación colectiva se apoyan de una manera absolutamente fundamental en la desvinculación entre el ente físico que actuó como primer generador del sonido y el sonido mismo, que ha pasado a ser grabado, filtrado, reordenado, empaquetado, transportado, vendido, comprado, amplificado y reescuchado, una o mil veces, como algo completamente independiente de su producción inicial.

Paralela y directamente vinculadas a todas estas nuevas formas de comunicación de masas, la posibilidad de trabajar con el sonido aislado permite a los narradores audiovisuales establecer nuevas asociaciones virtuales entre sonidos e imágenes que no existen en el universo referencial. Esta nueva forma de trabajar vinculada al tratamiento tecnológico de la *acusmatización* ha abierto un universo expresivo revolucionario dentro de la comunicación audiovisual.

Revisemos algunas de las opciones narrativas que ha desencadenado la posibilidad de *acusmatizar* el sonido:

- *El doblaje*, por ejemplo, ha permitido la construcción de nuevos personajes partiendo de una recomposición del par voz-imagen y posibilita mezclar los rasgos físicos y de expresión gestual de un actor con la capacidad de expresión oral y sonora de otro. Un caso paradigmático de este tipo de uso del doblaje en español es el de la asociación de las imágenes del actor americano Clint Eastwood, que es propietario de una voz considerablemente aguda, con la voz del actor español Constantino Romero, poseedor de una

voz sumamente densa y grave, construyéndose así un nuevo personaje que sólo existe en el universo audiovisual, mucho más «duro» que el del original estadounidense. Aquí los directores de doblaje españoles no han optado por imitar el modelo original eligiendo una voz similar a la de Eastwood, sino que han preferido crear un ente dramático completamente nuevo a partir de los materiales audiovisuales de que disponían. Este tipo de recursos permite también la creación de personajes mágicos o monstruosos manipulando técnicamente la voz de los actores o sustituyéndola por determinados efectos sonoros. Pensemos por ejemplo en animales parlantes, en mujeres poseídas diabólicamente que hablan con voces masculinas fortísimas y rotas, etc.

- *La ambientación musical* en el cine, la televisión o el teatro ha permitido asociar melodías y ritmos instrumentales a situaciones en las que desde un punto de vista realista sería absurdo que sonaran allí. Imagine el lector, por ejemplo, la secuencia fílmica de un jinete solitario que cabalga sobre un campo por el desierto mientras una ambientación orquestal construye una intensa música épica. En ningún caso vamos a interpretar esta asociación audiovisual pensando que hay una orquesta escondida detrás de las dunas tocando de forma entusiasta bajo el sol abrasador del desierto. Lo que nos indica de forma inequívoca esta «incoherente» asociación entre imagen y sonido es que nuestro jinete se dispone a realizar un gran acto heroico. Este tipo de uso no naturalista, o no realista, de la música se utiliza también para narrar los estados emocionales de los personajes; para crear efecto de previsibilidad (algo va a ocurrir); para dar sensación de tensión a un momento determinado de la historia; etc.

- *La creación de efectos sonoros*. Mediante la asociación de sonidos pregrabados en distintos lugares a situaciones visuales filmadas en plató televisivos o cinematográficos ha sido posible conferir a las imágenes resultantes sensación de realismo. Un ejemplo clásico es el de las secuencias ambientadas en pantanos brumosos o en densas selvas tropicales en el cine de los años cincuenta y sesenta, normalmente filmadas en plató y sonorizadas después con ruidos de los animales naturales de ese tipo de zonas geográficas. Al vincular las imágenes a los efectos sonoros selváticos y tratar los sonidos con distintas intensidades se conseguía la sensación de espacio abierto y profundo lleno de misteriosos animales situados a diferentes distancias de los protagonistas. Otro recurso clásico es el de la asociación de ecos artificiales a determinadas imágenes para crear sensación de gran espacio cerrado. O, al contrario, eliminar en el momento de la sonorización las resonancias del plató para conseguir sensación sonora de espacio abierto.

Al poder trabajar a voluntad con una nueva recomposición entre sonido-imagen y, en consecuencia, con una recomposición virtual entre sonido y fuente sonora se han ampliado extraordinariamente las posibilidades expresivas del universo audiovisual. Así, lo que podríamos denominar ya como la *acusmatización tecnológica*, para diferenciarla de la *acusmatización pitagórica*, ha supuesto una

auténtica revolución tanto en el ámbito estricto del lenguaje audiovisual como en el universo de la comunicación de masas en el sentido más amplio y genérico del término, incluida su vertiente industrial.

## 2.2. Los conocimientos necesarios para dominar el audio

Todo este complejo universo audiovisual que acabamos de revisar un poco más arriba está repleto de posibilidades expresivas vinculadas al sonido acustmatizado. Pero ¿cuáles son los fundamentos del saber que pueden ayudarnos a dominar con eficacia la narración a través del audio? El contar cosas desde el sonido en las producciones audiovisuales tiene varias bases de conocimiento que actúan a la vez estrechamente vinculadas unas a otras. Pero, curiosamente, en algunas de ellas se piensa siempre, y en otras casi nunca.

### 2.2.1. La tecnología y la experiencia productiva

Sin duda, la disciplina más conocida entre las que apoyan todo este universo expresivo es la tecnología. Todo narrador o aspirante a narrador audiovisual se ha enfrentado a la necesidad de saber con qué sistemas, instrumentos, aparatos, técnicas, o trucos es posible hacer tal o cual sonido.

También se piensa muy a menudo en la experiencia productiva como generador del conocimiento acústico-narrativo. El realizador se plantea qué sensación percibe sobre sí mismo, o qué sensación ha percibido su público al escuchar tal o cual efecto sonoro en cada situación narrativa, contexto visual, etc., en los que trabaja. A partir de ahí se produce la acumulación progresiva de un saber generado desde la experiencia productiva desarrollada por uno mismo o por otros narradores. Obviamente, la experiencia productiva acumulada por «otros» es algo difícilmente transferible.

En suma, cuando nos aproximamos al universo de la producción sonora audiovisual, el dominio de este universo parece depender casi exclusivamente de un amplio saber sobre aparatos electrónicos y de la experiencia dilatada en su uso. Curiosamente, se suele pensar bastante menos en la acústica, es decir, en el conocimiento físico del sonido; y mucho menos todavía en la percepción e interpretación de las formas acústicas por parte del ser humano, o sea, en la psicología perceptiva. No obstante, paradójicamente, a pesar de que la tecnología es lo que actualmente parece deslunbrar de manera exclusiva a los estudiosos de la narración y a los propios narradores audiovisuales, la base de la tecnología del audio es la acústica, y la base de la experiencia productiva del cineasta o del realizador son sus propios mecanismos perceptivos.

Pero veamos de una forma algo más desarrollada en qué se fundamentan las afirmaciones que acabamos de hacer.

### 2.2.2. La acústica

La base fundamental en la ingeniería de la alta fidelidad sonora (*HI-FI*) consiste en garantizar que después de cualquier manipulación electrónica, grabación, transmisión, reordenación, etc., el audio conserva con exactitud su espectro sonoro original. O, dicho de otro modo, la garantía de que el sonido ha de seguir exactamente igual antes y después del tratamiento (Marras, 1977). La única forma de observar si esto es cierto es utilizando la acústica como disciplina de apoyo. Sin saber cómo actúan sobre el sonido los equipos electrónicos que tratan el audio, la electrónica no tiene ningún sentido. Y este saber sólo es posible a partir del análisis acústico del audio antes y después de su tratamiento tecnológico.

Es cierto que frente a esto se podría argumentar que basta con tratar los equipos de sonido como una *caja negra* de la que tenemos garantía de que lo que entra en ella es exactamente igual a lo que sale y, por tanto, que podemos despreocuparnos de lo que ocurre dentro de ella. Esto nunca es objetivamente cierto. Todo aficionado a la música sabe perfectamente que la misma interpretación musical suena de forma muy distinta en el teatro que en su equipo de alta fidelidad. Del mismo modo que todo aficionado al cine ha experimentado más de una vez que la misma película no suena de la misma manera en la sala de cine que en el televisor de su casa. Y no entramos aquí en el juicio de valor que supone decidir en cuál de los dos sitios suena mejor, sino, simplemente, en la constatación de que existe una diferencia claramente perceptible. Esto supone, consecuentemente, que los aparatos de audio, por muchas garantías que ofrezcan y muy sofisticados que sean, siempre alteran el sonido, y sólo la acústica nos permite saber de qué modo.

Pero tampoco es una cuestión de sutilezas auditivas a lo que nos estamos refiriendo. El objeto de estudio de esta obra es el papel del sonido en el contexto de la narración audiovisual; y en la producción audiovisual una de las funciones fundamentales de los tratamientos del sonido es su alteración, el cambio de sus formas. Me estoy refiriendo a tratamientos como el filtrado (eliminación de una parte de las frecuencias), la compresión (acortamiento de la duración en determinadas partes de sonido), la aceleración, la ralentización, la adición de resonancias, etc. Probablemente la forma más clara de ilustrar esto sea haciendo referencia al doblaje.

Todo el doblaje actual se hace en estudios acústicamente bien aislados y (le paredes interiores muy absorbentes para evitar reflexiones sonoras que puedan revelar el tamaño y la forma del locutorio. No obstante los personajes a doblar aparecen a menudo en las imágenes hablando metidos en cuevas, grandes salones, catedrales... Lógicamente, este tipo de espacios altera el sonido de las voces que se emiten en su interior añadiéndoles resonancias o, incluso, ecos. La única manera, entonces, de hacer coherente la calidad del sonido de la voz del doblador, que está metido en una sala en la que no hay ninguna resonancia, con el espacio fuertemente resonante que se muestra en las imágenes, es manipularlo electrónicamente el sonido. Es decir: cambiándolo, alterándolo acústicamente de forma muy importante. La única forma, entonces, de saber qué está ocurrien-

lo con cada sonido, de saber cómo hay que trabajar y qué tipo y qué nivel de resonancia hay que añadir, de saber cómo actuará cada uno de los aparatos de que disponemos en el estudio, es utilizando los instrumentos conceptuales y técnicos de la disciplina que estudia el sonido objetivándolo como ente físico, es decir: la acústica.

### 2.2.3. La psicoacústica

Por otra parte, la base de toda comunicación humana es la percepción, y en el caso que nos ocupa: la percepción del sonido. Sin conocer los mecanismos del ser humano para percibir y entender las formas sonoras, el estudio de los aspectos sonoros del lenguaje audiovisual tampoco tiene ningún sentido. Pero vayamos también a un caso concreto. Una pregunta paradigmática en un momento como el actual, en que el soporte analógico de las grabaciones comerciales de audio está sustituyéndose a marchas forzadas por el soporte digital, podría ser la siguiente: ¿es o no es capaz el consumidor medio de valorar la diferencia de calidad de sonido entre ambos sistemas?

La rápida implantación del disco óptico en la industria discográfica (Compact Disc) parece apuntar a una respuesta afirmativa, pero el rotundo fracaso comercial del casete digital (DAT)<sup>1</sup> indica justamente lo contrario. La mejora en la calidad sonora que ofrecía el nuevo sistema no fue suficiente para que el consumidor medio abandonara el casete analógico de alta fidelidad. Si pensamos en las importantes mejoras funcionales que supone el disco óptico frente al antiguo sistema de discos de vinilo: se deteriora mucho menos, ocupa un espacio mucho menor y su manejo es mucho más cómodo y más versátil; posiblemente no nos equivocamos al pensar que la mejora en la calidad del sonido que supone la tecnología digital no es el factor determinante de su consumo. De hecho la nueva alternativa digital al casete, el *mini disc*, centra ya su oferta técnica y publicitaria mucho más en las posibilidades de edición doméstica del sonido que en la mejora de su calidad acústica. La fidelidad de la grabación de audio no ha cesado de mejorar desde que Edison inventó en 1877 su gramófono, pero este proceso parece haber llegado a su límite que es el límite de la propia percepción humana.

Pero veamos lo que pueden aportar algunos conocimientos de psicología perceptiva a la solución de la pregunta planteada.

Desde el punto de vista de la percepción, la diferencia entre una grabación digital y una analógica con equipos de alta fidelidad que incorporen el sistema «Dolby» de reducción de ruido no va mucho más allá de la eliminación de una sutil cantidad de ruido de fondo. Concretamente, el ruido de fondo de una platina de casete analógica está en torno a los 18-20 decibelios y el ruido de fondo de un sistema digital está entre los 9-10 decibelios. Es decir, aproximadamente la mitad. Esta ventaja resulta bastante poco importante si tenemos en cuenta que no existe ningún entorno sonoro absolutamente silencioso. El nivel de ruido habitual en

1. Digital Audio Tape.

cualquier contexto de audición que normalmente calificaríamos como «silencioso» es de unos 30-40 decibelios. Así que en las condiciones domésticas normales, el ruido ambiente duplica o triplica en intensidad el ruido de fondo de una grabación analógica de buena calidad, que como decíamos es de 18-20 Db. Si a estos datos añadimos que estos dos tipos de ruido son, normalmente, coincidentes en el tiempo, se llega sin ningún problema a la conclusión de que lo normal es que el ruido ambiente enmascare perceptivamente el levísimo ruido de fondo del audio analógico. Es decir, que lo esconde tras él haciéndolo inaudible.

A la luz de la psicología perceptiva, la respuesta a la pregunta planteada más arriba es que los sistemas digitales de audio suponen un avance poco útil para el oído humano puesto que la capacidad de resolución auditiva, en condiciones que no sean las de un laboratorio, no permiten disfrutar de la mejora acústica que supone el tratamiento numérico del audio. Naturalmente, las posibilidades de tratamiento de señal, y por tanto de comodidad en la manipulación profesional del sonido (edición, filtrados, almacenamientos, etc.) que aporta esta nueva tecnología suponen un avance indiscutible, pero esto es ya una cuestión muy distinta.

En suma, la única forma de saber cuál es la relación que existe entre los efectos sonoros que producen las máquinas de tratamiento del audio y la interpretación que hace de ellos el ser humano es apoyándonos en la acústica. Necesitamos estudiar con precisión cómo son los sonidos para llegar a saber con exactitud qué clase de sensaciones perceptivas produce cada tipo de variación acústica. Consecuentemente, la acústica es un instrumento que nos permitirá estudiar con precisión las formas sonoras, descubriendo cómo el ser humano da sentido a cada sonido al ser escuchado e interpretado por él. Esta perspectiva de conocimiento desemboca en la *psicoacústica*, es decir, en la integración sistemática de la acústica y la psicología de la percepción, como una disciplina de apoyo esencial para todo trabajo sonoro vinculado a la producción audiovisual.

### 2.2.4. El sentido de aplicar la acústica al lenguaje audiovisual

Demasiado a menudo, tanto los narradores audiovisuales como los estudiosos del lenguaje audiovisual desprecian los modelos basados en las ciencias físicas y los instrumentos de medición precisa, considerándolos como algo no aplicable a la creación de productos audiovisuales. Las razones en contra suelen basarse en una serie de argumentos, más o menos peregrinos, que desembocan siempre en considerar la «creación» como algo basado en una suerte de mística imposible de objetivar y medir.

Como respuesta puramente pragmática a esa corriente, en este apartado haremos una amplia revisión de las posibilidades que puede suponer, o que está suponiendo ya en muchos casos, la aplicación sistemática de los instrumentos conceptuales de la acústica y de la psicología perceptiva a la producción audiovisual.

A menudo se han desarrollado estudios previos sobre las formas visuales para la producción de películas de ciencia-ficción o de dibujos animados. Seguramente el lector recordará las promociones de muchos de los filmes de la productora

Disney en las que se explicaba el trabajo de investigación visual realizado para la secuencia del baile en *Blancanieves*, para dibujar la lluvia en *Bambi*, para crear los espacios virtuales de *Tron*, para dar expresividad a la cara de *Aladín*, etc. Este tipo de estudios supone un trabajo previo de análisis sobre formas visuales aparentemente no perceptibles, pero que influyen de una manera definitiva en el efecto global de verosimilitud. Es el caso, por ejemplo, del chapoteo de las gotas de agua, con las que no se consiguió un efecto realista hasta que se observó, con macro-ampliaciones y a cámara lenta, el efecto físico objetivo de la fragmentación de las gotas de agua en las salpicaduras. A pesar de que este efecto sólo era visible en el laboratorio, cuando se reconstruyó en los dibujos animados pudo comprobarse que daba una verosimilitud definitiva al efecto visual del chapoteo.

El instrumental teórico y técnico que ha desarrollado la psicoacústica pone al alcance del narrador audiovisual la posibilidad de aplicar este tipo de investigación previa también a la producción creativa del audio.

Veamos algunos ejemplos:

*En la voz*, por ejemplo. En el momento en que un realizador se enfrenta a la necesidad de construir dramáticamente un personaje, de darle forma sonora y visual, aparece el problema de la elección y la adaptación de la voz del actor, o su sustitución por otra voz (doblaje). Nos encontramos, con el fenómeno absolutamente cotidiano en la ficción audiovisual de que ciertas voces transmiten unas determinadas sensaciones y otras no. Y que, a menudo, el efecto que buscamos no aparece completo en ninguna de las voces de las que disponemos. ¿Por qué ocurre esto? Profesionalmente, apenas se utilizan instrumentos conceptuales para su descripción y mucho menos para su análisis y su manipulación.

Se pide, entonces, a uno de los locutores que adopte una actitud aninada, pero su voz no llega a lograrlo. ¿Por qué? ¿Qué le falta o que le sobra a esa voz para conseguirlo?

El conocimiento de la acústica del espectro de la voz nos puede proporcionar toda esa información con mucha precisión. Es perfectamente posible saber cómo actúa, por ejemplo, el valor expresivo del tono, es decir de la frecuencia fundamental del espectro de la voz. A grandes rasgos, cualquier director de doblaje sabe que una «voz grave» sugiere credibilidad, madurez y seguridad. Y que una «voz aguda» transmite poca credibilidad, inmadurez e inseguridad. Pero ¿cuáles son los límites exactos de esto? ¿Dónde termina lo grave y empieza lo agudo? ¿Qué ocurre con las voces intermedias? ¿Qué pasa con las voces femeninas? ¿Qué efecto produce que un locutor de registro normalmente grave fuerce su voz hacia los agudos?

Las respuestas a todas estas preguntas son siempre absolutamente intuitivas. En cambio, el tono concreto de cualquier voz y su margen específico de variación son activamente cuantificables y diferenciables de los de cualquier otra. Y sus efectos en la construcción de los personajes son también rigurosamente analizables, hasta el punto de que es posible desarrollar una lista cerrada de categorías sobre este tipo de efecto narrativo (Rodríguez Bravo, 1989).

Sería perfectamente posible, por ejemplo, establecer una tipología acústica de los personajes que podría facilitar enormemente la tarea de la elección de los ac-

tores o los presentadores en función de lo que sugiere su voz. Bastaría con hacer una lista de los rasgos de personalidad que el director desea que sugiera la voz para buscar en un banco de datos el actor o presentador con una voz ajustada a esta demanda. También es viable, sin ningún tipo de problemas, estudiar acústicamente la voz de los personajes históricos, lo mismo que se hace con su imagen, con el objeto de encontrar un doblador o dobladora que tenga exactamente sus características o sea capaz de reproducirlas con la ayuda de determinados tratamientos técnicos.

*El reconocimiento sonoro del espacio* plantea un problema muy similar al de la voz. La identificación sonora del espacio consiste en reconocer las resonancias que se han añadido a sonidos que ya conocemos y que han sido situados en otros contextos espaciales: pasos, voz, toses, ruido de movimiento de papeles, de objetos de madera, etc., pero ¿cuál es el carácter acústico que hace unas resonancias distintas de otras? ¿Qué resonancias se asocian a qué espacios? ¿Por qué?

El análisis previo de los tiempos de reflexión del sonido en cada uno de los espacios sobre los que vamos a trabajar nos van a dar una información acústica valiosísima para definir la reverberación de cada espacio visual de acuerdo con la experiencia auditiva natural que tiene cualquier oyente. Este tipo de estudio previo permite diferenciar unos espacios de otros de manera verosímil y realista, incluso utilizando sistemas electrónicos de reverberación artificial.

Otro de los problemas clásicos de la narración audiovisual es resolver cuál es el tipo de sonido que le ha de llegar al espectador en cada momento. Pensemos, por ejemplo, en las situaciones en las que se produce un cambio del punto de vista de la cámara en el mismo espacio contemplando la misma fuente sonora, y no cambia en absoluto el paisaje sonoro. Este problema tiene una doble dimensión: la reconstrucción auditiva realista, y la ubicación del espectador en el espacio de la narración. El análisis acústico previo de los espacios que se van a recrear en una narración abre, también, un amplio universo de posibilidades expresivas en la ubicación del espectador en estos espacios.

Si el lector busca en su memoria el sonido de una sala, el de un aula por ejemplo, recordará que mientras el profesor habla desde su mesa, la voz no suena igual escuchándola desde la primera fila que desde la última. En consecuencia, no se trata lo mismo reconstruir en una película el sonido de la voz concreta haciendo que todos los espectadores del cine se sintieran en la primera fila, o que se sintieran en la última. A partir del análisis de algunos de los parámetros acústicos del espacio: intensidad, tono fundamental, resonancias, duración y evolución temporal, es posible dominar minuciosamente la situación espacial en la que queremos que se sienta el espectador mientras escucha. El método de trabajo partiría de estudiar el sonido del espacio desde puntos concretos de audición.

*La construcción de atmósferas emocionales* es otro de los ejemplos clásicos de la narración audiovisual en los que el papel del audio tiene un gran protagonismo y que la adecuada aplicación de la acústica supondría una mejora esencial. Los estados emocionales están asociados a los sonidos con la misma fuerza que a los olores, y disponer del instrumental adecuado para reconocer la forma de estos so-



nidos es tener a nuestro alcance la llave de las emociones y las sensaciones que están asociados a ellos.

Uno de los problemas habituales en la narración de estados emocionales es encontrar el sonido capaz de comunicar con fuerza, por ejemplo, una gran sensación de soledad. Todos hemos vivido situaciones similares, y esas situaciones están asociadas a sonidos concretos que no son nada fáciles de encontrar. Si conseguimos encontrarlos y reconstruirlos, esos paisajes sonoros son fuertes inductores de estados emocionales. Pensemos en que nos sugieren: «el tic-tac de un reloj en una habitación en silencio», «un goteo de agua en un espacio resonante amplio», «el sonido de los grillos», «el rumor del agua circulando con resonancias de túnel o bóveda», «el sonido del viento en un espacio abierto», «el sonido del viento a través de una ventana cerrada», «el sonido de la lluvia en la calle; desde una ventana; en la montaña». Todos estos paisajes sonoros inducen con fuerza sensaciones como soledad, frío, vacío, nostalgia, etc. Pero no basta con estas ideas; con mucha frecuencia los realizadores experimentan que por alguna extraña razón un determinado efecto no sugiere lo que se esperaba de él, no suena a aquello que buscamos a pesar de que en realidad lo es. ¿Por qué ocurre eso? ¿Por qué un sonido que identificamos perfectamente deja de actuar como debiera?

El sonido */grillo/* es mucho más que ese concepto simple de *insecto que canta por la noche*. Es una duración, es una composición de frecuencias que suena en un espacio concreto, es un ritmo, es una evolución de la intensidad que varía en el tiempo, es el resultado de una atenuación (filtrado) de una parte del espectro por la humedad del aire y la distancia que lo separaba de nosotros. Es, en suma, un fenómeno tan preciso y matizado como la luz y el punto de vista, y para definirlo el instrumental acústico es imprescindible si no queremos limitarnos exclusivamente a usar el mismo decorado sonoro indefinidamente.

Finalmente, *la edición y el control informatizado del sonido* es el ámbito donde la acústica resulta ya un conocimiento imprescindible para obtener rendimientos profesionales mínimamente aceptables.

Actualmente ya existen en el mercado sistemas que permiten la edición informatizada del sonido mediante la digitalización y tratamiento de su oscilograma, segmentándolo en paquetes temporales y manipulándolo en ficheros que permiten el montaje sobre el eje de la temporalidad siguiendo el modelo visual de los tratamientos de texto. Es decir, permiten cortar y reordenar los sonidos como si se tratara de cromos autoadhesivos.

Permiten también el tratamiento lineal del timbre del sonido mediante la manipulación del espectro de frecuencias en tiempo real. Y la aplicación de algoritmos de compresión o de dilatación del sonido mediante el tratamiento de las zonas estacionarias del oscilograma. Esta última función se aplica sobre todo al montaje sincrónico sonido-imagen. Resumiendo: cuando el sonido no cuadra con la imagen el sistema lo recorta de manera que los fragmentos eliminados se reparten proporcionalmente a lo largo de todo el paquete sonoro en las zonas en las que se nota menos. El resultado es que el sonido se acorta sin que se note donde está el golpe de tijera. Otra de las posibilidades que ofrecen es la manipulación del ataque o la caída al comienzo y al final de cada sonido. Esta función se aplica especialmente

a los efectos sonoros y a las composiciones de música electrónica y permite cambiar radicalmente la sensación perceptiva de los sonidos, posibilitando, por ejemplo, la transformación artificial de un sonido de piano en uno de violín.

Lógicamente, la correcta utilización de este tipo de sistemas, que comienzan a ser en estos momentos algo cotidiano en los estudios de audio de alto nivel, pasa necesariamente por dominar, por lo menos, los instrumentos conceptuales básicos de la acústica.

### 2.3. Del fenómeno acústico al sentido audiovisual

Revisadas ya las razones que nos mueven a tomar la acústica como un campo de conocimiento fundamental en la investigación sobre la expresión audiovisual, pasaremos a desarrollar la definición de los conceptos esenciales que nos permitirán vincular sólidamente la acústica con la percepción sónica del sonido. Así, el primer paso para desarrollar un modelo coherente que intente explicar con eficacia la acústica del sonido en el contexto del lenguaje audiovisual, será sistematizar cuatro conceptos básicos en los que se va a apoyar todo nuestro desarrollo teórico posterior. Concretamente los de: *sonido*, *fuentes sonora*, *objeto sonoro* y *ente acústico*.

#### 2.3.1. El sonido

La lingüística ha mostrado ya con mucho detalle que el sonido como tal, como fenómeno físico, es un fenómeno que no tiene nada que ver con las formas significativas que se pueden construir con él como sustancia modelable (Hjelmslev, 1980). Y nos ha enseñado, también, que una cosa es el hablante como ente capaz de producir sonido y otra muy distinta el habla como sonido estructurado.

La aplicación de este tipo de lógica estructural al sistema sonoro de la comunicación audiovisual, de modo que cada nivel de complicación funcional del sonido quede perfectamente delimitado y definido, como podrá comprobar el lector, resulta muy fértil.

Comenzaremos con la definición del nivel más elemental desde el punto de vista estructural: *el sonido*. Puesto que la aproximación que estamos haciendo al audio es esencialmente fenomenológica, intentaremos localizar los mecanismos que estructuran la sensación sonora, pero para ello no podemos ignorar en ningún momento que todas nuestras propuestas estructurales han de ser rigurosamente coherentes con el fenómeno mismo. Así, partiendo de este planteamiento, nos interesa el concepto de sonido como primera etapa de un proceso expresivo, como material físico perceptible en bruto sobre el cual se van a efectuar una serie de manipulaciones para modelarlo, transformándolo en material expresivo.

El sonido es ya en su esencia misma un puente entre la acústica y la percepción. El sonido es el resultado de percibir auditivamente variaciones oscilantes de algún cuerpo físico, normalmente a través del aire. El origen de un sonido es siempre la

vibración de un objeto físico dentro de la gama de frecuencias y amplitudes que es capaz de percibir el oído humano. Esta vibración empuja rítmicamente las moléculas de los otros cuerpos físicos que lo rodean generando a su vez vibraciones en ellas. Cuando estas vibraciones llegan a nuestro oído, normalmente a través del aire, las percibimos como un sonido. En suma, el fenómeno sonoro es la percepción de las oscilaciones rítmicas, normalmente, de la presión del aire, y que han sido estimuladas por otro objeto físico vibrante que actúa como fuente de emisión.

Definimos pues el *sonido* como:

El resultado de percibir auditivamente variaciones oscilantes de algún cuerpo físico, normalmente a través del aire.

Para comprender esta definición en todo su sentido es importante tener en cuenta que la percepción del sonido no se realiza exclusivamente a través del aire. Así, el sonido puede llegarnos, por ejemplo, a través de la vibración de nuestro propio cuerpo. Éste es el caso de la percepción que tiene todo locutor de su propia voz.

Todo hablante percibe su propia voz con una calidad sonora muy distinta a como la oyen los demás. Una experiencia reveladora, que seguramente el lector habrá vivido ya, es la de todo aquel que escucha por primera vez su voz grabada en un magnetófono. De repente, uno se encuentra con la desagradable sorpresa de que no reconoce su propia voz. De que no se reconoce a sí mismo. Y eso se vuelve todavía más desorientador cuando otras personas que han escuchado también esa grabación nos aseguran que el aparato está reproduciendo con total fidelidad el sonido y que, efectivamente, uno suena así de raro para todos los demás. La explicación de esta percepción distinta es, justamente, que mientras las vibraciones de la voz de un locutor llegan a su propio oído, fundamentalmente, a través de su estructura ósea, tanto al magnetófono como a todos los demás oyentes el sonido de la voz les llega a través del aire; con lo que las vibraciones de la fuente original (la laringe del locutor) se ven profundamente alteradas por el medio que las transmite.

Otra forma relativamente habitual de percibir la influencia del medio que transmite las vibraciones sonoras en la percepción del sonido es comparando como suena algo con los oídos dentro o fuera del agua. Por ejemplo, el sonido del motor de una canoa mientras nos bañamos en el mar. Es muy fácil comprobar cómo cambia radicalmente el ruido del motor según lo escuchemos con los oídos sumergidos en el agua o no. De nuevo aquí la diferencia se debe a la influencia del medio que transmite la vibración.

### 2.3.2. Fuente sonora

Imagino que el lector habrá deducido hace ya desde hace muchas líneas que una primera diferencia básica imprescindible desde el punto de vista audiovisual es la de distinguir entre sonido y fuente sonora.

La tecnología del audio nos permite tratar el sonido como un fenómeno que es posible empaquetar, separar y reproducir de forma completamente independiente del objeto físico que lo generó. En consecuencia, parece muy claro que, a pesar de que *sonido* y *fuente sonora* tienen entre sí una relación evidente, es necesario tratarlos como conceptos separados.

Definiremos *fuente sonora* como:

Cualquier objeto físico mientras está emitiendo un sonido.

Esta definición introduce en el concepto de fuente sonora la dimensión territorial en tanto que, normalmente, existe la posibilidad de que los objetos físicos que son capaces de emitir sonidos (la laringe humana por ejemplo) suenen o no de manera totalmente arbitraria en el tiempo. Así, cuando una laringe no está emitiendo sonido no debe ser considerada como una fuente sonora.

Ampliar el concepto a todos los objetos capaces de producir sonidos, estén sonando o no, nos llevaría a confusiones importantes en tanto que prácticamente todos los objetos físicos que existen son susceptibles de emitir sonido si se les suministra de la manera adecuada.

Como hemos comentado ya, y veremos con detalle más adelante, la separación conceptual entre sonido y fuente sonora es esencial desde el punto de vista narrativo ya que la acusmatización audiovisual abre la posibilidad de que sonidos y fuentes sonoras sean completamente intercambiables.

### 2.3.3. Objeto sonoro

Estamos trabajando con el sonido como objeto de estudio en un sentido genérico y amplio. Las formas acústicas contienen estructuras muy complejas que nos interesa estudiar en todas sus dimensiones. Esto nos lleva a la necesidad de encontrar métodos de acotar el sonido de forma objetivable, en unidades concretas que permitan su estudio sabiendo cuáles son los límites de aquello que estamos investigando.

De esta necesidad parte el concepto de *objeto sonoro*.

Siguiendo a Murray Shafer (Shafer, 1979, pág. 185) que, a su vez, toma este concepto de Pierre Shaeffer (Shaeffer, 1988, pág. 102), entendemos que un *objeto sonoro* es un sonido concreto que hacemos objeto de nuestra percepción y de nuestro estudio. Un sonido del que definimos sus límites físicos, su principio y su final para que sea posible estudiarlo de forma sistemática.

Así, podemos definir *objeto sonoro* como:

Cualquier sonido que aislemos físicamente o con instrumentos conceptuales, acotándolo de una forma precisa para que su estudio sea posible.

Esta definición se inspira de una manera muy clara en el concepto metodológico de *objeto de estudio*. Podríamos decir también, por ejemplo, que un objeto



sonoro es todo sonido que transformemos en nuestro objeto de estudio. Lógicamente, para transformar un sonido en objeto de estudio es imprescindible aislarlo físicamente o acotarlo conceptualmente con suficiente exactitud.

Como conceptos vinculados al de *objeto sonoro*, aunque de menor nivel en tanto que éste los englobaría, podemos hablar de *suceso sonoro* y de *forma sonora*.

Un *suceso sonoro* es cualquier sonido acotado en el tiempo (véase Shafer, 1979). Definiremos *forma sonora* como cualquier sonido identificable y reconocible a través de alguna de sus características acústicas.

Como el lector habrá deducido sin dificultad, cualquiera de estos dos conceptos puede ser utilizado para definir un *objeto sonoro*.

### 2.3.4. *Ente acústico*

Cuando escuchamos a través de un medio de comunicación audiovisual un sonido, el oyente puede reconocer o no reconocer la fuente sonora que lo ha generado. Cada una de estas dos posibilidades genera una fenomenología completamente distinta desde el punto de vista cognoscitivo.

Cuando el oyente reconoce la fuente generadora del sonido, el sonido deja de ser un *objeto sonoro* para pasar a actuar en la mente del receptor como si fuese la *fuentes sonora* misma. No obstante, la *fuentes sonora* ya no existe, solo existe el sonido como un ente independiente que ha adquirido para el receptor un valor signico aparentemente «sustitutorio» de la fuente sonora. En realidad el fenómeno comunicativo que esto comporta es bastante más complejo que una simple sustitución del sonido por la fuente. O, en todo caso, este valor sustitutorio es de un carácter muy similar al que tiene un signo lingüístico.

Intentaremos ilustrar esto con un ejemplo.

Mediante la forma sonora */trino de pájaros/* podemos evocar en la mente del oyente la sensación de presencia de un grupo de pequeñas aves, sin ninguna necesidad de que éstas estén presentes en realidad, sin necesidad siquiera de que éstas existan. Exactamente igual que hacemos con la lengua al decir por ejemplo «allí detrás hay unos pajarillos». La única diferencia objetiva es que el signo */trino de pájaros/* no es arbitrario sino indexativo. El par de los pájaros es reconocido como un forma acústica que está asociada de forma física al referente que evoca y esto le confiere un grado de verosimilitud muchísimo mayor del que tiene una secuencia lingüística.

Pero antes de seguir avanzando en el estudio detallado del concepto *ente acústico* intentaremos definirlo con la precisión suficiente:

Llamaremos *ente acústico* a cualquier forma sonora que habiendo sido separada de su fuente original, es reconocida por el receptor como una fuente sonora concreta que está situada en algún lugar de un espacio sonoro.

El hecho de que el oyente identifique una forma sonora con un objeto físico concreto que emite sonido conlleva necesariamente que este objeto físico esté situado en algún espacio volumétrico. Es decir, el acto mismo de la construcción de un ente sonoro desencadena, también, la construcción del espacio sonoro que lo contiene. No obstante, debemos tener en cuenta, también, que a pesar de que teóricamente todo sonido proviene siempre de una fuente física real, el oyente no siempre se preocupa de identificarla.

Pensemos de nuevo en la fenomenología de la escucha radiofónica.

Muy a menudo oímos el discurso informativo preocupándonos sólo por interpretar sus contenidos, sin prestar ninguna atención al lugar que ocupan los locutores en el espacio del estudio, a la distancia que hay entre ellos, o a cómo estaban colocados los instrumentistas que grabaron las ráfagas musicales utilizadas como separador entre las distintas secciones del informativo. Es decir, como oyentes, podemos optar perfectamente por no preocuparnos en ningún momento por reconocer a un locutor o a una música como un *ente acústico*.

Pero también podemos hacer lo contrario.

En el momento en el que un oyente sustituye intelectualmente el sonido de la voz de un locutor por el locutor mismo, automáticamente está creando un espacio para él y en torno a él. En este momento el sonido de la voz actúa como un *ente acústico*, es decir, como una forma acústica que sustituye significativamente a una fuente sonora, y que puede ser tratada dentro del espacio sonoro exactamente igual a como podría ser tratada la fuente sonora evocada en el espacio real.

Para comprender la lógica narrativa que organiza el espacio sonoro es importante tener en cuenta que los *entes acústicos* son siempre acusmáticos. Es decir, hablando de una forma menos crítica, son absolutamente independientes de su fuente sonora objetiva original. Se trata de sonidos aislados que sólo están vinculados a un objeto físico en el cerebro del receptor. Es decir, el sonido */trino de pájaros/* puede perfectamente haber sido creado mediante síntesis sonora por un ordenador. El valor expresivo de un *ente acústico* no depende de su origen productivo sino de que éste sea reconocido o no por parte del receptor como directamente dependiente de una fuente sonora. Así, el *ente acústico* no solo no depende de su fuente original, sino que puede construirse articulando sonidos con fuentes sonoras que nunca tuvieron nada que ver entre sí. Ejemplos muy claros de esto podemos encontrarlos en el radiodrama. Recuerdo en estos momentos un brillante ejercicio de narración sonora realizado por un grupo de alumnos míos en el que se asociaba un sonido sintético a los saltos de una planta que había adquirido misteriosamente la capacidad de moverse. La planta saltaba haciendo un */poooooing/* metálico extraordinariamente sugerente que aterrorizaba a una tercera señora cincuentona. El */poooooing/*, una vez asociado mediante una adecuada descripción a la exótica planta, actuaba ya narrativamente como la planta misma. A partir de este momento cualquier manipulación acústica del */poooooing/* (tratamiento de la intensidad, reverberancia, etc.) era ya interpretado por los oyentes como un movimiento espacial de la planta. En suma el */poooooing/* se había transformado en el *ente acústico /planta saltadora/*.

Naturalmente, el origen del efecto sonoro no tenía nada que ver con ninguna planta. Los estudiantes lo construyeron a partir del sonido de un muelle previamente grabado en un disco editado por la BBC,<sup>2</sup> y que ellos manipularon copiándolo en cinta magnetofónica y ralentizando después su velocidad de reproducción.

Evidentemente, un *ente acústico* es un signo en tanto que es una forma expresiva que al ser reconocida por el receptor desencadena en su mente un estímulo concreto con el que está asociada. Y, a la vez, esta forma sonora tiene un referente concreto en la realidad referencial.

Hasta aquí un *ente acústico* actúa exactamente igual que un signo lingüístico. No obstante, el *ente acústico* está justo a mitad de camino entre las concepciones peirceana y saussuriana del signo. Las formas sonoras que configuran un *ente acústico* no tienen un origen arbitrario como ocurre con el signo lingüístico, sino que la *forma sonora* que actúa como significante ha sido generada originalmente de un modo natural, y estaba vinculada de una forma objetiva al ente físico que la produjo mediante sus vibraciones.

Pero desde el momento en que arrancamos el sonido de su fuente física y lo tratamos en una narración de manera independiente, tampoco está actuando como un índice vinculado rígidamente a su origen físico.

Al enfrentarnos con el concepto de *ente acústico* nos encontramos con un ejemplo muy claro de una fenomenología signífica radicalmente nueva que tiene su origen específico en el lenguaje audiovisual y cuyas características fundamentales son las siguientes:

1. Partimos de la voluntad previa de transmitir una información de carácter naturalista.
2. Creamos artificialmente un mensaje totalmente nuevo que estimula sobre el sistema sensorial del hombre percepciones muy similares a las que producen las informaciones de origen natural.

El *ente acústico* es un signo específicamente audiovisual que tiene como característica esencial, igual que todos los signos específicamente audiovisuales, una enorme capacidad de verosimilitud. Como signo audiovisual es de carácter esencialmente naturalista, y se basa en su identidad con formas sonoras y visuales que originariamente fueron índices naturales de fenómenos físicos concretos. Este carácter de fragmentos de copia casi exacta de la realidad, creados y reordenados para narrar es lo que confiere a los signos audiovisuales su gran capacidad de realismo.

2. British Broadcasting Corporation.

### CAPÍTULO 3

## DEL ANÁLISIS ACÚSTICO A LA SENSACIÓN

El objetivo de este capítulo es proporcionar al lector una serie de conocimientos sobre acústica básica y su relación con el fenómeno perceptivo que le serán imprescindibles para comprender muchos de los razonamientos y los modelos que revisaremos a lo largo de esta obra.

IMP. FORM. TELA DO

ALT. TELA DO

NOTICIA  
(COPILADO)

VARIANTE

COMPARTE

COMPARTE

COMPARTE

COMPARTE

COMPARTE

### 3.1. Entre el fenómeno vibratorio y el perceptivo

Si bien es cierto que no todos los fenómenos vibratorios son perceptibles por el ser humano, si que lo es que el oído del hombre tiene una extraordinaria capacidad de análisis para toda vibración simple o compuesta que esté entre los 17 y los 20.000 ciclos por segundo y entre los 0 y los 120 decibelios (naturalmente con muchos matices como veremos más adelante).

En este apartado haremos una revisión de la física del sonido, es decir, de las vibraciones perceptibles por el oído humano, para extraer de esta disciplina una serie de instrumentos conceptuales que pueden ser de gran utilidad para la narración audiovisual. Para que esta aproximación a la acústica resulte útil desde el primer momento, relacionaremos siempre las dimensiones acústicas con el tipo de sensaciones que vehiculan.

La aproximación al sonido que se va a hacer en esta obra tendrá siempre dos puntos de vista: el físico y el perceptivo. Comenzaremos en este apartado estudiando el sonido desde la acústica, pero conectando siempre esta disciplina con su inevitable dimensión perceptiva; y en el capítulo siguiente desarrollaremos con detalle la perspectiva inversa, es decir, el estudio del sonido desde el punto de vista de la psicología de la percepción. El sentido de hacer este doble esfuerzo radica en las importantes contradicciones que aparecen sistemáticamente entre la medición acústica de las formas sonoras, y la percepción humana de aquello que hemos medido,<sup>1</sup> contradicción que no suele estar suficientemente recogida ni en los tratados de acústica ni en los de psicología perceptiva. El objetivo último de este método de trabajo va a ser siempre el mismo: intentar localizar los puentes que existen entre el sonido como fenómeno físico y las distintas interpretaciones que hace de él el ser humano.

#### 3.1.1. Sonidos puros y sonidos compuestos

Para empezar a ordenar en categorías el universo sonoro comenzaremos por esta dicotomía esencial. Hablaremos de *sonidos puros* cuando estemos haciendo referencia a sonidos que están constituidos por una sola frecuencia. Los sonidos puros producen una sensación auditiva extremadamente pobre y simple, poco agradable. El ejemplo clásico es la vibración de un diapason; los pitidos electrónicos suelen ser también sonidos simples; este tipo de sonidos es bastante escaso fuera de los estudios de audio y de los laboratorios de sonido. Los *sonidos com-*

1. Este fenómeno fue sistematizado ya en la *ley de Weber y Fechner*, que estudiaremos con detalle en el apartado 3.4.2 cuando se habla de la intensidad del sonido.

*puestos* son aquellos que están formados por dos o más frecuencias. Éste es el estado más natural del sonido y su gama de posibilidades es tan extremadamente amplia que no es posible hablar de ninguna sensación auditiva común en todos ellos.

La división entre sonidos *puros* y *compuestos* es puramente física. Su sentido tiene poco valor perceptivo pero, como veremos más adelante, es la base de una aproximación teórica que permite comprender los sonidos complejos partiendo de la suma de conocimientos obtenidos en el estudio de las distintas frecuencias simples que constituyen cualquier sonido compuesto.

Los sonidos puros están generados por la vibración de cuerpos que tienen una estructura interna muy homogénea, lo que hace que todas sus moléculas respondan frente a determinado estímulo físico vibrando exactamente del mismo modo. Contrariamente, las diferencias de densidad, de forma, de cantidad de masa, etc., entre las moléculas de un mismo cuerpo hacen que al ser estimulado físicamente, por ejemplo con un golpe, se originen dentro de él varias vibraciones distintas que al interactuar unas sobre otras generan una oscilación global compuesta y, por tanto, mucho más compleja.

Una forma simple para ejemplificar esta diferencia es imaginando dos cuerpos físicos ideales, compuesto cada uno de ellos solamente por dos moléculas. Imaginemos también que las dos moléculas del *cuerpo-1* (fig. 1) son exactamente iguales y tienen la misma forma, masa y tamaño, es decir, forman un cuerpo homogéneo. En cambio el *cuerpo-2* (fig. 2) es heterogéneo y, por tanto, está compuesto por dos materias distintas, así que sus dos moléculas tienen una masa, un tamaño y una densidad diferentes una de otra. En tanto que cada uno de los cuerpos tiene sus dos moléculas ensambladas entre sí elásticamente, tanto el *cuerpo-1* como el *cuerpo-2* tendrán al vibrar un movimiento global que será la combinación de los movimientos de sus dos moléculas respectivas (fig. 3).

Estudiemos en primer lugar el movimiento vibratorio del *cuerpo-1*.

Al golpearlo, cada una de sus dos moléculas idénticas oscilará exactamente del mismo modo, trazando dos movimientos vibratorio-armónicos simples. Estos movimientos pueden representarse mediante dos senoídes también idénticas y en fase. Si el resultado es la combinación de movimientos, y esto lo representamos mediante la suma lineal de las dos senoídes trazadas por la vibración de cada una de las moléculas A y B, el resultado será una nueva vibración simple (columna izquierda de la fig. 3). En suma, la vibración global del *cuerpo-1* (*vibración-1AB*) es una vibración simple. Vibra de un solo modo; si esta vibración fuera audible (entre 20 y 20.000 oscilaciones por segundo) estaría constituyendo un sonido puro analizable como una única frecuencia.

Pasemos ahora a estudiar con detalle el movimiento de nuestro segundo cuerpo ideal, el *cuerpo-2*.

Al recibir el estímulo, cada una de sus dos moléculas C y D responderá de un modo distinto en función de su propia masa y tamaño. Así, cada una de las moléculas que lo componen describirá un movimiento vibratorio-armónico diferente. La *molécula-D*, al ser de mayor masa y tamaño que la *molécula-C*, oscilará en menos ocasiones por unidad de tiempo pero alejándose más del centro de oscilación

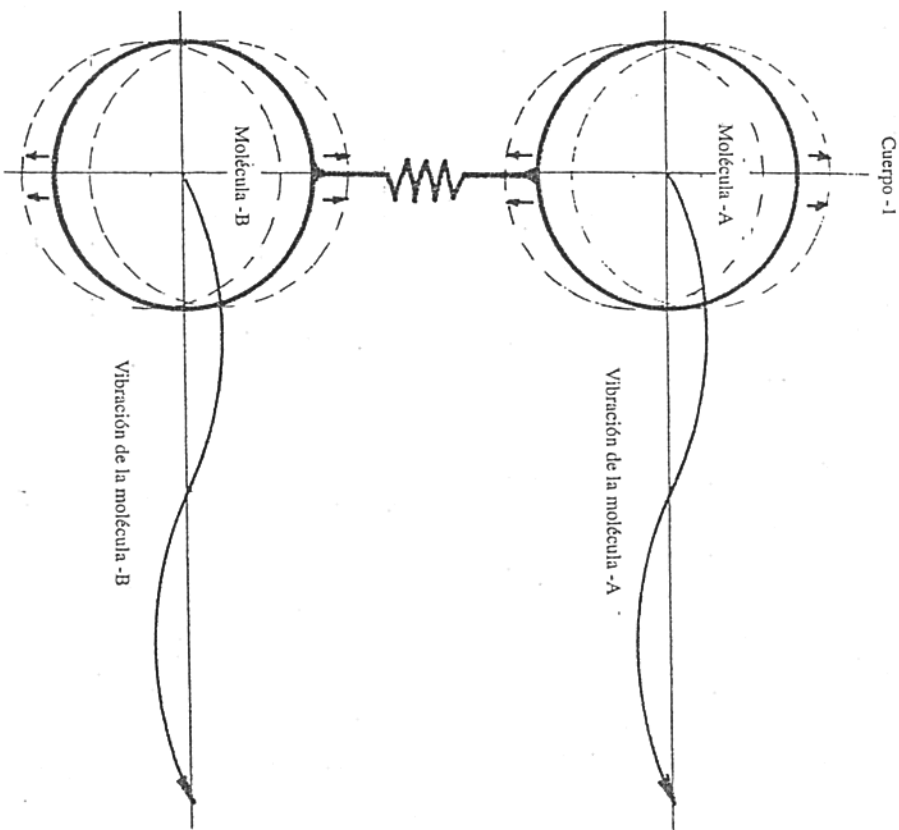


FIGURA 1. *Vibración de un cuerpo físico ideal homogéneo.*

que la *molécula-C* (fig. 2). Estamos ahora ante un sistema constituido por dos partes distintas y ensambladas que, aunque responde parcialmente de dos maneras diferentes al mismo estímulo, tiende, obligado por una unión elástica, a construir un único movimiento global. La oscilación de todo el *cuerpo-2*, en lugar de producir un movimiento vibratorio-armónico simple, lo producirá complejo (columna derecha de la fig. 3). Resumiendo, el movimiento global del *cuerpo-2* (*vibración-2CD*) es una única oscilación compleja, que si fuera audible estaría construyendo un sonido compuesto, y que al analizarlo instrumentalmente daría como resultado dos frecuencias perfectamente diferenciadas.

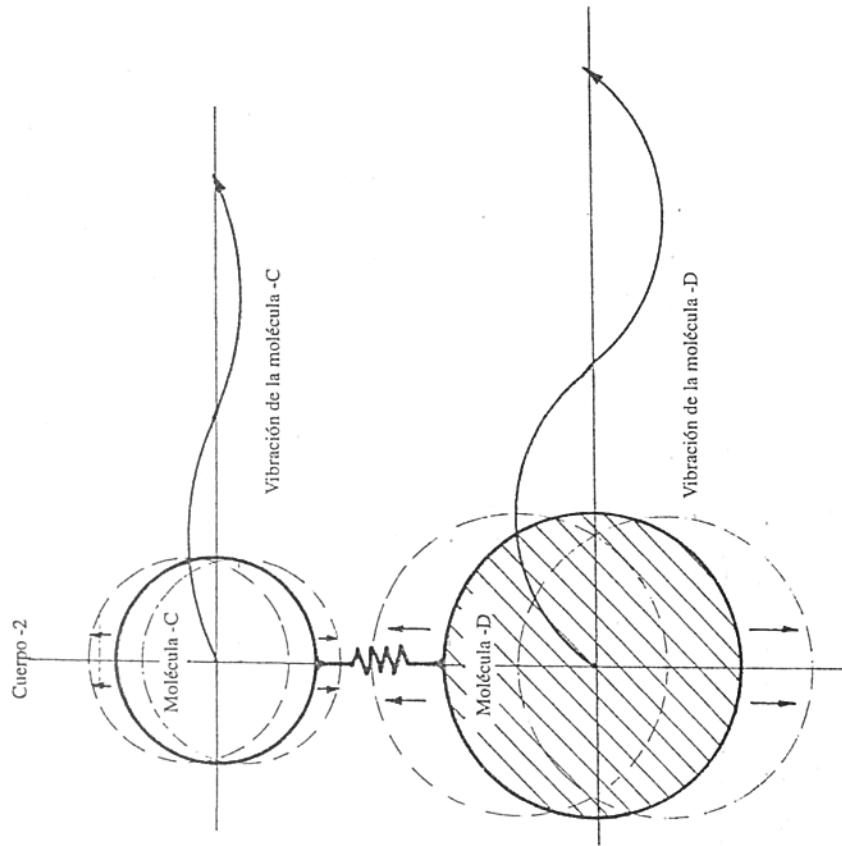


Figura 2. Vibración de un cuerpo físico ideal heterogéneo.

Normalmente, todo cuerpo físico es heterogéneo en su forma, o en ambas cosas a la vez. Al ser estimulados de modo que vibren con frecuencias audibles, las distintas oscilaciones internas de estos cuerpos heterogéneos componen un único movimiento complejo global que se traduce perceptivamente en un sonido compuesto. Estos sonidos compuestos son la única información habitual de que solemos disponer, así que solamente cuando se analizan esos sonidos compuestos con los instrumentos adecuados, éstos revelan cuales son las distintas frecuencias que los componen.

La que hemos descrito en la figura 2 no es la única forma de composición de una vibración compleja, y, en consecuencia, de un sonido compuesto. Cada molécula de un cuerpo físico puede responder, también, moviéndose simultáneamente de formas diferentes frente a estímulos distintos, organizando así vibraciones completas. Esto explica el caso de los objetos que suenan de forma completamente distin-

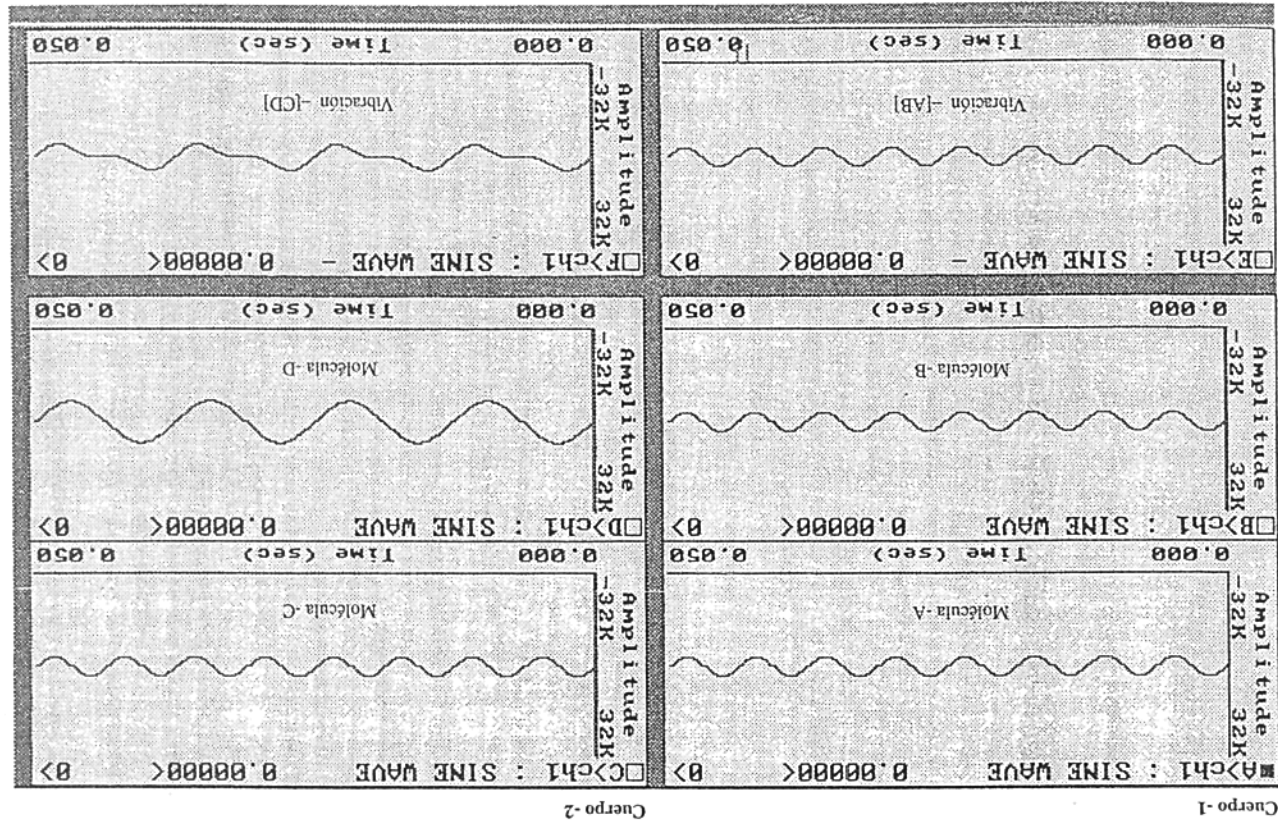
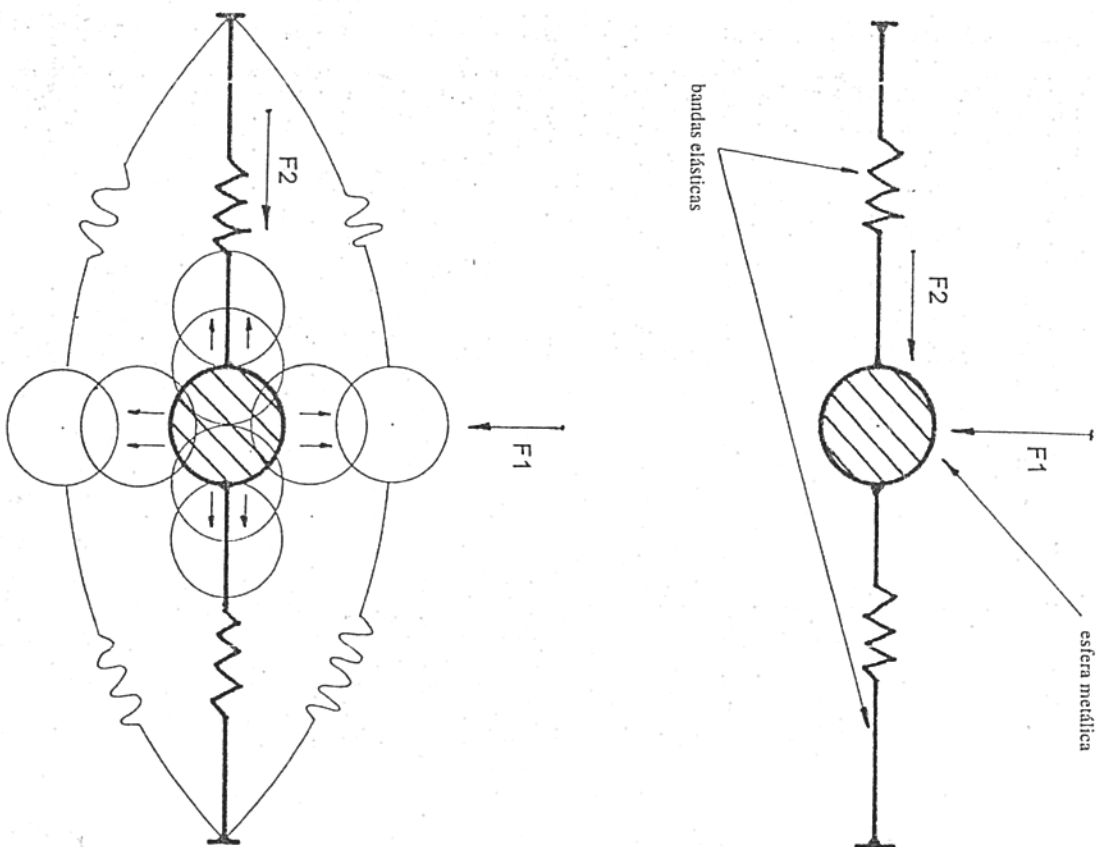


Figura 3. Senoides trazadas por las vibraciones de cada una de las moléculas de los cuerpos 1 y 2 (curvas superiores), y senoides resultantes de sumar las dos vibraciones de cada cuerpo (curvas inferiores).





FIGURAS 4 Y 5. Sistema físico estimulado por dos fuerzas diferentes  $F1$  y  $F2$ , que generan un movimiento vibratorio complejo.

ta según se les golpea en un lugar o en otro. Cada uno de los golpes hace vibrar al cuerpo en cuestión de una manera diferente y, en consecuencia, produce un ruido distinto. Pero veamos esto con un poco más de detalle.

Imaginemos una pequeña esfera metálica sujeta horizontalmente por dos bandas elásticas en la disposición que se muestra en la figura 4. Podemos aplicar a este sistema físico dos tipos de estímulos diferentes que lo harán oscilar de manera diferente. Si estiramos de la esfera, por ejemplo aplicando una fuerza  $F1$  de arriba hacia abajo para soltarla cuando las bandas estén tensas, se producirá una vibración de la bola de arriba hacia abajo y de abajo hacia arriba. El sistema permitirá también, aplicando una fuerza  $F2$ , que mueva la bola de metal hacia la derecha hasta tensar las gomas y soltándola bruscamente, que la esfera oscile de izquierda a derecha y de derecha a izquierda. Cada una de estas dos vibraciones tendrá una frecuencia y una amplitud diferentes, así que si sus frecuencias estuvieran dentro del espectro audible producirían dos sonidos distintos.

Imagine ahora el lector que aplicamos ambas fuerzas simultáneamente, es decir, que a la vez que estiramos de arriba hacia abajo tensamos la esfera hacia la derecha y volvemos a soltar bruscamente. Ahora la bola metálica se estará moviendo a la vez: de abajo arriba, de arriba abajo, de derecha a izquierda y de izquierda a derecha (fig. 5). En resumen, trazando un movimiento vibratorio sumamente complejo que de ser audible sería un sonido compuesto de dos frecuencias. A diferencia del ejemplo anterior, ahora la complejidad no depende de la heterogeneidad del cuerpo vibrante, sino de los distintos estímulos aplicados para hacerlo vibrar.

Finalmente, la tercera forma de construcción de un sonido compuesto es la suma de las dos anteriores. Es decir la de un cuerpo que, además de ser heterogéneo, está sometido a varios estímulos distintos que hacen vibrar cada una de sus moléculas de varias formas diferentes. El resultado es una vibración muy compleja compuesta por el sumatorio de tantas frecuencias simples como tipos de movimientos existen en el interior del cuerpo en cuestión.

### 3.1.2. La amplitud y la frecuencia en los sonidos puros

Creo que es importante empezar este apartado explicando las dos siguientes cuestiones previas.

En primer lugar, hemos visto que cualquier sonido compuesto es analizable: como una serie de frecuencias simples, o sonidos puros, que al sumarse linealmente dan como resultado el sonido compuesto estudiado. Esta posibilidad de análisis por descomposición en partes simples es lo que da sentido real al estudio de los sonidos puros. Sin partir de esta base conceptual no sería posible entender el funcionamiento de los sonidos cotidianos, que son siempre compuestos.

La segunda cuestión previa es que en el ámbito de la comunicación audiovisual se confunden a menudo el fenómeno físico de las vibraciones y la sensación que éstas producen en el sentido de la audición. Este problema tiene su origen en la relación directa que existe entre estos dos fenómenos: toda vibración que está

entre las 20 y las 20.000 oscilaciones por segundo es detectada por el hombre como un sonido, con lo que vibración se asimila a sonido, y sonido a vibración. No obstante, ni todos los fenómenos físicos vibratorios son audibles ni todo lo que oímos es traducible en fenómenos físicos vibratorios. Es necesario, en consecuencia, diferenciar siempre con toda claridad entre lo acústico como fenómeno físico y la sensación sonora como fenómeno psicológico ya que la relación entre ambos no es siempre isomórfica. En este apartado revisaremos los primeros y más simples puentes de relación entre estos dos fenómenos.

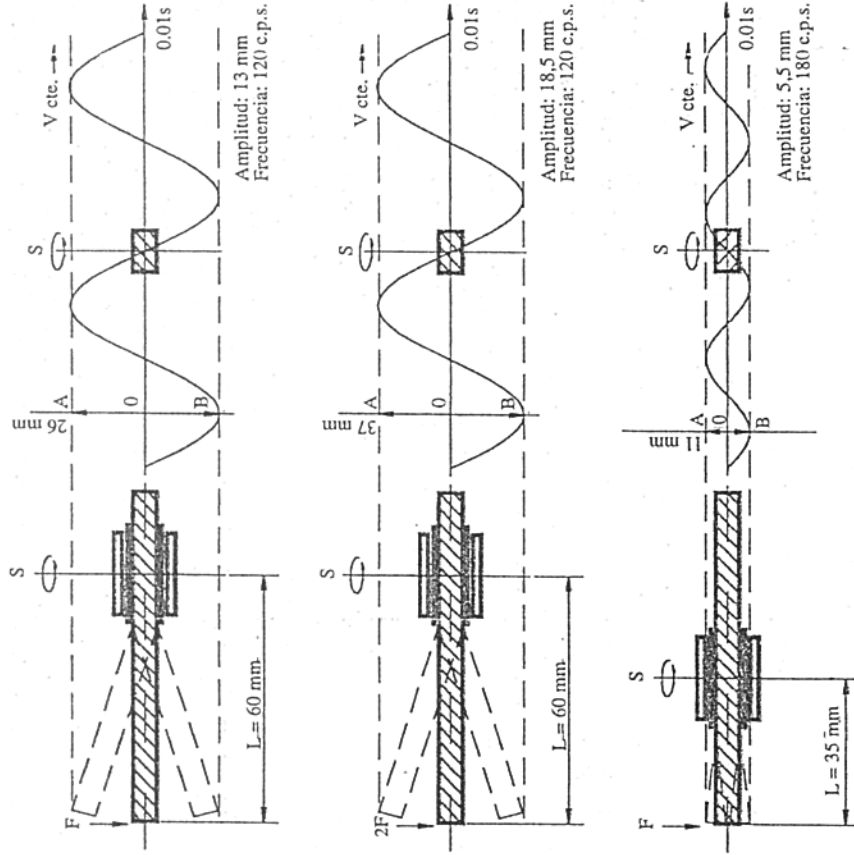
Explicados ya estos dos apriorismos que determinarán por completo nuestra forma de abordar y desarrollar la acústica, comenzaremos la aproximación a los sonidos puros con el concepto de *amplitud* y su vinculación a la sensación de *intensidad sonora*. Para ello recurriremos de nuevo a la imaginación auditiva del lector.

Suponga el lector que se encuentra frente a una barra metálica de masa y tensiones internas homogéneas, es decir, capaz de emitir sonidos puros, firmemente sujeta por uno de sus extremos, tal como se muestra en el dibujo de la izquierda de la figura 6. Si golpeásemos la barra aplicando una fuerza  $\cdot F$ , la barra se pondría a oscilar de arriba abajo emitiendo un sonido de frecuencia única. Imaginemos, también, que el sistema de mordazas que sujeta la barra permite el giro en torno a un eje de simetría  $\cdot S$ , tal como se muestra en la misma figura 6. Si hacemos girar la barra a la vez que oscila, su extremo seguirá una trayectoria ondulante describiendo un movimiento similar al que se ha dibujado en el lado derecho de la figura. Lógicamente, el pequeño rectángulo rayado de la derecha representa la punta de la barra de metal, y la línea ondulante la trayectoria que sigue al desplazarse a velocidad uniforme de izquierda a derecha mientras gira en torno al eje  $\cdot S$ .

Desde este tipo de representación del movimiento es posible relacionar con comodidad las distintas dimensiones perceptivas de la audición con la forma de vibrar la fuente sonora. Junto a la figura 6 se muestran la figura 7 y la figura 8 que nos permitirán desarrollar una serie de comparaciones para explicar estas relaciones acústico-vibratorias.

La propia experiencia vital nos muestra con frecuencia que cuanto más fuerte es el golpe que asestamos a algo (por ejemplo a nuestra barra de metal) más fuerte es el sonido que produce. Esto es lo que expresan las diferencias entre las figuras 6 y 7: Mientras la fuerza del imaginario golpe que hemos dado a la primera barra era  $\cdot F$ , la fuerza con que se ha golpeado la segunda es justamente el doble, es decir  $\cdot 2F$ . El resultado lógico de esta diferencia es que la segunda barra, la de la figura 7, vibra alejándose mucho más de su centro de reposo y suena, también, bastante más fuerte. Vemos, pues, que existe una relación evidente entre la *amplitud* con la que vibra la fuente sonora y la sensación de fuerza que produce el sonido generado por esta vibración. Dicho de otro modo: existe una

2. Recordemos que la *amplitud* es la distancia entre el punto de reposo del cuerpo vibrante y el lugar más alejado de éste al que llega el mismo cuerpo durante un ciclo oscilatorio. En el caso de las figuras 6, 7 y 8 sería la distancia entre los puntos  $\cdot O$  y  $\cdot A$  respectivamente.



FIGURAS 6, 7 y 8. Representación de la amplitud y la frecuencia desarrolladas por una barra de metal al estar sujeta y ser golpeada de formas distintas.

relación de dependencia directa entre la *amplitud* de cualquier vibración y la sensación de *intensidad*<sup>3</sup> que ésta produce en el oído humano. Como puede verse a la izquierda de las figuras 6 y 7, mientras la *amplitud* de la vibración producida por el primer golpe es de 26 milímetros, la amplitud producida por el segundo es de 37.

Así, la mayor o menor *amplitud* de las vibraciones que llegan hasta nuestros tímpanos queda traducida por el sistema perceptivo humano en sensación de más o menos *intensidad*. Entre ambos fenómenos existe, pues, una conexión directa;

3. Recordemos, también, que el concepto de *intensidad* sonora es el que sistematiza la sensación de sonidos más o menos fuertes.

no obstante, la relación entre ambos es muy compleja. Más adelante estudiaremos de qué modo es posible cuantificarla.

Pasaremos ahora a hablar de la relación entre la *frecuencia* de una vibración y la sensación auditiva de *tono*.

Recordaremos, en primer lugar, el sentido de estos dos conceptos: cuando hablamos de *frecuencia*<sup>4</sup> nos estamos refiriendo a la velocidad de vibración de cualquier cuerpo físico; y cuando hablamos de *tono* nos referimos a la sensación de agudo o grave que comporta siempre la percepción de cualquier sonido. Y para explicar de qué forma percibe el ser humano la frecuencia de vibración de un cuerpo, recurriremos de nuevo a la experiencia vital del lector.

Volvamos a la barra metálica del ejemplo anterior, pero ahora nos ocuparemos de la figura 8. Vamos a golpear la misma barra exactamente con la misma fuerza *F*, con que lo hicimos en el primer ejemplo (fig. 6), pero ahora la barra en lugar de estar sujeta colocando las mordazas en su extremo derecho está sujeta por el centro. En esta nueva disposición del sistema, la cantidad de masa que vibra es mucho menor, con lo que la recuperación elástica debida a las tensiones internas del metal es más eficaz. El sistema vibrante tiene menos masa y, por tanto, menos inercia, así que las fuerzas de recuperación interna le dejan alejarse menos de su centro de reposo (menor *amplitud*) y le hacen retornar a él mucho más rápidamente. Cuando ocurre todo esto, el golpe suena con menor intensidad pero mucho más agudo. En suma, el mismo golpe sobre la misma barra, ahora sujeta de otra manera, ha transformado parte de la energía vibratoria de modo que la barra se mueve ahora con mayor rapidez, generando un sonido más débil, pero mucho más agudo.

El fenómeno que describimos es exactamente el mismo que se desarrolla en una cuerda de guitarra cuando tocamos notas cada vez más agudas. A medida que bajamos la mano que trastea en el mástil de la guitarra, reducimos la longitud de la zona vibrante de la cuerda, haciendo así que la sensación auditiva que produce el tañido sea progresivamente más débil y aguda.

Observando la parte derecha de la figura 8 y comparándola con la de las dos figuras anteriores se puede comprobar cómo durante los 0,01 segundos representados gráficamente en las dos primeras situaciones, la barra desarrolla sólo 2 ciclos completos, mientras que en la tercera consigue desarrollar 3. Traduciendo esto en número de ciclos por segundo que es la unidad habitual, obtendríamos que las situaciones de las figuras 6 y 7 generarían una frecuencia de 120 hertzios<sup>5</sup> o ciclos por segundo mientras la situación de la figura 8 produce una vibración de 180 Hz.

Este tipo de situaciones tan cotidiano nos muestra de manera muy clara cómo se establece una relación entre la *frecuencia* de vibración de cualquier cuerpo (si

4. En el contexto de los movimientos vibratorios, el concepto de *frecuencia* se define como el número de recorridos completos que realiza un cuerpo oscilante durante un segundo.

5. El hertz (Hz) es la unidad de medida de la frecuencia de oscilación. Decimos que determinada vibración tiene una frecuencia de 1 hertz cuando su movimiento oscilatorio desarrolla un solo ciclo completo cada segundo. Así, una frecuencia de 2 Hz supone hablar de una vibración que oscila 2 veces por segundo, 10 Hz suponen 10 oscilaciones cada segundo, etc.

ésta está en la gama audible) y la sensación de tono. Cuanto más alta sea la frecuencia de una vibración, será percibida como una sensación sonora de tono cada vez más agudo. Nos encontramos, pues, con que la *frecuencia* de cualquier vibración simple es traducida por el sistema auditivo humano en sensación tonal. En el siguiente capítulo hablaremos de cómo se sistematiza esta relación frecuencia-*tono*, que es, también, considerablemente compleja.

### 3.2. Los instrumentos para analizar sonidos compuestos

#### 3.2.1. Sobre la concepción analítica de los sonidos compuestos

Más arriba decíamos que los *sonidos compuestos* son aquellos que están formados por dos o más frecuencias, es decir, son aquellas sensaciones sonoras que han sido producidas por vibraciones complicadas. También hemos visto en el apartado anterior cómo estas vibraciones complejas son siempre el resultado de que una misma masa vibrante se mueva simultáneamente con distintos tipos de oscilación.

Esta concepción de los sonidos compuestos tiene su origen en los estudios que hizo Hermann von Helmholtz en el siglo XIX utilizando esferas huecas de vidrio con dos cuellos tubulares cortos y abiertos. Cada esfera tenía sus dos cuellos situados en extremos diametralmente opuestos (Helmholtz, 1954). Uno de los cuellos se aplicaba al oído mientras el otro se acercaba a la fuente de sonido. Si la composición del sonido contenía una frecuencia igual o muy próxima a la frecuencia resonante de la cavidad del resonador, éste la amplificaba permitiendo percibirla aisladamente. Utilizando una serie de este tipo de resonadores era posible tener una idea muy aproximada de las distintas frecuencias que componían cada sonido estudiado. La deducción es directa y lógica: los sonidos complicados estaban compuestos por un conjunto organizado de otros sonidos más simples que era posible aislar y escuchar con los *resonadores de Helmholtz*.

Más tarde, la electrónica ha hecho posible transcodificar con comodidad el sonido en una señal eléctrica que varía del mismo modo que la señal sonora utilizando los micrófonos; y esta señal eléctrica variable puede transcodificarse a su vez en un dibujo en forma de onda que contiene prácticamente toda la información de la vibración: estamos hablando del *oscilograma*. A partir de este momento el *oscilograma* puede ser medido y tratado como una función matemática que varía al transcurrir el tiempo y que es descomponible en una serie de frecuencias simples mediante la herramienta matemática denominada *Transformada de Fourier*. Estas frecuencias resultantes al ser superpuestas dan siempre como resultado la vibración compleja original.

Finalmente, con la rapidez de cálculo y de toma de muestras que ha proporcionado la informática, y con el descubrimiento en 1965 del algoritmo denominado *Transformada Rápida de Fourier*<sup>6</sup> (véase Martí Roca, 1988, págs. 233-251),

6. En el ámbito más técnico es conocida por las iniciales FFT -Fast Fourier Transform-.

es posible disponer de equipos de análisis de audio capaces de analizar con gran rapidez la descomposición en frecuencias simples de cualquier sonido complejo.

Ciertamente, la concepción analítica que emana de todo el proceso histórico que acabamos de resumir es sumamente fértil. No obstante, como más adelante veremos, tiene limitaciones importantes. Su limitación consiste en que resulta extremadamente difícil utilizar conceptualmente un centenar de frecuencias distintas que evolucionan a la vez variando rápidamente en el tiempo.

Pero volvamos al origen. El punto de partida natural es siempre una vibración compleja que origina la sensación de sonido compuesto. Éste va a ser también nuestro punto de partida en el contexto audiovisual. Normalmente nos enfrentamos a sonidos compuestos extraordinariamente ricos y complicados que necesitamos tratar técnica y narrativamente para transmitir tal o cual efecto perceptivo a nuestro público. Pero para poder tratar un sonido necesitamos disponer de una cantidad de información sobre él lo suficientemente amplia. Veamos, pues, en primer lugar, de qué modo y con qué lógica estructural se pueden presentar actualmente los análisis de un sonido compuesto, y luego pasaremos a delimitar y explicar los instrumentos de trabajo que utilizaremos para obtener información sobre los sonidos en el resto de esta obra.

### 3.2.2. Los métodos gráficos de análisis acústico

Comenzaremos hablando del *oscilograma* ya que en la instrumentación actual éste es siempre el punto de partida de cualquier análisis acústico, seguiremos después explicando el *sonograma* como instrumento básico de análisis global, y pasaremos, por último, a estudiar el sentido puntual y «microscópico» de *espectrograma*.

El *oscilograma* es un instrumento gráfico fundamental para estudiar cualquier sonido compuesto. Como hemos dicho más arriba, el *oscilograma* es la representación gráfica en forma de onda de todas las variaciones eléctricas que ha generado un micrófono al ser estimulado por cierto sonido durante un tiempo concreto. Estas variaciones se extienden sobre ejes de coordenadas representándose la amplitud en el eje vertical y el tiempo en el horizontal. En la figura 9 podemos ver en la parte superior el oscilograma de una voz masculina pronunciando la vocal /AAA/ de forma continua, y en la parte inferior el oscilograma de la misma voz pronunciando la palabra /FONÉTICA/.

El primero es un sonido compuesto estacionario, es decir que no varía en el tiempo (naturalmente este carácter estacionario es sólo aproximado); y el segundo un sonido compuesto que varía en función del tiempo. Se puede observar perfectamente, comparando los dos oscilogramas, que mientras el primero está constituido por una oscilación compleja regular, el del gráfico inferior sufre variaciones muy importantes tanto en amplitud como en forma de onda a medida que nos desplazamos en el eje de tiempo (de izquierda a derecha). Ambos osci-

### OSCLOGRAMAS

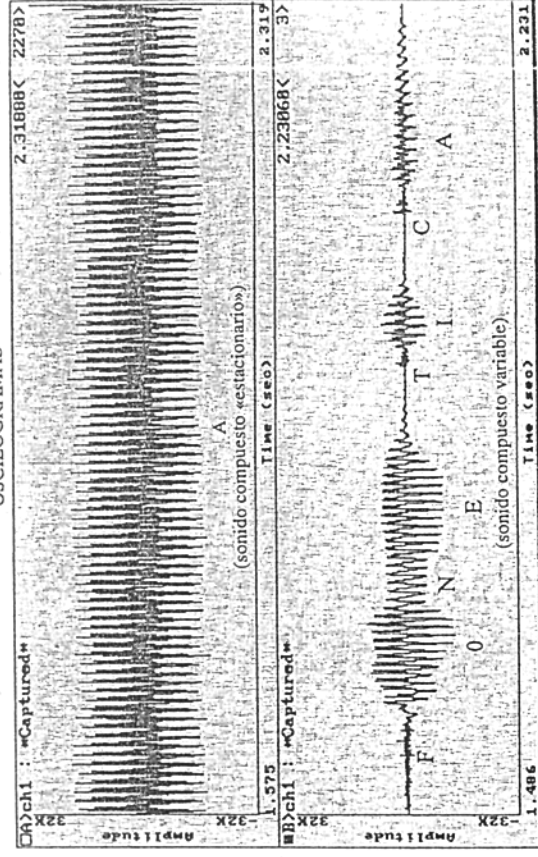


FIGURA 9. Oscilogramas del sonido de una voz pronunciando la vocal A (gráfico superior); y pronunciando la palabra FONÉTICA (gráfico inferior).

logramas representan sendos fragmentos sonoros de 0,74 segundos. El lector puede comprobar esto restando la cifra del extremo inferior izquierdo<sup>7</sup> de la del extremo inferior derecho<sup>8</sup> respectivamente en cada gráfico. En el segundo oscilograma se ve claramente cómo cada uno de los distintos fonemas de la palabra /FONÉTICA/ configura un tipo de representación vibratoria perfectamente diferenciada.

A partir del tratamiento matemático de estos dos oscilogramas es posible obtener información sobre su composición de otras dos formas distintas: generando *sonogramas* y generando *espectrogramas*. Pero antes de entrar con detalle en estas dos nuevas formas de presentación de la misma información acústica desarrollemos una propuesta gráfica global que proporcionará al lector una idea intuitiva de cómo realizan las distintas transformaciones de los datos del *oscilograma* los sistemas de análisis electroacústico.

Observe ahora el lector la figura 10. En este dibujo están representadas dentro de un espacio tridimensional las cuatro frecuencias simples que constituyen un

7. La cifra expresa el momento en que empieza la representación gráfica del oscilograma (1.486 seg.) respecto al momento (0 seg.) que sería el inicio absoluto del sonido que el sistema ha capturado y tiene completo en su memoria.

8. La cifra expresa el momento en que termina la representación gráfica del oscilograma (2.231 seg.) respecto al momento (0 seg.) que sería el inicio absoluto del sonido que el sistema ha capturado y tiene completo en su memoria.

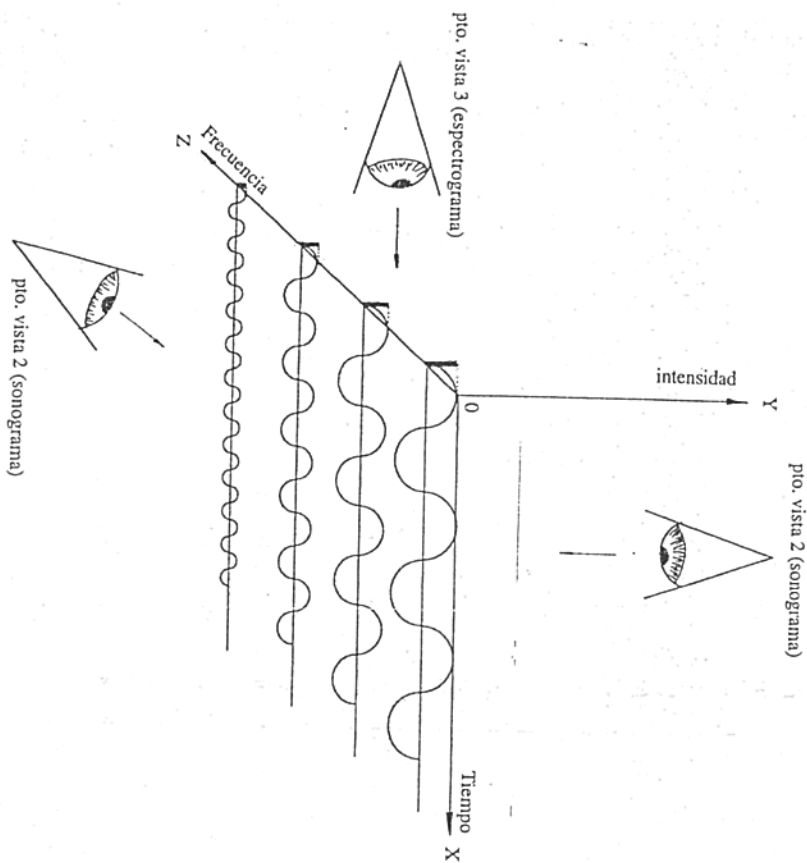


Figura 10. Representación tridimensional de un sonido compuesto por cuatro frecuencias simples.

sonido compuesto. Las ondas vibratorias están ordenadas de menor a mayor frecuencia según estén más o menos cerca del punto -0- en los ejes de coordenadas. Para poder observar y tomar datos de nuestro sonido compuesto, ya desglosado gráficamente en cuatro frecuencias, hemos de definir puntos de vista concretos y uno los tres planos (X-Y, X-Z, Y-Z), que definen geoméricamente el espacio tridimensional, determina uno los tres métodos clásicos de representación gráfica que se utilizan para estudiar los sonidos. Es decir: el *oscilograma*, el *sonograma* y el *espectrograma* respectivamente.

El *oscilograma*. Vayamos en primer lugar al *pto. de vista 1* de la figura 10. Si observásemos las cuatro curvas desde este punto de vista, o sea, colocando nues-



Figura 11. Medida de la duración del fonema /l/ en la palabra FONÉTICA utilizando el oscilograma.

tros ojos exactamente frente al plano X-Y, tal como se indica en el dibujo, veríamos todas las senoides aplastadas y amontonadas una sobre otra formando una única madeja de líneas que se extiende de izquierda a derecha en este plano según avanza el tiempo. La forma de que esta «madeja» de líneas aporte información coherente y ordenada es transformando las cuatro frecuencias simples en una única curva compleja que sea el resultado de la superposición lineal de las cuatro anteriores. Si hiciésemos esto, el *pto. de vista 1* nos proporcionaría la representación gráfica del sonido denominada *oscilograma*.

Este tipo de representación de los sonidos es especialmente eficaz para trabajar sobre la dimensión temporal. Dicho de otro modo: el *oscilograma* nos permite medir con mucha precisión la duración de los *objetos sonoros* (ver el apartado 2.3.3). Mediante el oscilograma del sonido /FONÉTICA/, por ejemplo, podemos saber cuál es la duración de cada uno de los fonemas que lo componen. Podemos situar los cursores del sistema de análisis en los dos extremos de la /l/ (véase fig. 11), tomar los tiempos, restarlos, y obtener así su duración exacta. Este tiempo exacto aparece en la primera cifra del ángulo superior derecho del gráfico; 0,04760 segundos. O lo que es lo mismo: 47 milésimas de segundo. La única forma de cuantificar tiempos de esta magnitud es utilizando este instrumento gráfico, por supuesto, sería imposible medir este *objeto sonoro* con un cronómetro mientras se escucha la palabra.

Mediante el mismo método podemos conseguir, por ejemplo, la duración media de los distintos sonidos de toda la palabra /FONÉTICA/.

Pero, ¿cuál es la utilidad de todo esto en el ámbito de la expresión audiovisual?

Pongamos por caso que decidimos obtener datos sobre la duración media de las vocales y las pausas de locuciones tensas y agresivas para compararlos con otros datos sobre locuciones tranquilas y relajadas. Este tipo de información proporcionaría referencias objetivas muy útiles en el doblaje. Cada actitud emocional concreta que perseguimos podría estar formalizada por una serie de datos de referencia que servirían, tanto al director como a los dobladores, para saber si se está construyendo exactamente, o no, la actitud que se intenta construir.

Veamos otro ejemplo. La reverberación artificial es un recurso expresivo especialmente habitual en el contexto audiovisual. Cuando el tiempo de reverberación supera la duración media de los fonemas, el texto sonoro reverberado se



hace completamente ininteligible. Así pues, conocer con precisión la duración media de los fonemas del texto concreto que queremos manipular es una información crucial para decidir el nivel de reverberación artificial que podemos aplicar en cada texto.

Un último ejemplo para ilustrar la utilidad de este instrumento gráfico. Sin duda, el uso más cotidiano de la medición sobre el oscilograma es el que se hace en el montaje audio-vídeo mediante la edición informatizada. Actualmente, para ajustar la sincronía entre imágenes y sonido se manipula informáticamente el audio recortando fragmentos sonoros aquí o allá de forma cuantificada y exacta. El lector debe tener en cuenta que éstos son sólo unos pocos ejemplos de una gama de posibilidades de uso extremadamente amplias de las que ahora sólo hemos citado alguna.

*El sonograma.* El «pto. de vista 2», que contempla las curvas de la figura 10 desde arriba (plano X-Z), es el que generaría la imagen del *sonograma*. Al mirar las líneas trazadas por las cuatro vibraciones a vista de pájaro tendríamos la sensación de que éstas se han transformado en rayas lisas (la perspectiva adoptada nos esconde el efecto de ondulación). No obstante, estas líneas que ondulan acercándose y alejándose de nuestro ojo según aumenta o disminuye la amplitud de la vibración, o lo que es lo mismo, la intensidad del sonido, sí que producirían un efecto visual de mayor grosor y oscuridad a medida que se acercan a nuestro ojo. Así, el sonograma traduce el aumento de intensidad del sonido (amplitud de la vibración) en mayor oscuridad y grosor de las líneas. A mayor intensidad, mayor oscuridad. Cuantas más frecuencias intensas se concentran en una zona concreta mayor es el grosor de la línea oscura.

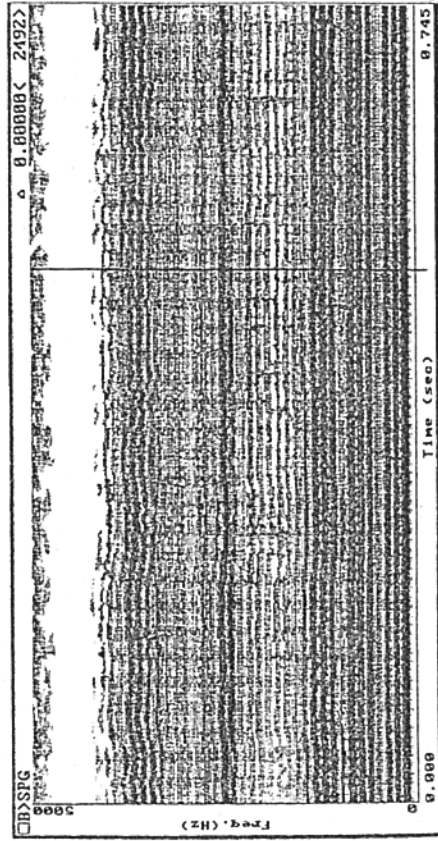


FIGURA 12. Sonograma del sonido de una voz pronunciando la vocal /A/. Es el mismo sonido que se estudió en la figura 9, gráfico superior.

Al ser vistas sobre el plano X-Z, las curvas de la figura 10 estarían representadas, entonces, por cuatro líneas rectas horizontales que se extenderían de izquierda a derecha, siendo la más cercana al eje «X» la más oscura de las cuatro, puesto que es la que tiene mayor amplitud. Y de acuerdo con el citado dibujo, cada una de las tres siguientes sería menos oscura a medida que disminuye su respectiva amplitud.

En la figura 12 podemos observar el sonograma real de un sonido estacionario, es el de la vocal /A/ que ya habíamos estudiado en la figura 9, ahora representado como una serie de líneas de mayor o menor oscuridad que se extienden de izquierda a derecha durante un tiempo de 0,000 a 0,745 segundos.

Pero ésta no es la única información que proporciona el *sonograma*. Siguiendo siempre la misma lógica representacional, las frecuencias simples obtenidas en el análisis estarán ordenadas de menor a mayor frecuencia según estén más o menos cerca del punto «0». Consecuentemente, cada una de las líneas que contempláramos al mirar la figura 10 desde el «pto. de vista 2» está ubicada a una distancia concreta de «0» al cortar el eje «Z». La distancia que va del punto «0» al punto en que cada línea intersecta el eje «Z» nos da información sobre la altura tonal de cada uno de los sonidos simples que componen el análisis. O lo que es lo mismo, sobre la frecuencia exacta de cada una de las vibraciones simples que constituyen la frecuencia compuesta. Volviendo, pues, a la figura 12, puede verse como el eje situado a la izquierda del *sonograma* (se corresponde con el eje «Z» de la figura 10) es el que indica la frecuencia en hertz «Freq. (Hz)» y se extiende, en este caso, de 0 a 5.000 hertzios.

El *sonograma* es la representación gráfica de un sonido compuesto que nos proporciona más cantidad de información simultánea. Sobre el eje vertical «Z» en la figura 10) obtenemos información de la altura tonal; el grado de oscuridad de las líneas nos proporciona información sobre la intensidad; y, finalmente, tal como ocurría con el *oscilograma*, el eje situado en la parte inferior del *sonograma* («X» en la figura 10) aporta información temporal. Es decir, nos encontramos frente a una representación que da información sobre los tres parámetros sonoros a la vez: *intensidad, frecuencia y tiempo*.

El *sonograma* ha sido muy utilizado en fonética y bastante menos en comunicación audiovisual. No obstante, es un instrumento absolutamente imprescindible en la primera etapa de cualquier análisis acústico. Cuando no sabemos absolutamente nada de un sonido que necesitamos tratar expresivamente, el punto de partida será siempre el análisis sonográfico.

*El espectrograma.* Finalmente, el «pto. de vista 3» para contemplar las curvas de la figura 10 sobre el plano Y-Z generaría el aspecto gráfico de una serie de barras verticales de distinta longitud, todas alineadas. A este tipo de representación es a lo que se denomina *espectrograma*. En el plano Y-Z de la figura 10 podemos observar la proyección de lo que sería el *espectrograma* de nuestro imaginario sonido compuesto por cuatro frecuencias simples. Al mirar «lateralmente» las curvas, solamente podríamos ver sus distintos perfiles como barras o líneas, tanto más alejados del punto «0» cuanto más alta sea la frecuencia de cada una de las vibraciones. Igual que en el sonograma, la inter-

sección con el eje «Z» nos aportará la frecuencia exacta de cada uno de los sonidos simples.

Por otra parte, cada una de las curvas, al ser vista de «lado» adoptará el aspecto de una barra vertical tanto más larga cuanto mayor sea su amplitud, es decir, cuanto mayor sea la intensidad del sonido simple que representa. La proyección de cada una de estas barras sobre el eje «Y» proporciona información precisa sobre la intensidad de cada uno de los sonidos simples.

Con el fin de que la representación espectrográfica sea visualmente más significativa, en lugar de representar las barras de cada una de las curvas alineadas por su centro, tal como correspondería a la imagen que generaría de forma natural el «plo. de vista 3», se representa sólo una de las mitades de la imagen, de modo que queden alineadas siempre con uno de sus extremos sobre el eje «Z». De esta manera, en lugar de presentar un aspecto simétrico a ambos lados de un eje central, el *espectrograma* muestra el aspecto de una cresta en la que las diferencias que-  
dan siempre acumuladas en el mismo extremo. Las diferencias de amplitud (intensidad) se hacen así más evidentes y fáciles de manejar.

Respecto al *espectrograma*, es muy importante observar que la información que proporciona respecto a la composición de frecuencias que tiene un sonido, y respecto a la intensidad de cada una de estas frecuencias, está siempre referida a un solo instante concreto de la evolución del sonido en el tiempo. Es un corte transversal del suceso sonoro, algo así como una fina rodaja que muestra con mucho detalle el contenido de un salchichón, pero sólo en el lugar exacto donde la hemos cortado, ignorando todo lo demás; a no ser que lo cortemos en su totalidad en infinitas rodajas sucesivas para estudiarlas una a una. Nos permite analizar al detalle la composición de la materia acústica en un punto temporal preciso, pero pierde completamente de vista el resto de la evolución sonora anterior y posterior al instante seleccionado.

Las barras representadas en el plano Y-Z de la figura 10 configuran el *espectrograma* del sonido-ejemplo de cuatro frecuencias con el que estamos trabajando. Puede observarse cómo cada una de las barras es más alta que la anterior a medida que nos acercamos al punto «0», respondiendo estas alturas a las respectivas amplitudes de cada una de las vibraciones. El *espectrograma* representado respondería a un análisis realizado justo al iniciarse el sonido, digamos que en el segundo 0,0001.

Veamos ahora un *espectrograma* real. En la figura 13 se muestra el *espectrograma* de la misma /A/ que hemos estudiado ya varias veces. Observe el lector que el espectro de frecuencias se muestra bajo un *oscilograma*.

El *oscilograma* muestra un suceso sonoro de 0,744 segundos de duración, con un cursor situado aproximadamente en el centro, exactamente en el segundo 0,33195 (cifra de la esquina superior derecha). En este punto es en el que se ha realizado el análisis espectral que se muestra en el gráfico de debajo. Siguiendo el modelo de representación global con el que estamos trabajando, el *espectrograma* viene a ser la contemplación frontal de un corte transversal del *oscilograma* realizado en el punto T = 0,33195 seg.

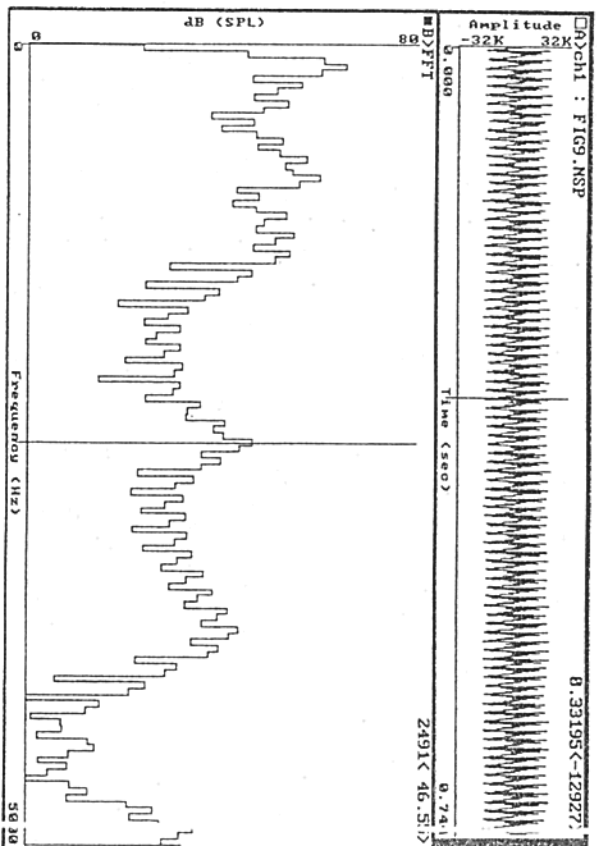


Figura 13. Gráfico inferior: espectrograma del sonido de una voz pronunciando la vocal /A/ desarrollado en momento: T = 0,33195 seg. Es el mismo sonido que se estudió en la figura 9.

Podemos observar que, en este instante temporal, la masa sonora de la /A/ está representada por una cresta de 41 puntas que indican la presencia de 41 frecuencias simples distintas entre los 0 y los 5.000 Hz, cada una de ellas con una intensidad diferente, componiendo el sonido complejo /A/. La proyección de cada una de las puntas de la cresta sobre el eje horizontal inferior («Z» en la figura 10) nos proporciona datos sobre las frecuencias respectivas «Frequency (Hz)» y su proyección sobre el eje vertical izquierdo («Y» en la figura 10) aporta datos concretos sobre cada una de las intensidades. El cursor que aparece en el *espectrograma* al cruzarse con el espectro indica un punto concreto del que se dese extraer datos. En la parte superior derecha del recuadro inferior pueden leerse «-stos datos, es decir, su frecuencia e intensidad exactas: «2491<46.55», es decir, 2.491 hertzios y 46,55 decibelios.

Hay que tener en cuenta, también, que el *espectrograma* de la figura 13 está representado sobre un fondo-escala de 5.000 Hz. O sea, analiza la composición de la /A/ entre 0 y 5.000 hertzios. Pero podría estar desarrollado, también sobre otros fondos-escala, por ejemplo, de 10.000, de 15.000 o de 20.000 Hz, con lo que obtendríamos mayor información sobre la composición de la masa sonora de la /A/ en las frecuencias altas. Más adelante, cuando hablemos de la percepción del tono en el capítulo 4, veremos otro tipo de re-

presentación del espectro de frecuencias más adaptada que éstas a la percepción humana.

El *espectrograma* es un tipo de representación analítica de los sonidos compuestos muy habitual en el ámbito técnico de la comunicación audiovisual, y veremos a lo largo de esta obra que puede serlo también el ámbito expresivo. Es muy común realizar análisis espectrográficos de los sonidos que están mezclados con ruidos no deseados. Esta «suciedad» suele introducirse en las grabaciones de efectos sonoros en entornos naturales, en las grabaciones de diálogos con sonido directo, etc. En estos casos se realizan análisis espectrales en varios puntos de la grabación para determinar cuál es la composición frecuencial de estos ruidos e intentar eliminarlos filtrando las frecuencias no deseadas.

Otro de los usos habituales de *espectrograma* es el estudio de las resonancias típicas de las salas de proyección cinematográfica, de audición, de teatro, etc. Toda sala, al proyectar dentro de ella un sonido y rebotar éste contra sus paredes, genera una serie de reflexiones. Cuando esto ocurre, las vibraciones reflejadas se superponen al sonido original reforzando unas frecuencias y cancelando o atenuando otras. Este fenómeno modifica la composición de las distintas intensidades originales, generando en el interior de la sala un sonido compuesto resultante que puede ser muy distinto del que había previsto el realizador de la banda sonora, o del que tendría la misma banda sonora en una sala distinta.

Proyectando dentro de cualquier local un ruido blanco<sup>9</sup> y realizando un análisis espectral de este mismo ruido después de que haya recorrido el local, se detecta inmediatamente sobre qué zonas de frecuencia influye la forma y el volumen de la sala. Una vez que se dispone de la información espectral sabremos qué tipo de modificaciones introducirá ese local concreto sobre cualquier sonido que suene dentro de él. Desarrollado ya este proceso de análisis, normalmente es posible corregir la influencia acústica de la sala mediante una manipulación electrónica del sonido que contrarreste esa influencia antes de su amplificación.

Actualmente existen sistemas de tratamiento del audio en el contexto de la industria audiovisual como el de *Lucas Film Entertainment*, denominado THX, que exigen a sus exhibidores comerciales que las salas de proyección cumplan con unas características acústicas muy estrictas de reducción de niveles de ruido y de tiempos de reverberación en función de la frecuencia (véase Siguero, 1995, pág. 54). De otro modo, todo el esfuerzo para el tratamiento expresivo del audio que ha realizado la productora podría resultar completamente inútil al ser desuido por la acústica propia de la sala de exhibición. Como hemos visto un poco más arriba, el modo de saber si la sala reúne o no las condiciones es siempre el análisis espectral, o algún método derivado de él.

9. El *ruido blanco* es un ruido homogéneo que se obtiene sumando un gran número de vibraciones sinusoidales muy próximas que han de cumplir las siguientes condiciones técnicas: a) las amplitudes de todas sus vibraciones han de ser muy similares; b) las frecuencias de estas vibraciones deben cubrir toda la banda audible (20 Hz a 20 kHz); c) las fases de todas estas vibraciones deben estar repartidas sobre todo el dominio angular de forma estadísticamente uniforme.

### 3.2.3. La interrelación entre los tres métodos clásicos de análisis acústico

De hecho hemos ido relacionando ya constantemente *oscilograma*, *sonograma* y *espectrograma* apoyándonos en el modelo tridimensional de la figura 10; no obstante, es evidente que ése es un modelo ideal que lo que aporta es una concepción global intuitiva. Así que nos queda, todavía, estudiar de qué modo se relacionan entre sí estos tres instrumentos sobre representaciones de sonidos reales:

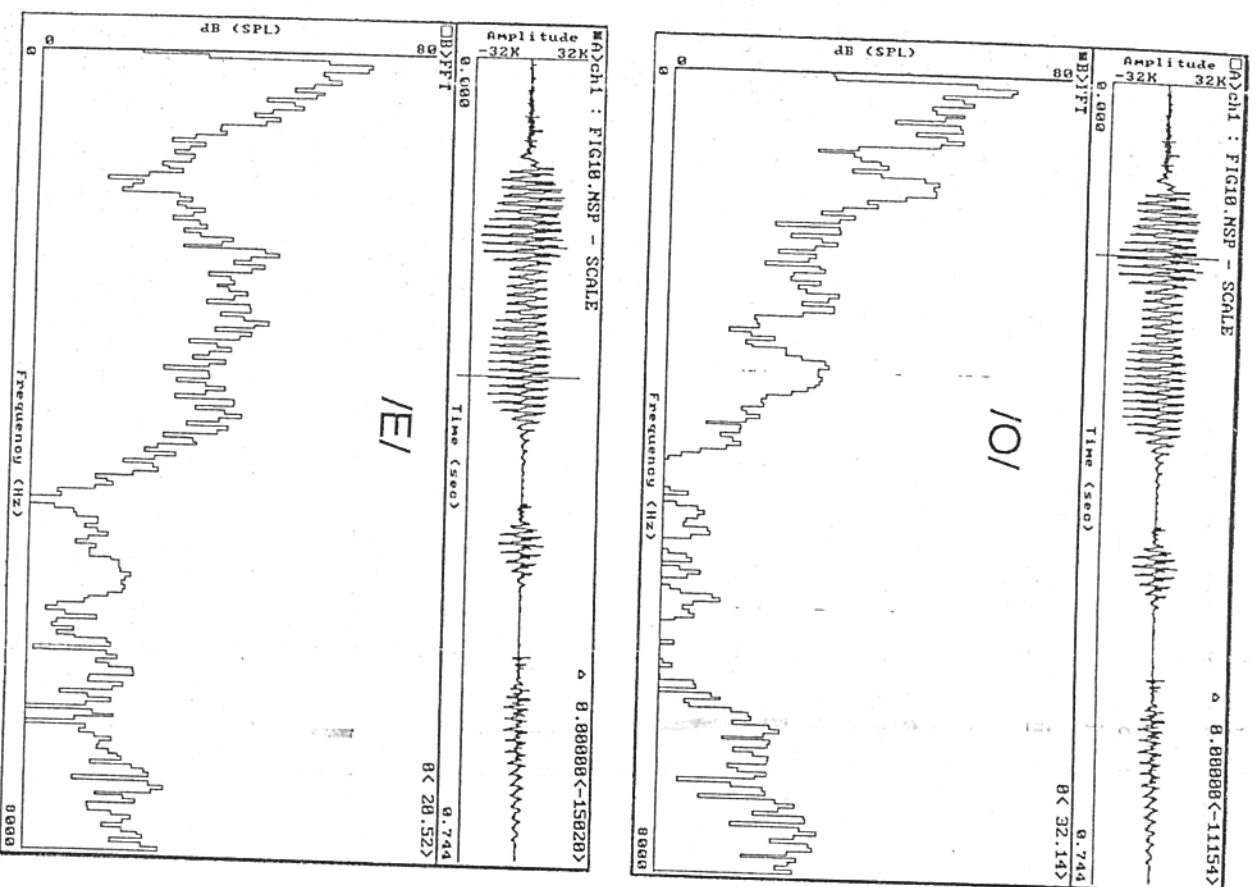
*Del oscilograma al espectrograma.* Hemos visto ya a grandes rasgos que el *oscilograma* era la primera etapa de referencia visual a partir de la cual se elige un instante concreto del suceso sonoro para obtener de él su espectro de frecuencias. Lo que ocurre, en realidad, es que el sistema de análisis toma una serie de muestras en el entorno inmediato en el que hemos situado el cursor y, aplicando la Transformada Rápida de Fourier (véase apartado 3.2.), calcula y representará gráficamente un *espectrograma*. La relación básica entre estas dos representaciones gráficas es que el *oscilograma* actúa como punto de partida sobre el cual el analista decide los puntos temporales que quiere estudiar, y a partir de ellos el sistema de análisis toma las muestras, calcula y construye el *espectrograma*.

Las figuras 14 y 15 ilustran con claridad cómo analizando el *oscilograma* del suceso sonoro /FONÉTICA/ se obtienen *espectrogramas* completamente distintos según el cursor esté situado en la zona de la /O/ o de la /E/, y volvería a resultar diferente en cualquier otra. Esto ocurre porque nos enfrentamos a un sonido variable. Naturalmente, si trabajamos con sonidos estacionarios, es decir, que mantienen su composición de frecuencias homogénea a lo largo del tiempo, el *espectrograma* será el mismo aunque realicemos análisis en distintos puntos temporales del *oscilograma*.

De todas formas, hay que señalar, también, que los conceptos de *variable* y de *estacionario* son siempre relativos al tipo de análisis que necesitamos desarrollar. En un suceso sonoro que durante 1 segundo ha sufrido 10 variaciones espectrales, seguramente podremos localizar zonas estacionarias de 0,005 segundos, y viceversa. Si tomamos un suceso sonoro que se muestra estacionario durante un segundo, y ampliamos lo suficiente la muestra antes y después, casi con toda seguridad localizaremos variaciones. Cuanto menos, encontremos el inicio y el final y, en consecuencia, las formas de variación que constituyen el *ataque* y la *caída*, que pueden ser muy diversas.

*Del espectrograma al sonograma.* Al describir la figura 10, decíamos que el *espectrograma* era una vista lateral del perfil que configura todo el conjunto de frecuencias de un sonido en un punto temporal concreto; y que el *sonograma* era algo así como mirar desde arriba el cuerpo del sonido viendo toda su extensión temporal. Veamos ahora esto con análisis gráficos de sonidos reales.

En la figura 16 hemos colocado juntos, uno debajo de otro, el *espectrograma* y el *sonograma* del sonido estacionario de vocal /A/ que hemos estudiado ya varias veces (véanse figs. 12 y 13). En esta imagen, el *espectrograma* (gráfico superior)



Figuras 14 y 15. Espectrogramas de las vocales /O/ y /E/ obtenidos analizando distintos puntos temporales en el oscilograma del suceso sonoro /FONÉTICA/.

aparece del mismo modo que lo habíamos visto hasta ahora. El sonograma, en cambio, ha sido manipulado respecto a cómo lo vimos en la figura 12 de dos formas diferentes: a) con un giro de 90 grados en el sentido de las agujas del reloj; y b) ampliando la longitud del eje que representa gráficamente las frecuencias de 0 a 5.000 Hz hasta hacerlo coincidir exactamente con la longitud del mismo eje de representación de frecuencias en el *espectrograma*.

Al comparar ambas representaciones del mismo sonido, podemos observar cómo en la situación sobre el eje horizontal (*Frecuencia*) cada una de las puntas de la cresta que constituye el *espectrograma* se corresponde de una forma muy aproximada con la situación en este mismo eje, ahora en el *sonograma*, de las franjas oscuras que representan las distintas frecuencias simples que componen la /A/. Y las franjas del *sonograma* aparecen tanto más oscuras cuanto mayor es la altura de la punta de cresta con la que se corresponde en el *espectrograma*, es decir, cuanto más alta es la intensidad de cada una de las frecuencias.

Resulta muy claro, por ejemplo, observando ambos gráficos de izquierda a derecha, que el *espectrograma* comienza con una intensidad muy baja en la zona próxima al 0, y que esta zona está representada con una estrecha franja blanca en el *sonograma*. Inmediatamente después observamos en el *espectrograma* cómo aparece la punta de mayor altura del espectro sonoro; si el lector mira el *sonograma* en la misma zona comprobará cómo ahí está también la franja más oscura y ancha de toda la imagen sonográfica. Siguiendo hacia la derecha las franjas del *sonograma* se oscurecen o se aclaran correspondiéndose respectivamente con las subidas y bajadas de la cresta espectral. Hacia el final, vuelve a verse con mucha claridad cómo la bajada global de la altura del espectro se representa en el *sonograma* como una zona blanca, bastante más ancha que las anteriores.

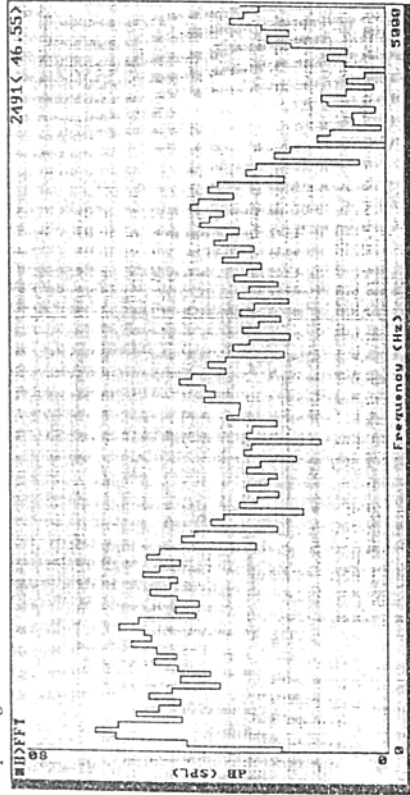
Un último ejercicio de imaginación.

Decíamos más arriba que el sonograma venía a ser la observación de sonido a vista de pájaro, es decir, desde arriba (*Pto. de vista 2* en la fig. 10). Podemos pensar, entonces, que ver el sonograma de la figura 16 es algo así como si estuviésemos mirando desde arriba un taco de madera en forma de cuña orlulada, que varía de 1 a 6 centímetros de grosor, y al que se le han hecho una serie de estrías profundas en su parte superior. Si el lector agarra este «taco de madera sonográfica» con sus dos manos de modo que los pulgares queden debajo del papel y lo gire levantando el perfil hacia sus ojos de manera que pudiese mirarlo frontalmente, verá exactamente el perfil que muestra el espectrograma de la figura 16.

*Del oscilograma al sonograma.* La correspondencia entre el *sonograma* y el *oscilograma* es esencialmente temporal. Ya vimos que el *oscilograma*, como instrumento de análisis, queda limitado básicamente al estudio minucioso de la duración de los sucesos sonoros. Sobre representaciones oscilográficas que muestran duraciones entre 0,5 y unos 6 o 7 segundos, suelen estar muy claras las variaciones de amplitud del *oscilograma*, y esto nos permite, normalmente, detectar con eficiencia la articulación sonora; es decir, el principio y el final de las variaciones del sonido y, por tanto, tomar medidas temporales entre el inicio y la terminación de los distintos sucesos sonoros. No obstante, observando directamente:



Espectrograma del sonido estacionario /A/



Sonograma del mismo sonido estacionario

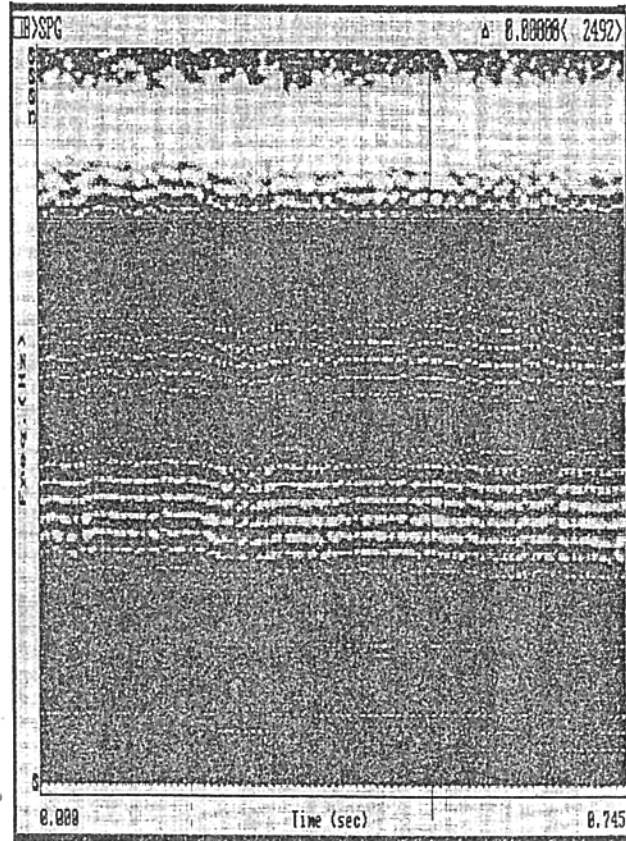


FIGURA 16. Representación espectrográfica (superior) y sonográfica (inferior) de la misma vocal /A/. Obsérvense las correspondencias visuales entre ambas.

la forma de onda no es posible obtener información útil sobre la composición acústica, a no ser que el sonido sea extremadamente simple.

En cambio, en la observación del *sonograma* ocurre justamente lo contrario. Aporta una información mucho más precisa y completa sobre la composición que sobre la duración de los sucesos sonoros. Es cierto que el sonograma presenta los sucesos sonoros que analizamos desarrollados en el tiempo. Pero, un sonograma muestra, simultáneamente, una serie de franjas oscuras (frecuencias simples) que evolucionan sufriendo cada una de ellas variaciones distintas, que muchas veces no coinciden en el tiempo. Esto hace que, a menudo, sea difícil decidir sobre el *sonograma* dónde empiezan y dónde acaban las variaciones sonoras. Veamos un ejemplo de esto sobre la figura 17.

En esta imagen presentamos de nuevo el *oscilograma* (gráfico superior) y el *sonograma* (gráfico inferior) del sonido oral /FONÉTICA/. Si el lector observa detenidamente, por ejemplo, la zona del sonograma correspondiente a la /O/, notará cómo este sonido no es homogéneo en el tiempo. Mirándolo de abajo hacia arriba podemos ver en primer lugar dos franjas que se extienden de manera continua y prácticamente sin variaciones desde la línea discontinua que señala el inicio de la /O/ hasta la que marca el final de la /É/. Posteriormente tres franjas se ajustan perfectamente a las líneas discontinuas que delimitan la /O/, pero, a partir de ahí y siguiendo hacia arriba, se forma algo parecido a un arco que no tiene ya nada que ver con lo que hemos marcado como límites del sonido /O/. Esta variación tan irregular en el tiempo de las distintas frecuencias se debe a que, en realidad, no existe un sonido estacionario /O/, sino un proceso de sucesivas transiciones sonoras, que el sonograma nos muestra como un conjunto de sonidos simples que evolucionan para componer la mutación de la /F/ a la /O/ pasando luego a ser /N/ sin que haya ningún momento de discontinuidad.

Así, para decidir solamente a partir del sonograma dónde comienza y acaba el sonido /O/ tendríamos algunos problemas. No ocurre esto en cambio con el *oscilograma*, en el que podemos observar con claridad cómo dos reducciones perfectamente definidas de la amplitud determinan el paquete de la primera vocal /O/. Si el lector vuelve a la figura 11, podrá ver el mismo oscilograma sin ninguna señal en torno a la primera vocal que entorpezca su observación y comprobar cómo, efectivamente, resulta mucho más fácil decidir dónde están el comienzo y el final de este sonido sobre el *oscilograma* que sobre el *sonograma*.

Es evidente que la combinación de ambos instrumentos gráficos de análisis del sonido resulta perfectamente complementaria. Aquella información que no puede aportar el *sonograma* la proporciona el *oscilograma* y viceversa. Por esa razón, la primera etapa de los análisis acústicos suele desarrollarse observando a la vez el *oscilograma* y el *sonograma*, y trabajando en el sistema de análisis<sup>10</sup> con ambos instrumentos gráficos lineados; es decir, sincronizados, ajustados en el tiempo.

10. Si es de su interés, el lector puede encontrar una información amplia y detallada sobre los fundamentos de este tipo de sistemas de análisis acústico en la obra *Fonética* de Martínez Celdrán (1984, págs. 91-130).



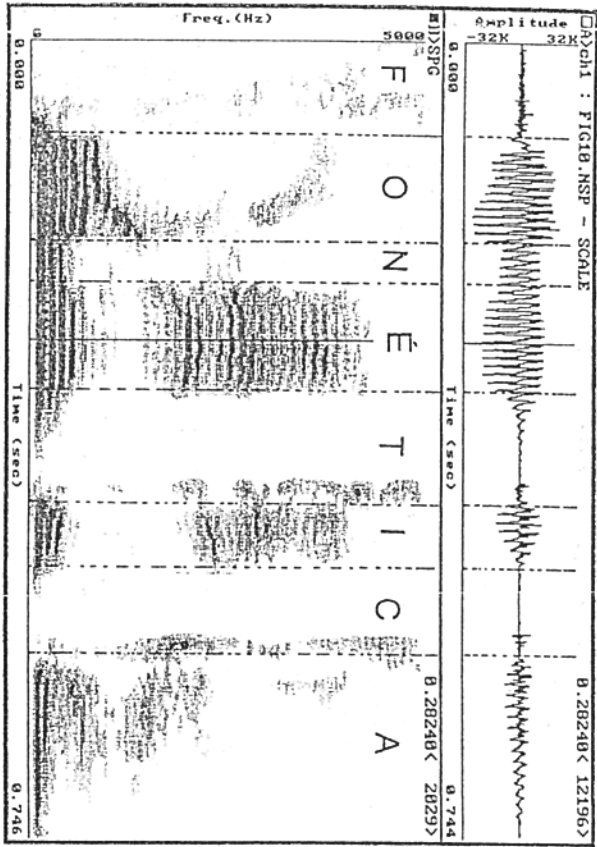


Figura 17. Representación oscilográfica (superior) y sonográfica (inferior) del mismo suceso sonoro /FONÉTICA/. Obsérvense las correspondencias visuales entre ambas.

Cuando *oscilograma* y *sonograma* están lineados, los cursores que se desplazan sobre ambos gráficos actúan como si fuese uno solo. Así, si trabajando sobre el *oscilograma* colocamos el cursor, por ejemplo, en el inicio de la /O/, puesto que ahí es perfectamente localizable el punto temporal donde comienza el suceso sonoro, tenemos ya la garantía de que el cursor indicará, también sobre el sonograma, cuál es el punto de partida temporal del fenómeno acústico que estamos estudiando. La gran ventaja de la asociación entre *sonograma* y *oscilograma* es que permite localizar con comodidad el inicio y final de los sucesos sonoros sobre ambos gráficos a la vez, solucionando el problema de que en el *sonograma* eso resulta visualmente mucho más confuso.

Tanto en el *oscilograma* como en el *sonograma* de la figura 17, aproximadamente en el centro de la /E/, pueden observarse sendos cursores representados con línea continua, que indican exactamente el mismo punto temporal en ambos gráficos. El lector puede comprobar, en el ángulo superior derecho de los dos dibujos, cómo el tiempo señalado por ambos cursores es exactamente el mismo: 0,28240 segundos. Es decir, puede comprobar que han actuado, en realidad, como un cursor único.

### 3.3. El tono y el timbre en los sonidos compuestos

En este apartado, estudiaremos y propondremos algunos conceptos que facilitarán la comprensión del espectro acústico de los sonidos compuestos.

Se trata de revisar conceptos como los de *armónico*, *parcial*, *formante* y *envolvente*, poniéndolos en relación con el de tono y, especialmente, con el de timbre; entendiendo que este último es una herramienta fundamental para la interpretación perceptiva de los sonidos complejos.

#### 3.3.1. La sensación de tono en los sonidos compuestos

Al escuchar un sonido compuesto estamos oyendo a la vez un conjunto muy numeroso de frecuencias que suenan a la vez. Cuando este sonido es *consonante*, es decir, cuando se percibe por las características de su composición como un sonido único y no como muchos sonidos a la vez, produce una sensación de altura tonal concreta. Cualquiera de los instrumentos musicales a los que estamos habituados se modula en altura tonal a pesar de que emite sonidos compuestos. Las teclas del piano, por ejemplo, cada una de ellas produce una sensación auditiva concreta que va de las más graves a la izquierda del teclado a las más agudas a la derecha. No obstante, el sonido de cada una de las notas es un sonido compuesto considerablemente complicado.

Esta sensación tonal unitaria que producen los sonidos compuestos viene determinada perceptivamente por la más baja de las frecuencias que componen el sonido. A esta frecuencia de valor perceptivo tan relevante se la denomina *frecuencia fundamental* o *pitch*. Encontraremos habitualmente el término *pitch* en la literatura técnica y el término *frecuencia fundamental* en la literatura lingüística, musical y psicológica.

Esa primera frecuencia o *frecuencia fundamental* actúa en los sonidos armónicos como pauta organizativa para el resto de frecuencias que componen el sonido. Así, los sonidos que denominamos *armónicos* son sonidos compuestos en los que todas las frecuencias que los constituyen son múltiplos de la *frecuencia fundamental*. Veamos un ejemplo. Una voz masculina muy grave puede tener una *frecuencia fundamental* de 80 Hz. Si esto es así y su voz está bien timbrada, el resto de las vibraciones que la componen tendrán unas frecuencias que serán siempre múltiplos de 80 Hz. Es decir, la voz estará compuesta por una *frecuencia fundamental* de 80 Hz y una serie de armónicos que se situarán de una forma bastante aproximada en torno a los: 160, 240, 320, 400, 480, 560, 640, 720, 800, 880, 960, 1040, 1120, 1200, 1280, 1360... Hz.

De todas formas, los sonidos no siempre están tan perfectamente organizados. A menudo, se da en los sonidos una mezcla entre frecuencias que son múltiplos de la fundamental, es decir de *armónicos*, con otras frecuencias que no tienen este tipo de relación numérica con el *pitch*. En estos casos los sonidos son auditivamente bastante más desagradables y producen un efecto de «suciedad» sonora, o de sonido ruidoso. En cualquier caso, todo sonido tiene siempre una frecuencia

inicial más baja que todas las demás, que será su *frecuencia fundamental*, y que es la que predomina perceptivamente en la sensación de altura tonal.

En la figura 18 podemos ver dos espectrogramas resultantes de analizar dos sonidos distintos de /E/ que ilustran lo explicado más arriba. Ambos tienen una *frecuencia fundamental* muy similar, unos 245 Hz, como se puede observar por la situación del cursor y la primera cifra en la esquina superior derecha de ambos gráficos.

Esto significa que ambas /E/ transmiten una sensación tonal muy similar. Perceptivamente, la diferencia entre ambas /E/ está precisamente en que el sonido de la primera (gráfico superior) es mucho más limpio y agradable que el de la segunda (gráfico inferior), que es un sonido «sucio». El sonido que representa en la figura 18 el *espectrograma* inferior está compuesto, a la vez, por frecuencias que son armónicas con la *frecuencia fundamental*, es decir, múltiplos de ella, y por otras frecuencias que no tienen ninguna relación numérica regular con la fundamental. Estas otras frecuencias no armónicas con la fundamental, pero que habitualmente forman parte también de la composición de los sonidos se las denominan *parciales*. Si el lector observa con detenimiento ambos *espectrogramas* y los compara, notará que la cresta del gráfico superior está formada por una serie de picos alargados, perfectamente definidos, separados por una distancia regular formando un perfil global ondulante. Éste es el aspecto habitual de los *armónicos*; cada una de las largas puntas de la cresta es una frecuencia armónica con la fundamental. En cambio, la cresta del gráfico inferior tiene una forma mucho más indefinida, los picos son mucho más cortos, la distancia entre ellos es irregular y la forma global de la cresta en lugar de ser ondulante es completamente «scarpada». Así es como se presentan normalmente los *parciales* en el espectrograma.

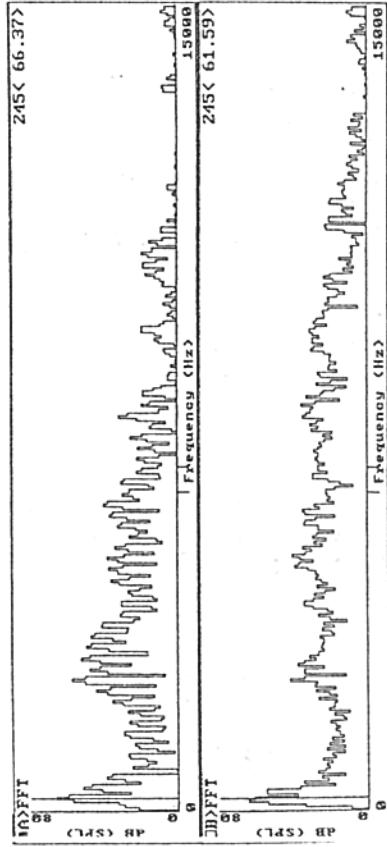


FIGURA 18. Representación espectrográfica de dos vocales /E/ distintas: la primera «superior», de gran armonización y transparencia; y la segunda (inferior), sucia y nada armónica.

Los sonidos están tanto más organizados armónicamente cuanto mayor es la parte del espectro ocupada por *armónicos*. Y viceversa, cuanto más baja es la frecuencia a la que comienzan a aparecer *parciales*, mayor es la desorganización sonora.

Pero veamos adónde nos lleva todo esto desde el punto de vista de la sensación tonal, llamada también *tonía*. Cuanto mayor es la organización armónica de las frecuencias que componen un sonido, mejor definida está la sensación de altura tonal que éste transmite, y podemos precisar su frecuencia auditivamente con bastante exactitud. Contrariamente, cuando la relación entre los componentes del sonido es desordenada e irregular, es decir, cuando predominan los *parciales* sobre los *armónicos*, la sensación tonal es mucho más confusa y resulta mucho más difícil precisar su altura tonal.

Si el lector compara de nuevo los dos *espectrogramas* de la figura 18, ahora observando ya hasta donde se extienden los armónicos, notará que mientras en el primero llegan más allá de la mitad del gráfico (aproximadamente hasta los 8.000 Hz), en el inferior prácticamente no hay presencia de armónicos.

La figura 19 muestra exactamente las mismas representaciones sonográficas que en la figura anterior, pero ahora con el cursor desplazado hasta el extremo más alto de la cresta después de la *frecuencia fundamental*. Si el lector cuenta sobre el *sonograma* superior el número de armónicos que hay desde la frecuencia *fundamental* o *pitch* hasta el lugar en el que está situado el cursor comprobará sin ningún problema que son 11. Si todo lo explicado hasta aquí es cierto, la *frecuencia fundamental* de este sonido debería ser aproximadamente la onceava parte de la frecuencia que indica ahora el cursor (primera cifra de la esquina superior derecha) es decir de 2.483 Hz. Realicemos la operación:

$$2.483 \text{ Hz} / 11 = 225,72$$

Hemos obtenido una cifra que se aproxima mucho a la frecuencia de 245 Hz que conseguimos colocando el cursor sobre el primer *armónico* o *frecuencia fundamental* (véase fig. 18). Este simple proceso de cálculo es uno de los que se utiliza habitualmente para localizar la frecuencia fundamental sobre un espectro, cuando el primer *armónico* no aparece bien definido visualmente. El lector puede, también, comprobar por sí mismo observando el gráfico inferior de la figura 19 cómo esta forma de calcular la *frecuencia fundamental* solamente es posible si estudiamos un sonido bien organizado armónicamente.

Todo lo explicado hasta aquí sobre el tono de los sonidos compuestos a partir de dos *espectrogramas* puede observarse también sobre los *sonogramas* correspondientes.

En la figura 20 podemos ver de nuevo el análisis de los mismos sonidos de /E/ limpia y armónica (izquierda) y de /E/ sucia e inarmónica (derecha), ahora en sendos *sonogramas* y sobre un fondo-escala de 10.000 Hz. En estas imágenes vuelve a ser muy evidente cómo las sombras que se extienden de izquierda a derecha, representando las distintas frecuencias simples, tienen un aspecto completamente distinto en cada gráfico. Mientras en el de la izquierda (/E/ «limpia») las sombras se ordenan con claridad en franjas horizontales representando los *armónicos*, en el

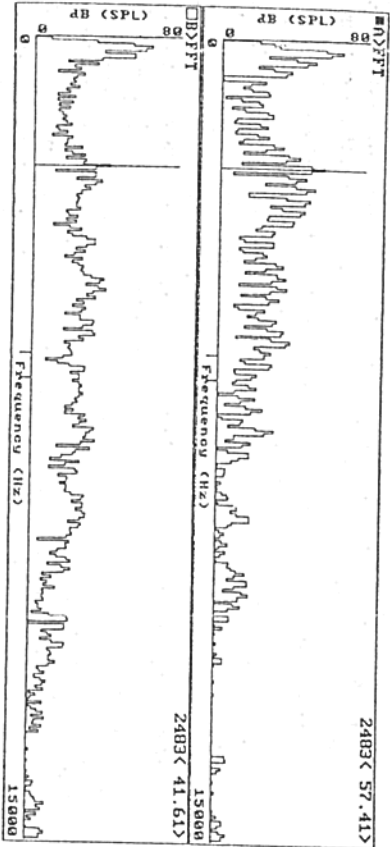


FIGURA 19. Representación espectrográfica de dos vocales /E/ distintas: la primera (superior) de gran armonicidad y transparencia, y la segunda (inferior) sucia y nada armónica.

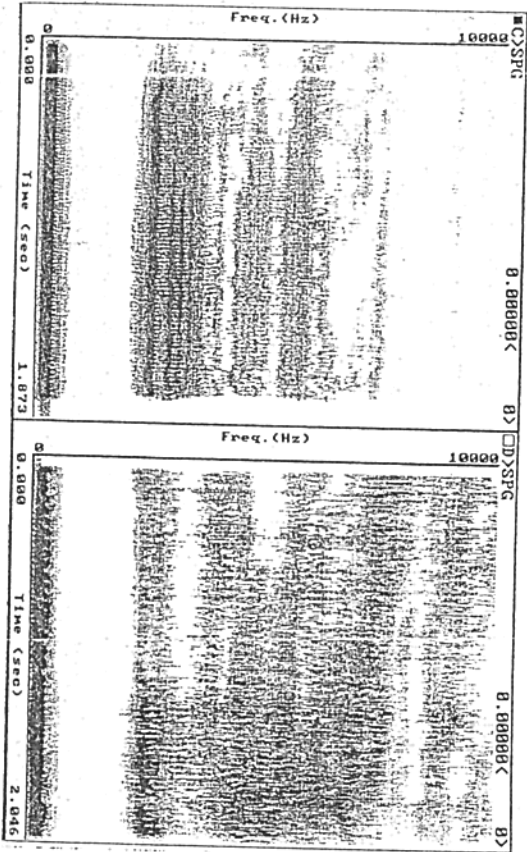


FIGURA 20. Representación sonográfica de dos vocales /E/ distintas: la primera (izquierda), de gran armonicidad y transparencia, y la segunda (derecha), sucia y nada armónica.

de la derecha (/E/ «sucio») la organización de las sombras es mucho más escasa representando frecuencias *parciales* que se distribuyen por el espectro sonoro (e una forma muy irregular.

3.3.2. El concepto de timbre

Antes de entrar en la problemática estricta de la sensación tímbrica y de estudiar los métodos para poder relacionarla con los resultados del análisis acústico, revisaremos el concepto mismo de *timbre*.

El concepto de *timbre* ha sido siempre uno de los que ha causado más problemas de comprensión al explicarlo en mis clases. Probablemente porque existe una idea social demasiado elemental y simplista respecto a qué es el *timbre*. El *timbre* suele entenderse como la dimensión acústica que da el carácter individual a los sonidos: cuando dos sonidos, a pesar de tener la misma intensidad y el mismo tono, se perciben como diferentes entre sí, su diferencia se debe al *timbre*. A menudo se define también como aquella parte del sonido de la voz o del sonido de un instrumento musical que es invariable, para oponerlo al *tono* y la *intensidad* como dimensiones sonoras variables y fácilmente manipulables.

Esa concepción del *timbre* como una dimensión sonora invariable e inherente en exclusiva a las características de la fuente sonora supone, por ejemplo, que una misma voz no podría variar jamás su *timbre*. Paradójicamente, una forma esencial de organización sonora de la lengua es el reconocimiento de las variaciones del *timbre* fonético. Cuando un hablante cualquiera construye oralmente, por ejemplo, una serie de vocales cambiando la posición de su boca, lo que hace es modificar el *timbre* del sonido que inmediatamente antes ha producido su laringe. Por lo tanto, el *timbre* no es una característica invariable de la fuente sonora, sino que varía constantemente, como oyentes cotidianos y permanentes del habla de aquellos que nos rodean, identificamos los distintos sonidos que componen cada palabra, lo que estamos haciendo es diferenciar y reconocer los respectivos timbres de cada uno de los fonemas que han construido nuestros interlocutores.

Solamente la observación de este fenómeno tan cotidiano del uso de la variación tímbrica en el habla debería ser suficiente para que se cuestionase profundamente y a nivel general el valor del concepto */timbre/*, e intentásemos corregir adecuadamente su definición. Sin embargo, es posible encontrar esa concepción simplista del *timbre*, en lugares tan diversos como el Diccionario de la Real Academia de la Lengua Española,<sup>11</sup> obras sobre foniatría,<sup>12</sup> obras técnicas sobre audio,<sup>13</sup> obras sobre radio,<sup>14</sup> etc. No obstante, la idea de que el *timbre* es algo ba-

11. Modo propio y característico de sonar un instrumento músico o la voz de una persona. (Real Academia Española, 1992, pág. 1.979).  
12. Ese algo sutil e indefinible que hace que dos voces, al cantar la misma nota, conserven su individualidad. (Percelló, 1975, pág. 69).  
13. La calidad o color de tono única de un sonido. (Allen, 1994, pág. 635).  
14. Varios sonidos de igual intensidad y tono y cuya onda sonora tenga diferente estructura, se distinguirán por el timbre y permitirán reconocer diversas características del emisor. (Cebrián Herrero, 1994, pág. 276).

tante más complicado que un simple rasgo diferenciador ha sido defendida con mucha claridad por algunos estudiosos desde hace más de 40 años. Edouard Garde en su obra *La Voix*, editada por primera vez en 1954, definía ya la noción de timbre como una dimensión muy compleja, y diferenciaba en el timbre de una misma voz: *color, volumen, densidad y mordiente* (Garde, 1960, págs. 37-40).

En general, en el estudio del canto hay una larga tradición de clasificación del timbre de la voz con términos más o menos afortunados pero siempre numerosos y diversos. Al margen de la potencia o volumen (*intensidad*) y de la extensión tonal, se habla de *voz blanca, negra, brillante, metálica, redonda, blanda, dura, sorda, oscura, nasal, gutural, cálida, ...* Y cada voz suele ser descrita con combinaciones concretas de varios de estos calificativos (véase Rodríguez Bravo, 1984, pág. 73).

Pero cuando el tipo de instrumentos de análisis que hemos descrito a lo largo del apartado 3.2 llegan a la fonética y a la investigación musical es cuando la concepción elemental de *timbre* que estamos discutiendo queda definitivamente obsoleta y se abre todo un universo de investigación en torno a esta tercera dimensión del sonido. El *timbre* pasa a contemplarse, entonces, como lo que es en realidad: un entramado sumamente complejo de caracteres acústicos y sensoriales auditivos. De hecho, las últimas investigaciones experimentales sobre los caracteres acústicos y perceptivos que configuran la sensación auditiva del *timbre* desembocan sistemáticamente en un tipo de definiciones que reflejan muy claramente lo inadecuado, o mejor, lo insuficiente de este término:

... lo que diferencia una vocal de otra u otras, aunque la frecuencia de sus componentes sea igual, es la distinta estructuración de sus armónicos, cuya percepción es lo que denominamos timbre (Quilis, 1981, pág. 142).

Altura y sonoridad son características definidas comparadas con la complejidad del timbre. [...] El timbre es mucho más complejo. Es verdad que en cierta manera depende del espectro del sonido [...]. Nuestros conocimientos sobre el timbre son aún incompletos, pero se ha aprendido mucho en los últimos años gracias al análisis de las ondas sonoras [...] (Pierce, 1985, pág. 170).

La sensación de intensidad sonora no tiene más que una dimensión. Un sonido puro o un sonido compuesto no puede ser más que igual, más fuerte o más débil que otro. Incluso la altura de un sonido no tiene más que una dimensión. Hablamos de sonidos agudos o graves. El timbre del sonido, al contrario, de múltiples dimensiones, corresponde a la gran variedad de organización posible de los parciales. No se han encontrado todavía nombres apropiados para caracterizar con precisión el timbre de un sonido [...] (Zwicker, 1981, pág. 5).

Probablemente esa concepción del *timbre* como una dimensión invariable del sonido que depende exclusivamente de la fuente sonora que lo produce, tiene su origen en la sistematización de la música occidental. Nuestro sistema musical es esencialmente tonal. El código de notación musical permite dar cuenta por escrito de una serie cerrada de categorías tonales concretas, y los instrumentos musicales de Occidente están diseñados también para reproducir exactamente esa mis-

ma gama de categorías tonales. La mayor o menor energía con la que se sopla en un instrumento de viento, o el grado de fuerza con que se pellizcan, golpean o frotan los instrumentos de cuerda determinan inevitablemente variaciones de intensidad; los cambios de intensidad se contemplan también en el solfeo, pero de una manera muchísimo menos precisa que la evolución tonal. Lo que en ningún caso contempla nuestro código de signos musicales es la variación tímbrica.

La variación del timbre en la música occidental se limita al cambio de instrumento. Para conseguir otro tipo de matiz sonoro sobre una misma altura tonal sencillamente se elige otro instrumento. Esa forma de organizar y dar cuenta escrita de los sonidos determina muy claramente una tendencia a no observar ni clasificar los matices tímbricos que van más allá de las diferencias globales entre un instrumento y otro. Una tendencia que ha impregnado toda nuestra cultura sonora.

Tal como constatabamos un poco más arriba, la preocupación por las variaciones y por un conocimiento más desarrollado del *timbre* proviene del estudio de la voz cantada y del estudio de los sonidos de la lengua.

Todo hablante ha observado en más de una ocasión cómo, en determinadas situaciones de salud, o cuando vive ciertos estados emocionales, su voz cambia de matiz, de «color sonoro». En esos estados, la voz sigue pudiendo variar su tono y su intensidad sin ningún problema, pero hay algo que la hace sonar globalmente distinta. Lógicamente, este fenómeno tiene mucha más trascendencia en alguien, como un cantante, que utiliza profesionalmente su voz y que necesita que su rendimiento sonoro sea homogéneo. Así, la preocupación por aspectos de la voz como el *brillo*, la *dureza*, el *color*, la *claridad*, la *calidez*, el *mordiente*, etc., que son conceptos independientes de los de *tono e intensidad*, suponen de forma inequívoca una concepción del timbre completamente distinta por parte de los estudiosos de la voz cantada. Esta concepción, mucho más abierta, es la que conecta directamente con las corrientes de investigación actuales sobre el timbre, especialmente en el contexto de la psicoacústica y la fonética experimental.

### 3.3.3. La sensación tímbrica

El concepto de *timbre* no es acústico sino psicológico. Es la denominación de un tipo de sensación auditiva. Pero como el lector habrá observado ya en el apartado anterior, no es en absoluto un concepto unidimensional, sino un concepto considerablemente ambiguo y complejo que debe ser desarrollado, a su vez, en varias dimensiones. Eso será lo que haremos en este apartado. Pero, antes, proponemos una nueva definición para el *timbre*:

El timbre es una sensación auditiva compleja (independiente de las de duración, tono e intensidad y simultánea a ellas) que nos permite percibir la estructura acústica interna de los sonidos compuestos.

Es cierto que esta definición es, también, considerablemente ambigua. No obstante, creemos que tiene las siguientes virtudes frente a la concepción tradicional:



a) es coherente con ella; b) no contempla el timbre como algo simple; y c) permite ser desglosada y ampliada en función de los nuevos conocimientos que se obtengan sobre esta dimensión perceptiva, conservando el concepto genérico de *timbre*.

La dificultad para enfrentarse a este concepto sobre la percepción sonora es tan notable que algunos autores como Bruce Goldstein (Bruce Goldstein, 1993, págs. 380-429) evitan usarlo directamente y optan por utilizar sólo los conceptos como «*resonancia*», «*percepción categorial*», «*formante*», etc., para tratar la percepción de la estructura sonora interna; fragmentando así una sensación que, a nuestro modo de ver, debe ser tratada finalmente como algo global, a pesar de la descomposición que exige su complejidad. Como vamos a ver más adelante, éstos son, precisamente, aspectos puntuales que componen la sensación tímbrica global.

En una primera aproximación, podemos encontrar en la sensación tímbrica tres dimensiones bien diferenciadas: *armonicidad*, *impresión espectral* y *definición auditiva*. Veamos con detalle cada uno de estos tres conceptos utilizando los instrumentos de análisis acústico que tenemos ya a nuestra disposición.

• *Armonicidad*. La sensación de *armonicidad* o *inarmonicidad* será probablemente la más familiar y fácil de comprender para el lector. De hecho hemos estado haciendo ya referencia a ella cuando en el apartado 3.3.1 habíamos de la sensación de *tono* en los sonidos compuestos, e introducíamos el concepto de *parcial* para diferenciar entre *armónicos* (frecuencias múltiplo de la frecuencia fundamental) y *parciales* (frecuencias no organizadas respecto a la fundamental). En este apartado mencionábamos sonidos más o menos ricos en armónicos y concluíamos que la sensación de *tono* podía ser más o menos difusa o sucia en función de lo bien o mal organizado que estuviese el sonido en armónicos. Lógicamente, si hemos definido el *timbre* como la sensación auditiva que transmite la estructura acústica interna de los sonidos compuestos, el grado de organización de un sonido en armónicos y parciales ha de estar necesariamente asociado a la sensación tímbrica. Lo está concretamente con la *armonicidad*.

Denominaremos *armonicidad* al distinto grado de limpieza y agradabilidad que percibimos al escuchar un sonido compuesto, dependiendo de la relación que existe en su espectro entre armónicos y parciales. Cuanto mayor sea la gama de frecuencias organizada armónicamente, mayor será la sensación de limpieza y agradabilidad, es decir: mayor será la *armonicidad*.

Hemos llamado a esta primera dimensión de la sensación tímbrica *armonicidad* para diferenciarla del concepto musical de *armonía*. En la terminología musical se denomina *armonía* al conjunto de normas que ordenan y regulan las relaciones entre las partes de una composición, para que sea posible articularlas entre sí en función de unos principios acústicos (véase Valls Gorina, 1986, pág. 16). De hecho, las normas de la armonía musical se apoyan también en las re-

laciones de multiplicidad respecto a la frecuencia fundamental. No obstante, la composición musical occidental trabaja mezclando sonidos compuestos de instrumentos musicales que son acústicamente muy complejos y distintos entre sí. Esto complica enormemente el análisis acústico de las composiciones sonoras; y hace que los problemas de la armonía musical desemboken en soluciones mucho más relacionadas con los hábitos culturales que con los mecanismos perceptivos.

El análisis acústico, concretamente el espectrograma, refleja con mucha claridad el grado de armonicidad del sonido. Los sonidos compuestos que perceptivamente son claros, limpios y agradables son sonidos que tienen una gran parte de su espectro ocupada por armónicos. Éstos se presentan siempre de izquierda a derecha. Es decir, desde la *frecuencia fundamental* hacia las frecuencias más altas de una forma contigua. Cuanto más armónico es un sonido, tanto mayor es la superficie del espectro de izquierda a derecha ocupada por una cresta formada por dientes de sierra muy largos, ordenados y bien definidos.

Si el lector vuelve a revisar la figura 19, ahora desde este nuevo punto de vista, podrá observar claramente cómo el espectrograma superior representa, sin lugar a dudas, un sonido con un grado de *armonicidad* mucho mayor que el espectrograma inferior. Ya explicamos antes que tanto el gráfico superior de esta figura como el inferior correspondían a sendos sonidos de /E/, pero mientras el superior corresponde a una voz limpia, resonante y agradable, el inferior está hecho a partir de una voz desagradable, sucia y ensordecida. Mientras en el espectrograma inferior (voz de escasísima armonicidad) sólo son claramente visibles 2 armónicos en la zona izquierda, al principio del espectro; en el espectrograma superior (voz de mucha armonicidad) son perfectamente visibles 38 armónicos, que ocupan más de la mitad del gráfico.

Parece existir, pues, una relación directa entre la sensación tímbrica de *armonicidad* y la relación proporcional entre *armónicos* y *parciales* que podemos observar visualmente al hacer el análisis espectral de un sonido. Cuanto mayor sea la superficie espectral ocupada por *armónicos*, mayor será la sensación de *armonicidad*.

• *Impresión espectral*. La organización de las *resonancias* o *formantes* a lo largo del espectro configuran una sensación tímbrica que podríamos denominar genéricamente como *impresión espectral*. Definiremos este concepto en el siguiente modo:

Llamaremos *impresión espectral* a la sensación de diferente matiz auditivo que percibe un receptor, cada vez que escucha un sonido de idéntica composición de frecuencias, pero con distinta envolvente espectral. Entendiendo como envolvente espectral la forma obtenida al trazar una línea que une todas las puntas de cresta de cualquiera de sus espectrogramas posibles.

Hemos visto ya con detalle distintas representaciones espectrales y hemos estudiado sus relaciones con el sonido. Pero para poder desarrollar adecuadamente



esa dimensión de la percepción tímbrica que hemos denominado *impresión espectral* es necesario que utilicemos como punto de partida el concepto de *resonancia* o *formante*. Éste es un concepto muy habitual en el ámbito del análisis acústico que permite comprender e interpretar algunos aspectos importantes en el espectro de los sonidos compuestos.

Se denomina *formante* a la gama de frecuencias de un sonido compuesto que han quedado reforzadas en amplitud por la forma y el volumen del espacio en el que se ha producido, o la forma y el volumen del espacio en el que se está produciendo el sonido en cuestión.

Cuando un sonido compuesto se emite dentro de un espacio cerrado o semicerrado, este sonido incide en todas las paredes, resultando modificado en su composición al sumarse linealmente con su propia reflexión múltiple, de modo que unas frecuencias quedan reforzadas, otras se debilitan y algunas se cancelan. A los espacios o cavidades que producen este tipo de efecto se les llama resonadores y pueden ser de tamaños y formas totalmente diversos: cavidades bucales, cajas de resonancia de los instrumentos musicales, paredes, techo y suelo de las salas de audición... Perceptivamente, este tipo de fenómeno matiza el sonido originario de modo que éste genera una impresión auditiva más brillante y luminosa, más oscura y difusa, etc. No obstante, este tipo de matización sonora mantiene constante la sensación de *tono* y la *intensidad*.

Desde el punto de vista acústico, una *resonancia* es perfectamente reconocible tanto en el espectrograma como en el sonograma como un grupo de frecuencias contiguas que han aumentado de intensidad en torno a una frecuencia concreta a modo del perfil de una montaña. Volvamos a nuestra ya familiar *A/estacionaria* para observar esto en sus representaciones espectrográfica y sonográfica.

En el espectrograma (gráfico de la página siguiente) de la figura 21, de izquierda a derecha, podemos observar, en primer lugar, el *armónico* de frecuencia más baja o *primer armónico*, que a partir de ahora nombraremos ya siempre como *frecuencia fundamental*. A continuación, se extienden el resto de los *armónicos*, a modo de dientes de sierra perfectamente equidistantes entre sí a lo largo de todo el espectro. El lector notará cómo los armónicos aumentan de intensidad como si se tratara de la cresta de una montaña en torno a cuatro puntos concretos del espectro. Estas zonas bien definidas de acumulación de energía acústica son las *resonancias*, y a su forma tamaño y distribución en el espectro se debe la sensación tímbrica de *impresión espectral*.

Si observamos, ahora, el mismo sonido representado en el sonograma (gráfico inferior), podemos ver que las franjas estrechas y regulares que representan los armónicos se oscurecen también progresivamente definiendo bandas más oscuras. Estas bandas son las resonancias. Igual que en el espectrograma, en el sonograma están señaladas, también, la frecuencia fundamental y las cuatro resonancias.

La situación en el espectro de cada resonancia y su anchura (franja de frecuencias que la componen) influyen de una manera determinante en la *impresión espectral* concreta que produce un sonido. Cuanto más hacia la zona alta del es-

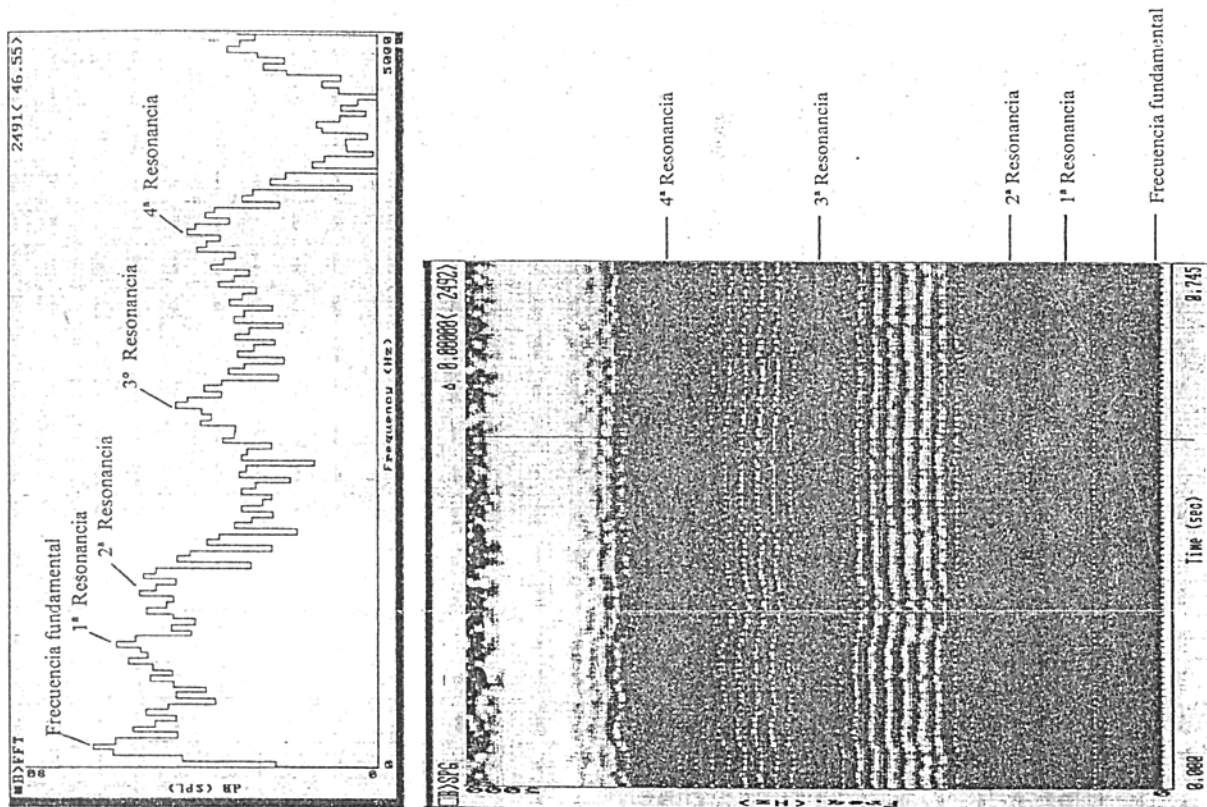


FIGURA 21. Se señalan sobre un espectrograma (arriba) y un sonograma (abajo) la frecuencia fundamental y las cuatro primeras resonancias de un sonido complejo.

pectro (frecuencias altas) se sitúa una resonancia, mayor es la sensación de brillantez y luminosidad del sonido; contrariamente: a medida que la resonancia se sitúa hacia la zona de las bajas frecuencias, mayor es la sensación de matización oscura del sonido. Sin duda, el mejor modo de ilustrar esto por escrito es recurriendo a nuestra voz, el instrumento de producción sonora más familiar del ser humano. El lector puede comprobar por sí mismo cómo es posible matizar el timbre de su propia voz sin variar el *tono* ni la *intensidad* emitiendo un sonido fuerte con una energía de espiración pulmonar que no varíe (*intensidad constante*) y sin alterar en ningún momento el grado de tensión muscular de la laringe (*tono constante*) pero cambiando la posición de la boca. A medida que los labios se cierran el sonido se oscurece. Cuando los labios se abren y la mandíbula se separa el sonido es claro. Si entrecerramos la mandíbula con los labios abiertos, y a la vez acercamos progresivamente el dorso de la lengua hacia el paladar con lentitud, escucharemos cómo nuestra voz se va haciendo estridente, mate y perdiendo profundidad.

Acústicamente, lo que está ocurriendo es que modificamos la situación de las resonancias del espectro de nuestra voz al ir cambiando la forma de la cavidad bucal. Al ir cambiando la forma y el volumen de espacio en el que se estrellan, reflejan y entremezclan las frecuencias del sonido compuesto producido por nuestra laringe, alteramos la *impresión espectral* que produce nuestra voz.

Desde el punto de vista conceptual, las resonancias producidas en la cavidad bucal y las producidas en las bóvedas del techo de una sala de conciertos no tienen ninguna diferencia entre sí. No obstante, cuando se nombran las resonancias producidas en la boca se habla sistemáticamente de *formantes*, dejándose la denominación «*resonancias*» para los instrumentos musicales y las salas de audición. Así, para que las resonancias de nuestra familiar /A/ en la figura 21 estén nombradas correctamente, en lugar de hablar de primera segunda, tercera o cuarta resonancia, debemos hablar respectivamente de *primer formante* o *F1*, *segundo formante* o *F2*, *tercer formante* o *F3*, etc. De hecho, una parte esencial del conocimiento acústico de la lengua está apoyado en la ubicación dentro del espectro de los tres primeros formantes (*F1*, *F2* y *F3*) que se observan al analizar las unidades mínimas del habla (fonemas). Estamos hablando ya de una forma muy concreta del *timbre fonético*, que es uno de los terrenos acústicos más investigados en los últimos veinte años en las principales lenguas del mundo. Especialmente en lo que se refiere a los sonidos vocálicos.

En función de como hemos definido el concepto de *impresión espectral*, el *timbre fonético* es una categoría de éste. Concretamente, denominamos *timbre fonético* al subconjunto de *impresiones espectrales* que configuran los sonidos de las lenguas. Así, la sensación de sonido /a/, por ejemplo, es una *impresión espectral* que pertenece a la categoría de los *timbres fonéticos*.

Para formalizar el *timbre fonético* de las vocales, la fonética traduce toda la complejidad del espectro acústico sólo en dos o en tres datos. Estos datos son las frecuencias concretas en las que se sitúa el punto de máxima intensidad de cada uno de los dos o los tres primeros *formantes*. Veamos esto en un ejemplo concreto.

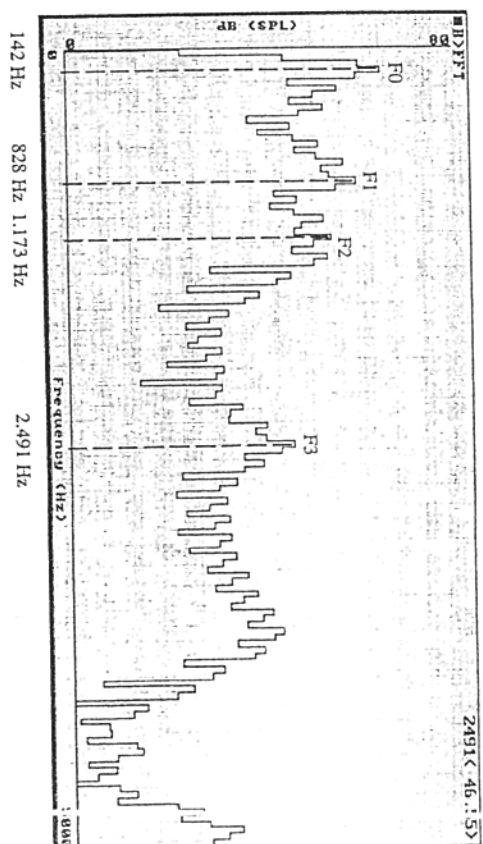


FIGURA 22. Medición en Hz de la frecuencia fundamental ( $F_0$ ) y los tres primeros formantes ( $F1$ ,  $F2$  y  $F3$ ), sobre el espectrograma de una /A/.

En la figura 22 seguimos interpretando el espectrograma de la /A/ estacionaria. Ahora hemos localizado cuál de las frecuencias que constituyen cada formante es la de mayor intensidad, y la tomamos como frecuencia representativa. Es decir, tomamos cada uno de estos valores frecuenciales como dato representativo de cada uno de los formantes del fonema /A/. Tenemos, entonces, que nuestra /A/ tiene la *frecuencia fundamental* ( $F_0$ ) a 142 Hz, su primer *formante* ( $F_1$ ) a 821 Hz, el segundo ( $F_2$ ) a 1.173 Hz y el tercero ( $F_3$ ) a 2.491 Hz.

En la sensación perceptiva de escuchar una /A/ la situación de F1 y F2 es lo que determina el *timbre fonético*, es decir lo que hace que identifiquemos («eterminado fragmento del sonido de la voz como una categoría sonora concreta) del habla denominada fonema «a». En cambio, la situación de  $F_0$  (*frecuencia fundamental*) no afecta en absoluto a la sensación de *impresión acústica*, sino a a tonal. Según esté  $F_0$  más alta o más baja percibiremos que la voz es más grave («más aguda, no obstante, mientras la situación de los formantes no varíe, seguiríamos aguda, no obstante, mientras la situación de los formantes no varíe, seguiríamos percibiendo como constante la *impresión espectral* de /A/. En este caso, el tono (142 Hz) es el de una voz masculina media, ni muy grave ni demasiado aguda. Una  $F_0$ , por ejemplo, de 80 Hz sería el de una voz masculina sumamente grave.

El tercer formante (F3) se incluye también, en ocasiones, como valor de la configuración del *timbre fonético*, no obstante no existe un acuerdo unánime al respecto entre los estudiosos. Quilis, por ejemplo, cuando define listas de *v*-lores sobre la realización acústica de las vocales del español sólo presenta datos (e los dos primeros formantes (Quilis, 1981, págs. 157-158). Es habitual encontrar comentarios en la literatura sobre fonética que asignan a los formantes altos (F3, F4, F5, etc.) un valor expresivo relacionado con las características individuales (el lo-

ductor: con la forma del paladar, la disposición de los dientes, el tamaño del velo del paladar, etc. No obstante, esto no ha sido probado y las investigaciones al respecto navegan todavía sin un rumbo fijo.

Es interesante observar lo limitado de la información acústica a la que se asigna valor expresivo: solamente tres puntos frente a la extrema complejidad del espectro. Incluso estos tres puntos ( $F_0$ ,  $F_1$  y  $F_2$ ) comportan una serie de informaciones que, normalmente, son despreciadas. Cada uno de los formantes tiene una intensidad concreta, que varía en cada una de las distintas realizaciones del fonema «a». Cada uno de los formantes, además de la frecuencia de mayor intensidad, está constituido por una franja más o menos ancha de otras frecuencias contiguas a ella; y en cada realización de una /a/ esta franja de frecuencias tiene un perfil de intensidades distinto. Posiblemente, esta información acústica sea la que permite, al escuchar un sonido de /a/ lo suficientemente largo, que el oyente pueda deducir si el locutor está deprimido, tiene una actitud agresiva o siente miedo. Pero éste es un campo muy poco explorado aún, probablemente porque los estudiosos de la narración audiovisual están todavía poco preocupados por encontrar instrumentos que puedan objetivar la eficacia expresiva de la voz.

Hemos revisado con cierto detalle la *impresión espectral* en función de las posiciones de la boca y cómo se relacionan estas sensaciones con el análisis acústico, concretamente, con los *formantes*.

Volvamos, ahora, de nuevo, al concepto de *resonancia*, para estudiar la sensación tímbrica de *impresión espectral* en las salas de audición (cines, aulas, salas de conciertos, etc.).

Volveremos a hacer referencia a la voz al ser siempre éste el más familiar y conocido de los sonidos y, por tanto, el sonido respecto al cual estamos en mejores condiciones para observar auditivamente cualquier tipo de cambio por leve que éste sea.

Probablemente el lector recordará alguna situación en la que ha oído hablar a alguien conocido, una vez que se ha alejado lo suficiente, en una sala de gran tamaño: una sala de conferencias, un aula, una catedral... Su voz se torna de repente oscura, o metálica, o se hace ininteligible. Este tipo de fenómeno tiene exactamente el mismo punto de partida que el del *tímbre* fonético: la reflexión de la voz, ahora en las paredes de la sala, y la modificación de su espectro de frecuencias por la suma del sonido original con sus propias reflexiones.

Cuanto mayor es el tamaño del resonador, más indefinida es la configuración de los formantes y, por tanto, más difusa es la *impresión espectral*. No obstante, es muy fácil diferenciar entre el *tímbre* que introduce en la voz un aula alargada y estrecha, que resulta frío, metálico y estridente, del que produce la nave central de una catedral gótica, que da a la voz un matiz mucho más cálido y oscuro. Acústicamente, estas sensaciones se traducen en el espectrograma como *resonancias* con un aspecto muy distinto que el de las producidas por nuestro «diminuto» resonador bucal. Las *resonancias* producidas por una sala son siempre mucho más anchas y de menor intensidad que los *formantes* producidos por la cavidad bucal. La influencia de los resonadores de gran tamaño actúa siempre sobre una cantidad mucho mayor de frecuencias del sonido compuesto original, y esto es lo que

genera esa impresión sonora de difusión, o dispersión, tan característica de los resonadores de gran volumen.

En la figura 23 podemos observar dos espectros de un mismo sonido, antes y después de ser emitido dentro del aula magna de la Facultad de Ciencias de la Comunicación de la Universidad Autónoma de Barcelona.

El sonido en cuestión es un ruido blanco (véase nota 9) generado digitalmente por un equipo «Computer Speech Lab» de la empresa Kay. El gráfico superior es un espectrograma del ruido blanco en su estado original, y el inferior es un espectrograma del mismo ruido blanco una vez emitido dentro de la sala y grabado durante la emisión con un micrófono situado en el centro del aula magna.

Comparando ambos perfiles se puede comprobar, en primer lugar, que el ruido blanco antes de ser emitido en la sala (gráfico superior) tiene un perfil muy regular a lo largo de todo el espectro, con variaciones de intensidad de las frecuencias que lo componen que oscilan entre 5 y 10 decibelios (entre línea y línea punteada hay un salto de 10 dB). En cambio, el mismo ruido blanco una vez que ha recorrido la sala (gráfico inferior) presenta ya variaciones en su cresta de hasta 30 decibelios. Puede observarse, también, cómo las diferencias en el gráfico inferior se acumulan en zonas concretas: la intensidad de las frecuencias aumenta globalmente en torno a los 1.030 Hz y disminuye alrededor de los 1.868 Hz.

Por otra parte, si el lector compara el perfil del espectrograma inferior de la figura 23 con el de la figura 22 podrá comprobar, como decíamos un poco más arriba, que las *resonancias* que configuran el perfil del primero son mucho más anchas y suaves que las que configuran el segundo. En el caso de la /a/ los *formantes* (o *resonancias*) trazan un perfil más acentuado y mucho mejor definido visualmente que el que han producido las resonancias del aula magna.

En suma, y recuperando todo lo revisado hasta aquí sobre el papel perceptivo de las resonancias, la organización de las *resonancias* o *formantes* a lo largo del espectro configuran una sensación tímbrica que hemos denominado genéricamente como *impresión espectral*.

- *Definición auditiva*. Para desarrollar y explicar este concepto perceptivo estableceremos un símil con la terminología utilizada para hablar de la sensación de detalle visual en la reproducción de imágenes:

Denominaremos definición auditiva a la sensación de máximo grado de precisión, exactitud o detalle sonoro que percibe el oyente al escuchar atentamente un sonido.

Esta sensación no está vinculada con la de altura tonal ni con la de intensidad, sino con la percepción de la estructura interna del sonido. Es, también, por tanto, una sensación tímbrica. La sensación de *definición auditiva* está ligada a la gama de frecuencias que componen un sonido. La riqueza de frecuencias con que está compuesto cada sonido transmite al oído humano una sensación equiparable a la de grano o textura visual. Cuanto mayor es la gama de frecuencias con que está compuesto un sonido, más grande es la sensación de detalle y de precisión sonora que produce. Así, la sensación de *definición auditiva* es tanto mayor cuando el

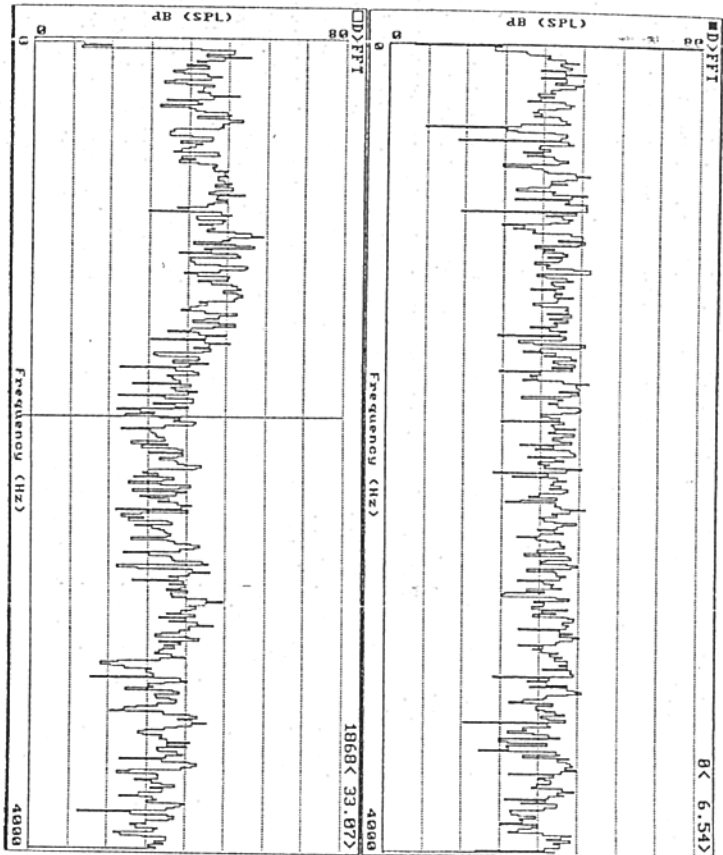


Figura 23. Espectrograma de un ruido blanco antes (arriba) y después (abajo) de ser emitido en un auditorio.

oído dispone de más elementos acústicos capaces de transmitir información sonora. Inversamente, cuanto más pequeña es la gama de frecuencias con la que se compone un sonido, menor es la sensación de precisión al escucharlo, y, por tanto, menor es el efecto de *definición auditiva*.

Con la evolución de la tecnología ha ido aumentando la gama de frecuencias que son capaces de recoger los sistemas de grabación y transmisión de sonido. Actualmente, los sistemas de alta fidelidad recogen perfectamente toda la gama de frecuencias que puede percibir el oído humano. No obstante, a menudo escuchamos sonidos con una *definición auditiva* bastante baja. El lector habrá observado que cualquier voz conocida cambia radicalmente de calidad sonora, es decir, de *timbre*, al ser escuchada a través del teléfono. La gama de frecuencias que es capaz de transmitir el teléfono está entre los 300 y los 3.000 hertzios. Teniendo en cuenta que la capacidad perceptiva del oído humano se extiende de 20 a 20.000 Hz, la *definición auditiva* del sonido telefónico es bastante escasa. Dicho de otra forma: el *timbre* característico del sonido podría ser descrito perfectamente como de *baja definición auditiva*.

Otro ejemplo característico podría ser el de la radio de onda larga. La gama de frecuencias que es capaz de transmitir este sistema va de los 160 Hz a los 6.300 Hz. Evidentemente, su *definición auditiva* es mayor que la del teléfono, pero es todavía bastante pobre. En suma, esa sensación tímbrica tan característica que nos hace identificar el sonido telefónico, o el sonido emitido por cualquiera de los equipos de audio domésticos, y diferenciarlo inmediatamente del sonido natural directo es su inferior grado de *definición auditiva*.

En el ámbito de la comunicación audiovisual se manipula muy a menudo el grado de *definición auditiva*. No obstante, en esta labor suelen predominar los criterios técnicos sobre los expresivos. De hecho, el término habitual dentro del ámbito de tratamiento del audio en las producciones audiovisuales es el de *ancho de banda*.<sup>15</sup> No es ningún secreto para un técnico de audio que el *ancho de banda* del teléfono es mucho menor que el de cualquier equipo doméstico de alta fidelidad. Pero la denominación *ancho de banda* tiene un carácter exclusivamente acústico y en ningún caso hace referencia a una dimensión psicológica perceptiva. No van a tener en absoluto el mismo papel perceptivo, por ejemplo, dos anchos de banda de 3.500 Hz, si el primero va de 100 Hz a 3.600 Hz y el segundo de 5.000 Hz a 8.500 Hz. A pesar de que ambos tienen un *ancho de banda* idéntico la sensación de *definición auditiva* del primero no tendrá absolutamente nada que ver con la del segundo. Es por eso que hemos considerado pertinente acuñar el concepto de *definición auditiva*. Podemos decir que existe una relación directamente proporcional entre el concepto acústico de *ancho de banda* y el perceptivo de *definición auditiva*. Ciertamente, cuanto mayor sea el *ancho de banda*, mayor será el grado de *definición auditiva*, y viceversa. No obstante, esta relación no es isomórfica, ya que cuanto más alta es la zona de la banda de frecuencias escuchada, menor es la capacidad de discriminación perceptiva del oído humano.

Antes de seguir adelante con esta reflexión sobre la *definición auditiva*, primero será útil para el lector que hagamos una descripción elemental sobre los mecanismos básicos de la técnica del filtrado, puesto que en la industria audiovisual esta es la forma más habitual de manipulación del *ancho de banda* de los sonidos, y, por tanto, también de su *definición auditiva*.

La forma más simple de filtrar es controlando la sensibilidad del sistema en la recogida del sonido. Un sistema, por ejemplo un micrófono cuya membrana sólo puede vibrar entre 0 y 5.000 Hz, al ser estimulado con un sonido compuesto por frecuencias de 0 a 20.000 Hz, traducirá la señal sonora original ignorando todas las frecuencias que no estén dentro de esta gama. Es decir, no recogerá ninguna de las frecuencias que están situadas entre 5.000 y 20.000 hertzios. Del mismo modo, un sistema que no pueda vibrar a frecuencias bajas, por ejemplo entre 0 y 300 Hz, al ser estimulado por un sonido compuesto por frecuencias entre 0 y 20.000 Hz, ignorará las frecuencias bajas y actuará como un filtro, recogiendo el sonido sólo desde los 300 Hz hacia arriba. En definitiva, el filtrado consiste en realizar una trans-

15. Se denomina *ancho de banda* a la diferencia entre los límites de frecuencia superior e inferior de un componente de audio. Por ejemplo, si la gama de frecuencias que es capaz de transmitir el teléfono va de 160 Hz a 3.600 Hz, el ancho de banda del sonido telefónico será de: 3.600 - 160 = 3.440 Hz.



misión de la señal sonora de un lugar a otro utilizando un sistema de captación que sólo es sensible para una determinada banda de frecuencias. Naturalmente, esta gama de frecuencias ha de ser selectiva y controlable por un operador. Las denominaciones más clásicas de filtrado son:

- De paso bajo (sólo deja pasar frecuencias graves).
- De paso alto (sólo deja pasar frecuencias agudas).
- De paso de banda (sólo deja pasar algunas frecuencias centrales).

Recuperemos ahora, de nuevo, la idea de que la manipulación del grado de definición auditiva es algo habitual en el ámbito audiovisual.

En la posproducción de audio, se recurre muy frecuentemente al recurso del filtrado para quitar la «suciedad» sonora. Es decir, para borrar sonidos no deseados cuando éstos están concentrados en una zona bien definida del espectro. Hemos visto ya que filtrar un sonido no es más que la eliminación selectiva de una gama determinada de frecuencias del espectro. Si, por ejemplo, detectamos mediante un análisis espectral que los ruidos no deseados que se escuchan en una grabación están en la zona más baja del espectro, por ejemplo en torno a los 50 Hz (el umbrido característico de las lámparas fluorescentes), podremos filtrar el sonido borrando la banda de frecuencias de 0 a 60 Hz. Este tratamiento de la señal sonora, efectivamente, eliminará el ruido, pero también cualquier otra frecuencia significativamente expresiva que se encuentre en esa misma zona del espectro.

En suma, todo filtrado, sea del tipo que sea, altera el *timbre* del sonido. Más concretamente: hace bajar el grado de *definición auditiva*, del sonido tratado.

Desde el punto de vista expresivo, los filtrados se utilizan a menudo en la realización audiovisual para simular sonido telefónico, o sonido radiofónico a partir de sonidos que han sido grabados con equipos de alta fidelidad y, por tanto, que tienen un *ancho de banda* entre 15.000 o 20.000 Hz. Evidentemente, lo que se está haciendo en estos casos es reconstruir el *timbre* del sonido telefónico o el del sonido radiofónico reduciendo la *definición auditiva* de una grabación original de gran definición.

Para concluir ya este apartado sobre la sensación tímbrica retomaremos la idea del *timbre* como sensación global compleja.

Hasta ahora, hemos dimensionado el timbre en tres aspectos diferenciados: la *armonicidad*, la *impresión espectral* y la *definición auditiva*. No obstante, las tres dimensiones sensoriales hacen referencia y dependen de la estructura interna del sonido, de su composición espectral y, por tanto responden a la concepción global de *timbre*. Así, podemos decir perfectamente que el *timbre* es una sensación compleja en la que influyen de forma simultánea y unívoca armonicidad, impresión espectral y definición auditiva, es decir:

$$\text{timbre} = (\text{armonicidad}) + (\text{impresión espectral}) + (\text{definición auditiva})$$

Para cerrar con cierta solidez esa concepción global del *timbre* que proponemos, el papel perceptivo de la composición interna del sonido también debería ser estu-

diado con detalle en su dimensión dinámica; es decir, en su actuación sobre las sensaciones auditivas en función del tiempo. Pero, desgraciadamente, el nivel de conocimiento acumulado en este campo todavía no ha generado las condiciones para eso.

### 3.3.4. Sobre la dimensión dinámica del timbre

Consideramos que en esta obra se aportan algunos elementos que pueden ser útiles para avanzar un poco más en el conocimiento y la comprensión de los sonidos compuestos, especialmente en la dimensión tímbrica. No obstante, nuestro planteamiento sobre el *timbre* se presenta pretendidamente como algo abierto en tanto que aún no podemos tratar su dimensión dinámica con la suficiente profundidad. En cualquier caso, sí que tenemos la certeza de que éste es un camino que necesariamente se deberá recorrer.

Más arriba, cuando estudiábamos el *sonograma* (apartado 3.2.3) habíamos visto que la composición de frecuencias de un sonido es un fenómeno dinámico que evoluciona constantemente en el tiempo. Y este fenómeno de evolución dinámica se da también, aunque en un grado mucho menor, en los sonidos estacionarios. El lector estará ya en condiciones de observar el sonograma de cualquier sonido estacionario y notar que las franjas que definen su composición de frecuencias tienen ligeras variaciones tanto de altura (frecuencia) como de grosor u oscuridad (intensidad). Puede comprobar todo esto, por ejemplo, en la figura 12, o en el centro de la /E/ de la figura 17. Y ver, también, que estas variaciones tienen una cierta regularidad cíclica a medida que el sonido evoluciona en el tiempo. Estas pequeñas variaciones regulares en la evolución dinámica del espectro actúan sobre la sensación tímbrica. Pero el carácter fugaz y múltiple de este tipo de variaciones de la estructura interna del sonido hace sumamente difícil y complicado su estudio.

Risset, intentando sintetizar digitalmente un timbre de trompeta, descubrió en 1965 que la variación dinámica del espectro de frecuencias era fundamental para que el sonido que estaba construyendo se pareciese perceptivamente al de una trompeta real. Y, en consecuencia, llegó a la conclusión de que la reconstrucción sintética de un sonido a partir de modelos espectrales completamente estacionarios no era válida (véase Mathews y Pierce, 1987, págs. 82-90). Quien escribe estas líneas piensa como Risset, como Mathews y como Pierce, que la evolución dinámica del espectro es una pieza clave para comprender cómo actúa la percepción sonora humana de los sonidos compuestos. Pero éste es todavía un campo demasiado poco conocido. Quizá, el campo más difícil e inexplorado del complejo y fascinante universo de las sensaciones sonoras.

### 3.4. Las magnitudes físicas para la medición del sonido

La acústica aporta dos unidades esenciales para poder medir y comparar entre sí los sonidos. Estas unidades son el *hertzio* para la frecuencia y el *decibelio* para



la amplitud. Ambas unidades están siendo utilizadas de un modo muy generalizado en todos los ámbitos en los que se trabaja con el sonido. Se han convertido en el punto de partida para cualquier investigación en la que las vibraciones audibles forman parte del objeto de estudio, y son también unidades muy cotidianas en el ámbito de la industria audiovisual. No obstante, es importante tener muy presente que aunque estas dos magnitudes dan cuenta adecuadamente de las dimensiones estrictamente físicas de los fenómenos vibratorios, están muy lejos de resolver satisfactoriamente la medición de las sensaciones de *tono e intensidad*.

En el primer capítulo de esta obra concretábamos ya que nuestro objeto de estudio es el *lenguaje audiovisual*, es decir: *los modos artificiales de organización de la imagen y el sonido que utilizamos para transmitir ideas o sensaciones*. Esto supone que el sonido no nos preocupa en tanto que entidad física, sino como instrumento para la transmisión voluntaria de sensaciones y de ideas. Así, lo esencial para nosotros no es la dimensión física del sonido sino su dimensión perceptiva. Consecuentemente, es importante que revisemos estos instrumentos de medida a la luz de nuestras necesidades concretas, mucho más cercanas a las de la psicología perceptiva que a las de la física.

### 3.4.1. La medición del tono

Técnicamente, la medición física de la frecuencia suele utilizarse en todos los ámbitos como medida de la sensación tonal, a pesar de que ambos son fenómenos perfectamente diferenciados.

La medida de la frecuencia es sumamente simple. Se mide contando el número de oscilaciones por segundo que desarrolla cualquier objeto al sufrir una vibración. La frecuencia de vibración de un objeto que actúa como fuente sonora se responderá con el número de oscilaciones que sufren las moléculas del aire que son estimuladas por él. Y el aire transmite esta frecuencia al oído, que se encargará de traducirla en sensación tonal; o al diafragma del micrófono, que se encargará de traducir la vibración sonora en variaciones de amplitud eléctrica.

La física toma como magnitud de referencia para medir cualquier frecuencia una vibración que tarda un segundo en hacer el recorrido completo desde que parte de su punto de reposo en un sentido y retorna a este punto en el mismo sentido. A esta relación entre la rapidez de la vibración y el tiempo se la denomina: *1 ciclo por segundo (1 cps)*. Haciendo referencia siempre a esta magnitud es posible comparar con exactitud la rapidez de cualquier vibración con la de cualquier otra. El *ciclo por segundo* es, pues, la magnitud que se utiliza como unidad para el estudio del grado de rapidez de las vibraciones. A esta unidad se la denomina también *hertz (Hz)*. Éste es el nombre de su creador.

$$1 \text{ hertz} = 1 \text{ ciclo} \times 1 \text{ segundo.}$$

Así, del sonido producido por unas moléculas de aire que oscilan en 150 oscilaciones cada segundo se dice que tiene una frecuencia de 150 Hz.

Esta unidad no responde al umbral de audición humana, sino que éste es de unos 17 cps. Una vibración de 17 cps produce una sensación de intensidad que sólo se ajusta parcialmente a la dinámica de la audición. En la audición contemplamos la pérdida de finura en la sensibilidad que tienen todos los seres humanos a medida que va aumentando la intensidad del estímulo percibido (ley de Weber y Fechner).<sup>16</sup>

Ciertamente, existe una relación entre la frecuencia de una vibración medida en Hz. y la sensación tonal que ésta produce, ya que cuando aumenta la frecuencia de la vibración sube, también, la sensación tonal. Pero esta relación no es lineal sino geométrica: cada vez que se dobla la frecuencia sólo aumenta en un grado nuestra sensación auditiva de tono. Estudiaremos con más detalle estas cuestiones en el capítulo 4.

El sistema musical de Occidente ha organizado sus unidades de una forma mucho más ajustada a la sensibilidad del oído humano que la física acústica. La escala tonal estructura su gama de unidades en *semitonos, tonos y octavas* y de este modo se aproxima bastante más a la estructura de la percepción humana que el estudio en cps o Hz. A continuación desarrollaremos un cuadro en el que se establece la equivalencia sistemática entre medida física de las vibraciones y escala tonal. En el cuadro se muestra el valor en Hz de todas las notas en 5 octavas de la escala tonal, y el salto de frecuencia que se produce entre cada nota musical y la siguiente.

En la primera columna, las cifras ordenadas de arriba abajo en negrita y que son de mayor tamaño muestran la frecuencia que se corresponde, respectivamente, con cada una de las siete notas de la primera octava (DO de la 1ª octava = 64 Hz; RE de la 1ª octava = 73 Hz; MI de la 1ª octava = 82 Hz; etc.). Las cantidades que están en cursiva y son de tamaño menor expresan el salto en Hz que hay entre cada nota y la inmediata siguiente. En la segunda columna se muestran, también de arriba abajo las respectivas frecuencias en Hz que corresponden a cada una de las notas de la 2ª octava y los saltos de frecuencia entre notas. En la tercera columna los de la 3ª octava; y así sucesivamente hasta la 5ª octava.

Mirando este cuadro, el lector puede observar cómo las notas de la escala musical, al ser escuchadas, producen la sensación de ser lineales y equidistantes entre sí; cuando son medidas físicamente, las cifras en Hz que las representan reflejan que la distancia entre ellas no es regular sino que aumenta muy rápidamente a medida que subimos en la escala. Mientras el salto de frecuencia entre las notas de la primera octava está en torno a los 10 Hz, en la quinta octava los saltos de frecuencia varían ya entre 150 y 200 Hz. Probablemente, lo que resulta perceptivamente más claro es la diferencia entre las variaciones de frecuencia y de sensación tonal de una octava a otra. Cuando escuchamos sucesivamente los DO de la 1ª, 2ª, 3ª, 4ª y 5ª octavas de un piano tenemos una sensación auditiva muy clara de que el salto tonal es regular y de que cada una de las notas es auditivamente equidistante de la anterior. Sin embargo, los datos de sus respectivas frecuencias nos

16. Esta ley se revisa con detalle en el apartado 3.4.2.

Frecuencias y saltos de frecuencia en la escala tonal

|     | 1ª  | 2ª  | 3ª  | 4ª  | 5ª   | Octavas | Frecuencia de cada tono en Hz.         |
|-----|-----|-----|-----|-----|------|---------|----------------------------------------|
| DO  | 64  | 128 | 256 | 512 | 1024 |         | Salto de frecuencia entre tonos en Hz. |
|     | 9   | 18  | 37  | 75  | 150  |         |                                        |
| RE  | 73  | 146 | 293 | 587 | 1174 |         |                                        |
|     | 9   | 18  | 36  | 72  | 144  |         |                                        |
| MI  | 82  | 164 | 329 | 659 | 1318 |         |                                        |
|     | 5   | 10  | 20  | 39  | 78   |         |                                        |
| FA  | 87  | 174 | 349 | 698 | 1396 |         |                                        |
|     | 11  | 22  | 43  | 85  | 171  |         |                                        |
| SOL | 98  | 196 | 392 | 783 | 1567 |         |                                        |
|     | 12  | 24  | 43  | 97  | 193  |         |                                        |
| LA  | 110 | 220 | 435 | 880 | 1760 |         |                                        |
|     | 13  | 26  | 58  | 107 | 213  |         |                                        |
| SI  | 123 | 246 | 493 | 987 | 1973 |         |                                        |

muestran cómo el número de hertz se duplica cada vez que pasamos del DO de una octava al de la siguiente. Si esta variación tuviese una correspondencia numérica directa en la sensación tonal, deberíamos percibir que la distancia tonal entre los DO sucesivos es cada vez mayor a medida que ascendemos en la escala. Y esto no es así.

Finalmente, la magnitud acústica para medir la frecuencia, por sí sola, tampoco da cuenta de un modo satisfactorio de los sonidos compuestos. De hecho, el concepto de *frecuencia* en su sentido estricto sólo es satisfactorio para los sonidos simples. En el momento en que nos enfrentamos a un sonido compuesto, como hemos visto ya en el apartado 3.3.1, es necesario diferenciar entre *frecuencia fundamental* y *armónicos o parciales*. Cuando se habla de *frecuencia* de un sonido compuesto, normalmente, se está haciendo referencia a su *frecuencia fundamental*, al ser ésta la que tiene una influencia perceptiva más clara en la sensación tonal. No obstante, hemos visto que el resto de las frecuencias que componen los sonidos complejos influyen también de una forma determinante en la sensación auditiva (*timbre*) que éstos producen. Así, el concepto de *frecuencia*, por sí solo, resulta demasiado simple para cuantificar adecuadamente este fenómeno perceptivo.

3.4.2. La medición de la intensidad

Mientras que la *frecuencia* hacía referencia exclusivamente al fenómeno vibratorio, y por tanto físico, el concepto de *intensidad* hace referencia ya de forma directa a la sensación psicológica de energía del sonido. Veremos que la acústica ha resuelto bastante mejor la medida de la amplitud de las vibraciones sonoras

que la medida de su frecuencia, al introducir en la medición de la energía sonora la pérdida progresiva de finura del sistema sensorial humano con el aumento de la energía de los estímulos. No obstante, encontraremos también algunos problemas en la cuantificación de la intensidad.

La definición de la unidad de medida de la intensidad es bastante más compleja que la de la frecuencia y hemos de partir de algunas consideraciones previas sobre la percepción antes de llegar a ella.

Como hemos dicho ya, el ser humano pierde finura en la sensibilidad de todos los sentidos a medida que aumenta físicamente la intensidad del estímulo percibido. Si comparamos, por ejemplo, el peso de dos paquetes pequeños de 100 g y 150 g respectivamente, poniendo uno en cada mano somos perfectamente capaces de percibir la diferencia y decir cuál de los dos es el más pesado. Es decir afirmamos a distinguir 50 g de diferencia. Pero si esta comparación la hacemos con paquetes de 10.000 g y 10.050 g utilizando la misma técnica seremos incapaces de notar diferencia alguna. Para percibir la misma sensación diferenciadora que tuvimos entre 100 y 150 gramos deberíamos tener en una mano 10 kg y en la otra 15 kg.

Es decir, para sentir lo mismo que antes podíamos percibir con 50 gramos, al aumentar la intensidad del estímulo necesitamos una diferencia de peso 100 veces mayor. Exactamente la misma proporción con la que ha aumentado globalmente el estímulo a comparar.

Este fenómeno fue sistematizado en la ley de Weber y Fechner, que se formula de la forma siguiente: *La percepción es proporcional al logaritmo de la excitación*. O lo que es lo mismo: *La sensación crece sólo en progresión aritmética mientras que la excitación que la provoca crece proporcionalmente en progresión geométrica* (véase Bruce Goldstein, 1993, pág. 20).

Volviendo al ejemplo de la comparación de 100 gramos y 150 gramos de peso entre las dos manos, si la percepción es proporcional al logaritmo de la excitación, calculando el logaritmo de los pesos tendremos una idea de la sensación que éstos transmiten. Veámoslo:

$$\log 100 = 2 \quad \log 150 = 2,17$$

así, la sensación de diferencia percibida es de

$$2,17 - 2 = 0,17.$$

Si lo que comparamos son 10 kg con 10 kg más 50 g

$$\log 10.000 = 4 \quad \log 10.050 = 4,002$$

la diferencia percibida es de

$$10.050 - 10.000 = 0,002$$

para conseguir la misma diferencia de sensación (0,17), el peso en la mano que aguanta una carga mayor debería ser igual a la inversa del logaritmo de  $(4 + 0,17)$ . O sea, el peso debería ser:  $10^{+17}$ . Haciendo el cálculo:

$$10^{+17} = 14.791 \text{ g (unos 15 kg).}$$

Estudiaremos ahora la unidad clásica de intensidad sonora que es el *decibel* (dB). Esta magnitud intenta relacionar la percepción humana con la cuantificación física de la presión que producen las vibraciones sonoras del aire al incidir sobre el oído. Es una medida relativa que toma como referencia la mínima presión sonora que es capaz de percibir el oído humano. Y compara la presión ejercida por cualquier sonido con esa magnitud.

A partir de aquí desarrollaremos paso a paso todo el proceso de construcción conceptual y matemática del *decibel* o *decibelio*.

Esta unidad trabaja desde el sistema CcGsInal (centímetros, gramos, segundos) y toma como referencia una frecuencia pura de 1.000 cps. La razón por la que se toma una referencia de 1.000 Hz es que la percepción de la intensidad no solamente varía con la amplitud de la vibración, sino que varía, también, en función de la frecuencia. Escuchando una frecuencia de 1.000 Hz estamos en la zona frecuencial para la que el oído humano tiene una respuesta sensible más regular a la amplitud de las vibraciones sonoras.

El punto de partida es el fenómeno físico: concretamente, la presión que ejercen las moléculas de aire, que han sido estimuladas por una fuente sonora, al vibrar sobre el tímpano humano. Aproximadamente,<sup>17</sup> la presión mínima que puede percibir el oído es de:  $0,0002 \text{ dinas/cm}^2$ . Esta cantidad se tomó como magnitud de referencia ( $P_0$ ) con la que se compararía cualquier otra presión acústica ( $P_1$ ) sobre el oído. Es decir, como unidad. Así, se podría conocer el grado de presión auditiva de cualquier sonido calculando el número de  $P_0$  que contuviera. La presión auditiva sería pues:

$$\text{presión auditiva} = P_1 / P_0$$

A partir de aquí era necesario encontrar el otro extremo: la máxima presión sonora perceptible, y construir una escala entre estos dos límites de la sensibilidad humana a la intensidad.

La máxima presión sonora que el oído podía aguantar antes de llegar al umbral de dolor era 1.000.000.000.000 de veces la presión mínima. Es decir  $10^{12} \times (0,0002 \text{ dinas/cm}^2)$ . Con lo que desde el umbral mínimo hasta el umbral máximo aparecía una escala de un billón de grados, que resultaba absolutamente innegable. No obstante, trabajando matemáticamente con potencias de 10 se podía

17. La sensibilidad mínima del oído a la presión auditiva varía de unas personas a otras; no obstante, se puede establecer sin problemas el umbral mínimo medio dentro de unos márgenes de variación relativamente estrechos. El lector puede encontrar en la obra *Introducción a la audíometría* (Quirós y Delia, 1982) un amplio y profundo estudio sobre estos temas.

diseñar una escala sólo de 12 grados desde la sensibilidad mínima hasta la máxima. O sea:

$$\text{si } 1.000.000.000.000 = 10^{12}$$

los grados podían ser:

$$10^1, 10^2, 10^3, 10^4, 10^5, \dots, 10^{12}.$$

Además, se había observado ya que la sensación que percibe el ser humano es siempre proporcional al logaritmo de la cantidad de excitación; por lo tanto, el cálculo de la sensación de sonoridad, o de potencia sonora, debería ajustarse también a esta ley. Así, en realidad, la sensación de *intensidad* se ajustaría al logaritmo del cociente entre la presión sonora que incide sobre el tímpano ( $P_1$ ), dividida por la presión sonora mínima que puede percibir el oído ( $P_0$ ). Es decir:

$$\text{intensidad} = \log (P_1 / P_0).$$

Dando al umbral mínimo de presión auditiva el valor de unidad: 1, y hacer este cálculo tomando como  $P_1$  la máxima de presión audible, se obtuvo la cifra: 12 como valor que expresaba la máxima *intensidad* audible. Veámoslo:

$$\text{intensidad} = \log (10^{12} / 1) = 12.$$

Quedaba, pues, definido así el máximo de una escala relativa de 12 grados que se aproximaba a la sensación de intensidad. A la unidad de esta escala se la denominó *bel*. La máxima intensidad que soporta el oído es, pues, de 12 *bels*. No obstante, el *bel* no era adecuado como unidad para medir las pequeñas variaciones de intensidad; en realidad el *bel* resultaba una unidad demasiado grande, demasiado gruesa para la sensibilidad auditiva humana. Se decidió entonces dividir el *bel* en 10 unidades más pequeñas, con lo que se obtenía definitivamente el *decibel*. Tenemos pues que:

$$1 \text{ bel} = 10 \text{ decibels.}$$

Lógicamente, para calcular la sensación de intensidad en *decibels* (dB), será necesario multiplicar por 10 el número de *bels*. El umbral de dolor, por ejemplo, es de 12 *bels*  $\times 10 = 120 \text{ decibels}$  o *decibelios*. Ahora, la sensación de intensidad que produce un sonido se podía ya predecir de forma bastante aproximada calculando el número de *decibelios* que tiene el sonido en cuestión mediante la siguiente fórmula:

$$\text{intensidad (en decibels)} = 10 \log (P_1 / P_0).$$

Siendo:  $P_0$  la magnitud o unidad de referencia (en este caso 0,0002 dinas/cm<sup>2</sup>), y  $P_1$  la presión que ejerce el sonido en cuestión expresada en las mismas unidades que la unidad de referencia.

Finalmente, cuando la impedancia acústica es constante, o dicho de otro modo, cuando la capacidad de penetración del sonido en el aire es constante, tal como ocurre en las investigaciones acústicas, las potencias acústicas resultan proporcionales al cuadrado de las presiones acústicas. Aplicando esto, el cálculo de la *intensidad* en dB queda del siguiente modo:

$$\text{intensidad (en dB)} = 10 \log P_1^2 / P_0^2$$

o lo que es lo mismo

$$\text{intensidad (en dB)} = 20 \log P_1 / P_0$$

El *decibelio* ha sido homologado internacionalmente como unidad de medida, dándole un valor concreto de presión, exactamente el de 0,000204 dinas/cm<sup>2</sup>. Éste es el valor medio de la mínima presión audible, determinado a 1000 cps.

Al estudiar el *decibelio*, hemos visto que todo el desarrollo de esta unidad está hecho tomando como referencia una vibración simple de 1.000 Hz. Lógicamente, eso supone, también, que aunque obtengamos, por ejemplo, dos medidas idénticas de 50 dB de presión sonora en sendos sonidos de 100 y de 1.000 Hz, no tenemos ninguna garantía de que estos dos sonidos vayan a transmitir la misma sensación psicológica de intensidad. De hecho esto es justamente lo que ocurre. La sensibilidad humana a la presión auditiva es menor cuando el sonido escuchado es de frecuencia más baja, y va aumentando a medida que aumenta la frecuencia del sonido. Esto ocurre aproximadamente hasta los 3.000 Hz. A partir de esa frecuencia, la sensibilidad del oído a la presión sonora vuelve a disminuir progresivamente, hasta que desaparece en torno a los 15.000 o 20.000 Hz. Sucede, entonces, que un sonido que tenga una frecuencia de 100 Hz y una presión sonora de 20 dB, no puede ser percibido por el oído humano, mientras que si esta misma presión sonora de 20 dB llega a nuestro oído asociada a un sonido de 1.000 Hz, sí que será perfectamente audible.

Este fenómeno fue formalizado por Fletcher (véase Pierce, 1985, pags. 110-111) mediante un diagrama de curvas de igual audibilidad. La base de esta formulación gráfica es psicoacústica. Para construirlo se solicita a un amplio número de sujetos experimentales que comparen entre sí sonidos de distinta frecuencia a medida que se modula su potencia en *decibelios*, hasta que tengan la sensación de intensidad idéntica. Una vez igualadas las sensaciones de intensidad entre sonidos de distinta potencia se anotan exactamente los dB de presión sonora con los que se estaba emitiendo cada uno de los sonidos. El resultado de este trabajo es la construcción del diagrama de la figura 24.

En el diagrama de Fletcher se ponen en relación sobre una escala logarítmica la presión sonora en *decibels* (eje vertical) con la frecuencia en hertz (eje horizontal). Y sobre él se trazan curvas isofónicas. Las curvas isofónicas expresan cuál

es la relación entre potencia sonora en dB y frecuencia en Hz para que el oído humano tenga la misma sensación de intensidad a medida que va variando el tono.

Revisemos, por ejemplo, la primera de las curvas observando el diagrama de abajo hacia arriba. Esta curva indica la mínima intensidad perceptible por el oído humano. Es decir, el sonido más débil posible que empieza a ser audible. En el diagrama podemos ver que cuando la frecuencia de un sonido es de unos 35 Hz, para que empiece a ser percibido es necesario que se emita con una potencia de unos 60 dB; sin embargo, cuando el sonido es de 100 Hz, para que sea audible ya sólo es necesaria una potencia de unos 40 dB; al subir la frecuencia del sonido a 500 Hz, con que la presión sonora sea de unos 5 dB basta para comenzar a percibir el sonido; a 1.000 Hz el sonido se hace audible a 0 dB; sobre los 3.000 Hz para percibir el sonido la presión sonora necesaria es mínima, aproximadamente, -15 dB, etc.

La segunda curva de abajo arriba expresa el mismo tipo de información, pero ahora tomando como punto de referencia la sensación psicológica de intensidad que produce un sonido de 1.000 Hz emitido con una potencia sonora de 10 dB. La curva indica las distintas potencias en dB que son necesarias en cada frecuencia para que la sensación subjetiva de *intensidad* sea siempre la misma.

Todas las curvas de abajo arriba siguen siempre el mismo criterio y expresan el mismo tipo de información: la tercera curva indica la relación potencia-frecuencia necesaria para que la sensación de intensidad sea la misma que la que produce un sonido de 20 dB a 1.000 Hz; la cuarta curva, la relación potencia-frecuencia necesaria para que la intensidad sea idéntica a la de un sonido 30 dB a 1.000 Hz, etc. Finalmente, la última curva indica el umbral de dolor.

En suma, la sensibilidad del oído humano se adapta a la medición de decibelios a los 1.000 Hz, pero cuando esta frecuencia varía, la percepción humana no se ajusta ya a esta forma de cuantificar la intensidad. La conclusión vuelve a ser la misma que en el caso de la frecuencia. Aunque bastante más adaptada a la percepción humana, el *decibel* no deja de ser una magnitud centrada en el fenómeno físico de la presión y no en el fenómeno psicológico de la *intensidad*.

Para resolver este problema se estableció el *fon*, que es una medida que da el mismo valor a todos los puntos de la curva isofónica. El valor en *fon*es de cada curva es el que se obtiene en decibelios cuando la curva cruza la línea de los 1.000 Hz. Para saber cuál es la intensidad en *fon*es de un sonido hay que situarlo, pues, como un punto en el diagrama de Fletcher y buscar la curva isofónica que pasa sobre él. Así, por ejemplo, un sonido de 50 Hz y 60 dB de potencia sonora tendría un valor de 10 *fon*es, otro de 100 Hz y 40 dB equivaldría a 0 *fon*es, etc.

Pero esta unidad de medida tampoco resolvía las diferencias sensitivas. El aumento de un *fon* no se ajusta a los saltos naturales de la sensibilidad auditiva. Se inventó entonces otra medida que sí respondería a los saltos de intensidad que reconoce el hombre como regulares: el *son*. En la frecuencia de 1.000 Hz se establece la equivalencia de:

$$40 \text{ fon} = 1 \text{ son}$$

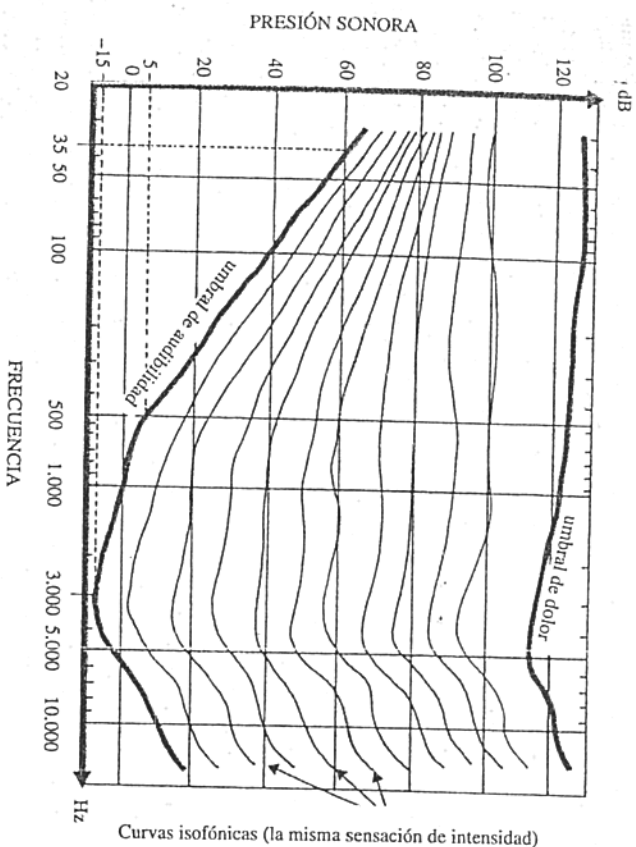


Figura 24. Diagrama de Fletcher.

A partir de esta equivalencia, cuando el oído perciba un sonido el doble de fuerte diremos que la intensidad es de 2 *sones*, si la diferencia es triple serán 3 *sones* y si es de la mitad serán 0,5 *sones*.

Finalmente, trabajar a nivel técnico con unidades como el *fon* y el *son* dificulta y encarece los instrumentos de medida para el cálculo de la intensidad. Esto ha hecho que el *decibel* se haya implantado internacionalmente como la unidad de medida utilizada en todos los ámbitos. Nos encontramos, pues, con que los instrumentos convencionales miden siempre la presión sonora en *decibelios*, y, lo que es peor, que demasiado a menudo se establece una asociación rígida entre la medición de la presión sonora en dB y la sensación psicológica de *intensidad*, cuando, cómo hemos visto, esta vinculación supone arrastrar errores muy importantes.

Todo creador de narraciones audiovisuales debería tener muy en cuenta este problema, y aprender a desconfiar de los omnipresentes vúmetros. Es imprescindible tener absolutamente claro que las indicaciones en dB de los instrumentos de medida en los estudios de grabación sonora sólo son válidas para las molestias calibradas a 1.000 Hz y para indicar los niveles de saturación desde un punto de vista exclusivamente técnico. Pero que nos aportan muy poco cuando se trata de decidir cuáles son las intensidades adecuadas para equilibrar una composición de sonidos distintos, complejos y diversos. Es el oído humano y no el vúmetro el que debe de-

cidir sobre el nivel de intensidad de los distintos sonidos para diseñar una perspectiva sonora. Es la sensibilidad auditiva del narrador y no el número de dB que n puede y debe precisar con exactitud si un determinado nivel de intensidad es o no suficiente para que el público de una sala de cine perciba el sonido de unos pasos en la lejanía.

A modo de resumen global, creo que es importante insistir en que la comprensión científica de los sonidos compuestos complejos es un problema que dista todavía mucho de estar completamente resuelto. Es cierto que los modernos instrumentos que hemos expuesto a lo largo de todo este capítulo han supuesto un enorme avance sobre el conocimiento del sonido en los últimos cincuenta años. Y que en poco tiempo se han desarrollado una serie de herramientas conceptuales, que permiten ya navegar con una cierta eficacia en el océano de los sonidos complejos.

Técnicamente, se ha avanzado y obtenido resultados muy satisfactorios sobre los sonidos estacionarios, concretamente en filtrado, eliminación y cancelación de ruidos; y sobre la manipulación y el tratamiento en general de la señal de audio. Pero hay en estos momentos numerosos frentes de investigación en los que se están invirtiendo grandes esfuerzos para conseguir sólo avances muy reducidos, especialmente en el ámbito de la comprensión del funcionamiento acústico expresivo de la voz humana.

La investigación en fonética experimental justo comienza ahora a trabajar sobre la evolución dinámica del habla en el tiempo, a pesar de ser ésta una de sus características esenciales; los estudios sobre seguridad e investigación policial no han resuelto aun el modo de identificar una voz con una garantía de acierto satisfactoria; el reconocimiento automático del habla no ha conseguido pasar todavía a de las palabras aisladas dichas por un único locutor; la diagnosis psiquiátrica y foniátrica mediante el análisis del sonido de la voz es todavía muy reducida y se utiliza sólo en un número reducido de casos; y el conocimiento general sobre los mecanismos de interpretación del sonido que utiliza el ser humano son todavía muy escasos. Por desgracia, se ha investigado sobre ellos muchísimo menos que sobre los de la comprensión visual.

Consecuentemente, lo que hemos estudiado en este capítulo no son en absoluto herramientas definitivas para resolver todos los problemas de comprensión del audio. Se trata simplemente de que tanto los investigadores como los profesionales de la comunicación audiovisual puedan tener en sus manos aquellos instrumentos de análisis acústico que, a nuestro modo de ver, se han perfilado como los más eficaces tanto para la comprensión científica, como para la manipulación técnica y expresiva de los sonidos compuestos.

18. La *cancelación* es un sistema de eliminación de ruido consistente en oponer al conjunto de vibraciones sonoras que lo componen, otro exactamente igual pero de fase invertida, de modo que al sumarse linealmente ambas oscilaciones complejas el resultado final sea cero. Es decir, ausencia total de vibración y, por tanto, desaparición del ruido. Algo así como empujar una puerta entrecerrada exactamente con la misma fuerza por ambos lados a la vez. Lógicamente, la puerta no llegará a moverse.



## CAPÍTULO 4

### FUNDAMENTOS DE LA PERCEPCIÓN SONORA

*En este capítulo se revisan aquellos mecanismos perceptivos de la audición que el autor considera fundamentales para comprender cómo se organiza la expresión sonora. Y su objetivo es localizar relaciones que existen entre los procesos perceptivos y la capacidad humana para expresar y comprender mensajes sonoros.*

#### 4.1. La subjetividad perceptiva como objeto científico

En el universo de la narración y del lenguaje audiovisual, los fenómenos que están relacionados con el universo interior del receptor suelen calificarse denominándose a menudo como algo difuso, individual e informalizable. Existe, además, una fuerte tendencia a desestimar los estudios sobre la interpretación del discurso audiovisual que ponen su punto de vista en el receptor. A menudo, aquellos que estudian la comunicación de masas desde una perspectiva sociológica, argumentan que investigar en el campo de las relaciones entre los fenómenos físicos y la interpretación de los mensajes audiovisuales es algo «demasiado subjetivo». En este tipo de discurso, el concepto de *subjetividad* suele ser utilizado como un instrumento de descalificación.

Buena parte de la literatura sobre comunicación de masas trabaja desde una concepción demasiado reduccionista del concepto *subjetividad*, entendiendo que cada sujeto percibe e interpreta de un modo completamente individualizado y diferente de los otros sujetos el conjunto de fenómenos físicos objetivos que constituyen un mensaje audiovisual.

No obstante, la psicología perceptiva ha demostrado de forma experimental, desde hace ya más de medio siglo, que cuando no se dan unas condiciones esenciales como: ambigüedad, confusión, tensión, agotamiento, miedo, etc., se comprueba sistemáticamente que frente a estímulos idénticos los sujetos coinciden en sus percepciones dentro de unos márgenes muy concretos y estrechos. Es decir, que la relación entre las condiciones físicas externas y los resultados de la percepción interna son regulares (Guski, 1992, pág. 24). La consecuencia directa de esto es que el concepto de *subjetividad*, entendido desde el punto de vista de la psicología de la percepción, es perfectamente objetivable. Dicho de otro modo: toda sensación que se desencadena en el interior de un sujeto como respuesta a la recepción de estímulos físicos externos objetivos y toda interpretación de estas sensaciones que haga el mismo sujeto receptor son objetivables. Es decir, formalizables como un objeto de estudio científico.

Entenderemos pues, en esta obra, el concepto de *subjetivo* en el sentido que lo entiende la psicología de la percepción. Así, denominaremos subjetivo a todo aquello que es inherente a la fenomenología objetivable, analizable y mensurable que se produce en el interior del sujeto. Y, como corolario, entenderemos el concepto de *objetivo* como todo aquello que es inherente a los fenómenos físicos externos al sujeto.

Este planteamiento inicial nos proporciona un fundamento conceptual muy claro para estudiar el lenguaje audiovisual desde una perspectiva psicofísica. Es evidente que si no hubiese coherencia entre las percepciones de los distintos receptores no sería posible ningún tipo de comunicación. Que productos audiovisuales

suales como *Cantando bajo la lluvia*, *Blancanieves y los siete enanitos* o *La Guerra de las Galaxias* puedan recorrer el mundo conservando intacta su organización visual y sus bandas sonoras de música y efectos, sólo es comprensible si existe coherencia perceptiva entre los receptores; sin una base de apoyo en códigos universales ese fenómeno comunicativo transcontinental no sería posible. Y esto no es en absoluto exclusivo del lenguaje audiovisual. Toda comunicación, sin excepción posible, tiene su origen antropogénico en la interpretación *subjetiva*, coherente y unívoca, por parte de un grupo de individuos, de una serie de estímulos físicos *objetivos*.

#### 4.1.1.1. En busca de mecanismos expresivos universales

Hemos estudiado en el capítulo anterior cómo fenómenos físicos y experiencia perceptiva son procesos radicalmente distintos que mantienen entre ellos unas relaciones regulares aunque, a menudo, bastante complicadas. El sistema perceptivo traduce las variaciones físicas de nuestro entorno en sensaciones siguiendo unas leyes que investiga e intenta encontrar la psicología de la percepción. En el tema concreto que nos ocupa la disciplina que nos aporta información sobre esas leyes es la psicoacústica.

Así, hemos visto cómo la longitud de las ondas sonoras (frecuencia) desencadena la sensación tonal; cómo la presión sonora (amplitud) implica sensación de intensidad; y que las distintas formas de organización del espectro desencadenan sensaciones tímbricas. Éstos son sólo unos primeros pasos todavía tambaleantes en nuestro camino. Nuestro objetivo será ahora revisar los elementos esenciales de la percepción auditiva, con el fin de comprender mejor el funcionamiento global de los fenómenos expresivos audiovisuales. Cuando seamos capaces de descubrir de qué modo actúa la interpretación *subjetiva* de los estímulos físicos audiovisuales, estaremos en mejores condiciones para localizar algunos de los mecanismos expresivos universales en los que se basa el lenguaje audiovisual.

#### 4.2. La percepción de las dimensiones sonoras simples

En este apartado, revisaremos cada uno de los umbrales perceptivos relacionados con la intensidad, el tono y la duración. Es decir, estudiaremos con detalle el primer nivel del proceso interpretativo que desencadena el sistema perceptivo humano cuando nuestro oído está expuesto a estímulos sonoros.

##### 4.2.1. El concepto de umbral

El concepto de umbral está asociado al límite de sensibilidad perceptiva. La psicología de la percepción trabaja de una manera sistemática con tres tipos de

umbrales perceptivos: el *umbral mínimo*, el *umbral máximo* y el *umbral diferencial*. Veamos la definición de cada uno de estos conceptos.

*Umbral mínimo*: supone detectar la cantidad mínima de estímulo por debajo de la cual el fenómeno físico ya no es percibido por el ser humano.

*Umbral máximo*: es justamente el concepto antitético, supone detectar la cantidad de estímulo por encima del cual el fenómeno tampoco puede ser ya percibido por el hombre.

*Umbral diferencial*: se define como la cantidad mínima de variación de estímulo que puede ser percibida por el sistema perceptivo. Cuando esta cantidad de variación física es inferior a la del umbral diferencial, el ser humano deja de notar cambios en el fenómeno y lo percibe como constante, a pesar de que se haya producido una variación física objetiva.

#### 4.2.2. Umbrales de intensidad

El *umbral mínimo* de intensidad se define como de 0 dB a 1.000 Hz. Es decir, una frecuencia pura de 1.000 Hz con una presión sonora que esté por debajo de los 0 dB ya no es audible por el ser humano. No obstante, el umbral mínimo de la sensación de intensidad tiene una fuerte variabilidad en función de la frecuencia. Por ejemplo a 2.500 Hz podemos percibir hasta -16 dB, y a frecuencias bajas (50 Hz por ejemplo) son necesarios hasta 50 dB para llegar a escuchar el sonido. El lector puede observar todo esto gráficamente en el diagrama de Fletcher (fig. 24).

Es interesante señalar que el silencio absoluto no existe. Cuando se produce la ausencia total de sonidos exteriores se oyen los ruidos del propio cuerpo. Quien escribe estas líneas ha podido experimentar, dentro de la cámara anecoica<sup>1</sup> del Instituto de Acústica del C.S.I.C. en Madrid, cómo los ruidos de su propia respiración y de los pequeños movimientos internos de su cuerpo adquirirían, en una situación de ausencia total de sonido externo, un protagonismo inesperado y sorprendente.

El *umbral máximo* en audición se denomina *umbral de dolor*. Esta forma de nombrarlo se debe a que a partir de él la sensación perceptiva es dolorosa. El umbral de dolor auditivo es bastante regular, es decir, varía poco con la frecuencia. Está situado en torno a los 130 dB (véase también fig. 24). A partir de este nivel de energía sonora el oído humano pasa de percibir sensación sonora

1. Se denomina *cámara anecoica* a un espacio que cumple las condiciones de no dejar pasar absolutamente ningún sonido del exterior, y dentro de la cual no se produce ningún tipo de reflexiones sonoras. Es una sala situada en el interior de un gran cubo de hormigón. El cubo está separado del suelo por un sistema de soportes capaces de evitar la transmisión de vibraciones, y envuelto en una malla conductora (jaula de Faraday) que evita toda influencia radioeléctrica en su interior. El suelo, el techo y las cuatro paredes interiores están totalmente recubiertos por aislante acústico y sobre esta capa aislante se instalan grandes cuñas de material absorbente. El acceso al interior de la sala se realiza sobre una malla metálica que flota entre el suelo y el techo.

a percibir sensación de dolor y corre el riesgo de sufrir daños fisiológicos importantes.

El *umbral diferencial* auditivo vuelve a ser, también, muy variable en función de la frecuencia. El umbral diferencial medio es de +3 dB y oscila entre 2 y 10 dB en función del tono. Esto significa que para que el oído humano pueda percibir que se ha producido una variación de intensidad sonora es necesaria una variación mínima de la presión sonora de unos 3 dB; por debajo de esta cantidad de variación, el cambio no será percibido y la sensación será de intensidad constante. El umbral aumenta en las frecuencias bajas (es necesaria una mayor cantidad de variación) y disminuye a medida que la frecuencia escucha es más alta.

#### 4.2.3. La sensación de intensidad en los sonidos compuestos

Todo lo que hemos revisado hasta ahora sobre la intensidad en decibelios hacía referencia a una frecuencia pura de 1.000 Hz, pero hemos visto, también, que lo habitual era escuchar sonidos compuestos y, por tanto, constituidos por muchas frecuencias. Esto supone que la sensación real de intensidad está directamente relacionada con la suma de todas las frecuencias que componen globalmente el espectro de un sonido.

No obstante, acabamos de decir que tanto el *umbral mínimo* como el *umbral diferencial* de la intensidad son variables en función de la frecuencia sonora; lógicamente, esto significa que la sensación de intensidad que percibe el oído humano al escuchar un sonido compuesto no responde exactamente a la suma de intensidades de todos los componentes frecuenciales del sonido, sino a una suma ponderada en función de la sensibilidad que el ser humano tiene de cada frecuencia. Técnicamente, este problema se ha resuelto estableciendo curvas normalizadas para ponderar la sensación de intensidad correspondiente a cada zona del espectro. En la figura 25 podemos ver tres curvas clásicas utilizadas para ponderar la intensidad de los distintos componentes de frecuencia antes de sumarlos para obtener una lectura única global.

Según se haga la ponderación para obtener la intensidad global aplicando la curva A, la B o la C se obtendrá un valor en dB distinto, es por eso que en el ámbito técnico se habla de decibelios A (dBA), decibelios B (dBB), o decibelios C (dBC), según el cálculo de la potencia sonora se haya realizado con una u otra curva. Cuando las medidas en decibelios no están ponderadas, es decir, cuando hacen referencia directamente a la sensibilidad del oído humano a 1.000 Hz se habla de decibelios NPS<sup>2</sup> (dB NPS) significando NPS: Nivel de Presión Sonora.

De las curvas presentadas en el gráfico de la figura 25 nos interesa fundamentalmente la A en tanto que es la que se aproxima con más eficacia a la percepción

2. En textos técnicos se utiliza también con mucha frecuencia la notación inglesa *Sound Pressure Level*, que en siglas es S.P.L., es decir: dB SPL.

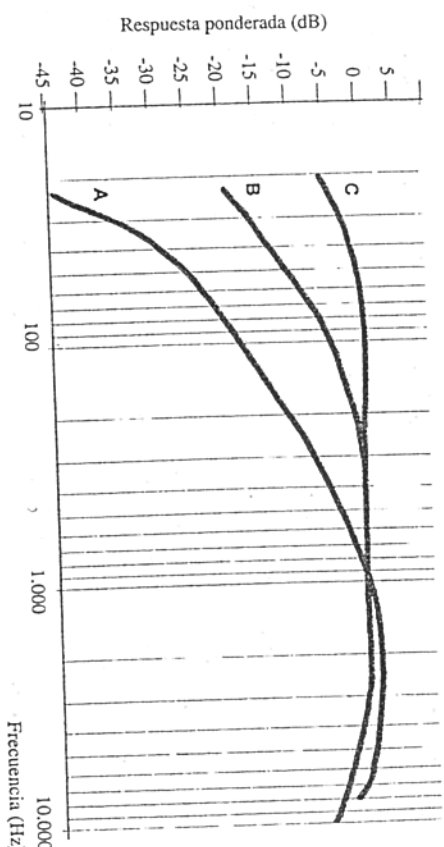


Figura 25. Diagrama de curvas de ponderación. Destinado a calcular la intensidad global en dB de los sonidos complejos.

humana de la sonoridad. La escala de ponderación para calcular la sonoridad que define la curva A es la más utilizada de las tres con las que pueden trabajar normalmente los sonómetros.<sup>3</sup>

Pero veamos con más detalle cómo funciona este gráfico. Éste es ya un instrumento adaptado a la percepción humana en tanto que sus dos escalas son logarítmicas. El eje vertical indica la cifra en decibelios que es necesario restar a la medida física de presión tomada por el sonómetro, para que el resultado final se adapte a la percepción psicológica de intensidad (los dB son ya unidades logarítmicas). El eje horizontal indica la frecuencia en hercios de los componentes espectrales del sonido compuesto. Como habrá observado el lector, el eje horizontal está graduado logarítmicamente para ajustar visualmente el papel de cada frecuencia a su protagonismo real en la percepción humana.

El modo de utilizar el diagrama es situando, una a una, en el gráfico, las respectivas medidas directas de intensidad en dB que toma un sonómetro sobre el sonido que estamos estudiando. Para situar cada medida en el gráfico y ponderarla necesitamos saber la frecuencia correspondiente a la que ha sido tomada. Supongamos, por ejemplo, que el sonómetro ha tomado una medida directa de intensidad de 80 dB en la banda de 90 Hz; nos hemos de situar, entonces, en el eje horizontal sobre la frecuencia correspondiente (90 Hz) y ascender verticalmente desde este punto hasta encontrarnos con la curva A, y una vez aquí, la proyección de este punto sobre el eje de ordenadas (vertical) nos dará la cifra en dB con la

3. Se denomina *sonómetro* al instrumento que mide la presión ejercida por las ondas sonoras. El sonómetro selecciona sistemáticamente las distintas bandas de frecuencia del espectro sonoro, mide la presión que ejerce cada una de ellas y, antes de sumarla, las modifica en función de una escala predeterminada, dando finalmente una medida resultante ponderada de intensidad global en decibelios.

que hemos de ponderar la medida directa. En este caso concreto la ponderación es de -20 dB. Es decir, la medida ya ponderada será de:

80 dB - 20 dB = 60 dB

Esta operación se realizará con cada una de las medidas directas, situándola en la banda correspondiente de frecuencia a la que ha sido tomada. Una vez ponderadas todas las medidas directas estaremos ya en condiciones de sumarlas y obtener la intensidad global, en este caso en dBA.

Es necesario, pues, realizar una serie de operaciones bastante complejas para llegar a obtener medidas de intensidad que se aproximen a la percepción humana de los sonidos compuestos. Todo este proceso de cálculo y ponderación que hemos descrito en los apartados 3.4.2 y 4.2.3 es el que se desarrolla normalmente de forma automática para obtener las medidas en decibelios que suelen manejarse en el ámbito audiovisual. Las cifras en dB que se usan normalmente son medidas ponderadas tomadas sobre sonidos compuestos; es decir suelen ser dBA.

Una tabla orientativa sobre la medida en decibelios A de algunos de los ambientes sonoros en los que cotidianamente estamos inmersos podría ser la siguiente:

| Descripción de la fuente              | Nivel (dBA) |
|---------------------------------------|-------------|
| Umbral de audición                    | 0 - 5       |
| Estudio de radio vacío                | 20 - 25     |
| Aula vacía                            | 25 - 30     |
| Ruido de fondo de una casa silenciosa | 30 - 42     |
| Conversación normal                   | 40 - 55     |
| Automóvil a 80 km/h                   | 45 - 50     |
| Radio o TV en una habitación          | 50 - 60     |
| Coches en el tráfico urbano           | 65 - 85     |
| Camiones en el tráfico urbano         | 70 - 90     |
| Martillo neumático                    | 90 - 105    |
| Avión a reacción                      | 105 - 130   |

Estas intensidades se toman como referencia escuchando la emisión de la fuente sonora a 1 metro de distancia.

4.2.4. Influencia de la distancia en la sensación de intensidad

La distancia entre la fuente sonora y el receptor influye también en la sensación de intensidad disminuyéndola de una forma regular. Igual que en todo fenómeno físico, existe un consumo de energía en el esfuerzo de propagación del so-

nido. Este consumo atenúa progresivamente el fenómeno vibratorio hasta llegar a eliminarlo. Concretamente: cuando un sonido se propaga a través del aire, cada vez que se dobla la distancia entre la fuente sonora y el receptor la intensidad disminuye en 6 dB. Y viceversa, cada vez que la distancia se reduce a la mitad la distancia aumenta en 6 dB.

Veamos esto en un ejemplo. Si escuchamos un sonido de 90 dB oído desde 3 metros de distancia, cada vez que doblemos la distancia entre nuestro oído y la fuente sonora alejándonos desde 3 hasta 6 metros, luego de 6 metros hasta 12 y así sucesivamente, observaremos que la intensidad disminuye 6 dB cada vez. Es decir:

| Distancia | Intensidad |
|-----------|------------|
| 3 m       | 90 dB      |
| 6 m       | 84 dB      |
| 12 m      | 78 dB      |
| 24 m      | 72 dB      |
| 48 m      | 66 dB      |
| 124 m     | 60 dB      |

Naturalmente el fenómeno de atenuación debida a la distancia recorrida por el sonido es independiente de la potencia con la que está emitiendo la fuente sonora, que en este caso concreto sería constante.

El fenómeno de la atenuación de la intensidad sonora en función de la distancia resulta especialmente interesante para el lenguaje audiovisual. Todo alejamiento entre fuente sonora y oyente se percibe mediante la sensación de atenuación de la intensidad sonora que percibe el receptor. Y a cada distancia entre una fuente sonora concreta y el oyente corresponde una intensidad determinada. Además, todo oyente, esté o no habituado a la comunicación audiovisual, tiene una larga experiencia vital en percibir cómo la distancia entre él y los objetos que sue-

nan influye disminuyendo la sensación de fuerza del sonido. Así, el alejamiento entre el micrófono que capta un sonido determinado y la fuente sonora que lo emite, se refleja de forma física siguiendo esta regla de los 6 dB; con lo que basta acortar o alargar la distancia entre micro y fuente sonora para controlar la sensación de distancia que va a percibir cualquier oyente cuando oiga su reproducción.

4.2.5. Influencia de la intensidad en la sensación de distancia

Este que acabamos de explicar es, sin duda, el nivel más elemental de uso de la «regla de los 6 dB». Pero resulta bastante más interesante la posibilidad de reconstruir sensaciones de acercamientos, alejamientos y distancias concretas (próximas o lejanas) sin que éstas existan en realidad. El universo de la producción de efectos sonoros es depositario de una larga y vieja tradición en manipular la sen-



sación de distancia entre los *entes acústicos* (véase apartado 2.3.4) y el oyente, que tiene su origen en el radiodrama. La posibilidad de manipular artificialmente la intensidad de cualquiera de las señales sonoras que se están manipulando en el estudio de audio, simplemente moviendo un potenciómetro, posibilita el control del número exacto de dB con el que va a emitir una *fente sonora* una vez grabada, y, en consecuencia, también el control de la sensación de distancia entre *fente sonora* y receptor, a pesar de que ésta permanezca quieta y muy cercana al micrófono.

Vayamos, como siempre, a algún ejemplo concreto. Imagine el lector el locutorio de un estudio de audio en el que hay tres personas conversando que llamaremos: L.1, L.2 y L.3, sentados alrededor de una mesa triangular, cada uno en un lado de la mesa y situado a 20 cm de distancia de un micrófono. Los tres micros son fuertemente direccionales con objeto de que cada uno de ellos capte exclusivamente la voz del locutor que está sentado frente a él. Finalmente, la señal de cada micro llega a la mesa de control de audio del estudio por un canal distinto, lo que posibilita la manipulación independiente del sonido de cada uno de los locutores. Este tipo de configuración técnica abre un interesante abanico de posibles aplicaciones del tratamiento artificial de la intensidad.

Podemos, por ejemplo, equilibrar las tres intensidades de forma que sean idénticas a 55 dB, con lo cual la sensación de distancia que tendríamos al escuchar el sonido tratado electrónicamente sería la de estar escuchando a los tres locutores desde muy cerca, a la misma distancia los tres, aproximadamente a un metro de ellos. Algo así como si el oyente estuviese en el centro de la mesa. Otra posibilidad es atenuar artificialmente las tres señales, de nuevo de forma equilibrada, por ejemplo, bajando su nivel de intensidad hasta 37 dB. Ahora la sensación al escuchar el sonido tratado será la de escuchar a tres personas que habían lejos de nosotros a unos 8 metros de distancia. Este efecto perceptivo se producirá sin que ellos se hayan movido físicamente ni un solo centímetro de su posición original. Podríamos también combinar ambos tratamientos manipulando la señal del locutor L.1 de modo que sonase a 55 dB, mientras que las de los locutores L.2 y L.3 fueran atenuadas hasta bajar a 37 dB. Escuchar el sonido resultante de esta última manipulación generará en el oyente la percepción de que el locutor L.1 está muy cerca (1 m), mientras que L.2 y L.3 se han alejado mucho y habían ambos desde la misma distancia (8 m).

Si realizamos tratamientos progresivos de la intensidad conseguiremos sensación de movimiento. Partamos del primer ejemplo: las tres señales a 55 dB, si desde esta situación realizamos una atenuación progresiva de la señal del L.3 bajando su intensidad lentamente hasta llevarla a 0 dB, la sensación que se generará será la de que el locutor L.3 se va alejando de nosotros y de los otros dos locutores, que permanecen cerca y quietos, hasta perderse en la lejanía.

Todas estas sensaciones se habrán producido sin que exista ningún movimiento real de las fuentes sonoras. Los locutores permanecieron quietos todo el tiempo, sólo fue alterada electrónicamente la información acústica que el ser humano utiliza para interpretar la distancia a la que se encuentra de él cualquier fuente sonora real. Se ha manipulado la esencia perceptiva sonora creando en la

mente del oyente interpretaciones de fenómenos que no existían en la realidad referencial. Hemos creado, en definitiva, *entes acústicos* y los hemos tratado para que actúen sobre la percepción sonora humana como lo habrían hecho las fuentes sonoras reales (en este caso los tres locutores sentados), si realmente se hubiesen movido alejándose y acercándose de los micrófonos.

Concluimos este apartado extrayendo algunas conclusiones que pueden resultar muy útiles para la narración sonora.

Si en el apartado 4.2.2 decíamos que el *umbral diferencial* medio de intensidad está en torno a los 3 dB, es decir, que por debajo de esta cantidad de variación de intensidad el cambio no es perceptible; y la regla de los 6 dB nos dice que cada vez que se dobla, o reduce a la mitad, la distancia entre fuente sonora y receptor se produce una variación de la sensación de intensidad de 6 decibelios. Una solución simple y lógica para garantizar que cualquier receptor, incluso los «duros de oído», perciben las variaciones artificiales de la intensidad es trabajar siempre a partir de la regla de los 6 dB. De este modo, dispondremos de una referencia empírica perfectamente definida y nos moveremos con unos amplios márgenes de seguridad que garantizarán la percepción y la comprensión de todas las manipulaciones de la intensidad.

En suma, de toda esta reflexión podemos deducir el siguiente principio expresivo: cuando un realizador intenta que los oyentes que contemplen su narración audiovisual perciban y comprendan sólo a través de la intensidad sonora las distancias y los movimientos de sus *entes acústicos*, es conveniente que los sonidos que constituyen esos *entes* sean manipulados de modo que sus alteraciones de intensidad se organicen en variaciones de 6 en 6 decibelios.

Intentaremos explicar la funcionalidad de esto con algunos ejemplos:

Un realizador intenta que los oyentes que contemplen su película perciban y comprendan solamente a través de la banda sonora que Jack el Destripador se ha alejado momentáneamente de su nueva víctima,<sup>4</sup> y manipula los sonidos que constituyen el *ente acústico* «Jack el Destripador» (por ejemplo: ruido de pasos y respiración agitada) de modo que sus alteraciones de intensidad se organicen en variaciones de 6 en 6 decibelios. La aplicación de esta norma supone tres claras ventajas: 1) garantiza que toda variación de la intensidad va a ser perceptible por cualquier receptor, porque está muy por encima del umbral diferencial; 2) durante la realización, mientras se producen y se graban los sonidos, da una referencia muy clara de cómo han de moverse las *fuentes sonoras* frente a los micrófonos, puesto que sabemos que duplicando o reduciendo a la mitad la distancia entre fuente y micrófono el resultado acústico es de 6 dB más o 6 dB menos; y, finalmente, 3) permite prever y controlar técnicamente, desde la manipulación artificial de la intensidad, la posición a la que vamos a colocar al asesino respecto al oyente.

Veamos con algo más de detalle esta última ventaja. Teníamos al *ente acústico* «Jack el Destripador» respirando, por ejemplo, a una intensidad de 25 dB, es decir,

4. Estamos imaginando una situación narrativa en la que el asesino está siempre oculto, y el espectador sólo lo identifica por el ruido de sus pasos y una característica forma de respirar.

aproximadamente a un metro del oyente y, por tanto, muy cerca de la víctima. Es decir, haciendo sentir al espectador que va a producirse un apuñalamiento en cualquier instante. Y, ahora, deseamos alejarlo acústicamente lo suficiente para que el espectador sienta que ha desaparecido el riesgo inminente y que baja algo la tensión dramática. Queremos alejar al «Destripador» a una distancia, por ejemplo, de 8 metros, desde la cual es ya muy improbable que apuñale a la víctima, para ello bastará con reducir la intensidad del sonido de «Jack» en 18 dB. El cálculo necesario es simple. Si tenemos al asesino a 1 metro y queremos alejarlo a 8 metros hemos de duplicar 3 veces la distancia:

$$1 \times 2^3 = 8$$

Traduciendo esto en potencias sonoras y aplicando la regla de los 6 dB, hemos de minorar la intensidad 3 veces 6 dB, es decir, reducir los 25 dB con los que está sonando «el Destripador» a sólo 7 (25 dB - 18 dB = 7 dB). De esta manera, a partir de un efecto pregrabado de pasos y respiración agitada, el realizador puede conseguir cómodamente un control virtual de los movimientos del *ente acústico* «Jack el Destripador» considerablemente exhaustivo y preciso.

Naturalmente, hemos de aclarar aquí que nos estamos refiriendo, todavía, única y exclusivamente a sensaciones virtuales de alejamiento y acercamiento. Mas adelante revisaremos la construcción global del espacio acústico e introduciremos otras dimensiones del movimiento y la espacialidad sonora. Recuerde el lector que ésta es solamente una ilustración sobre la influencia de la intensidad en la sensación de distancia.

#### 4.2.6. Umbrales de tono

**Umbral mínimo:** El oído humano comienza a percibir los fenómenos vibratorios cuando éstos se producen con un mínimo de 16 oscilaciones por segundo. Cualquier vibración de frecuencia menor no es percibida por el hombre. Se habla entonces de «infrasonidos». Una vibración de 16 cps será el tono más grave o bajo que puede percibir nuestro sistema auditivo. Este tono se corresponde con el que genera el *do* inicial del órgano.

**Umbral máximo:** La frecuencia máxima que es capaz de percibir el oído humano está en torno a los 20.000 cps a partir de aquí los sonidos son inaudibles y se habla de «ultrasonidos». El tono más alto o agudo que puede escuchar el hombre es, pues, el que genera una frecuencia de 20.000 Hz. Los instrumentos musicales clásicos no alcanzan esta frecuencia. El *re* de pícolo, por ejemplo, alcanza 4.700 cps, la gama de frecuencias del piano va de 27 a 3.500 cps y la frecuencia fundamental de la voz de 65 a 1.036 cps.

**Umbral diferencial:** El umbral diferencial de tono definido a 1.000 Hz es de 2 cps. Este umbral cambia haciéndose más grueso a medida que aumenta la frecuencia. Tal como hemos visto ya en otros fenómenos perceptivos de la audición, mientras la frecuencia de un sonido aumenta exponencialmente, la sensibilidad al

cambio de tono varía aproximadamente de forma lineal. Escuchando un sonido de 500 Hz podemos percibir sensación de cambio de tono cuando la frecuencia aumenta entre 0,5 cps y 1 cps. Es decir, cuando la frecuencia pasa de 500 a 501 Hz percibimos auditivamente que el tono ha subido. A medida que el estímulo aumenta, la sensibilidad auditiva al cambio va disminuyendo. Escuchando un sonido de 1.000 Hz el umbral diferencial del tono es de 2 cps. O sea, nuestro oído percibe que el tono ha subido cuando el sonido pasa de 1.000 a 1.002 Hz, pero ya no siente ninguna sensación de variación si el cambio es solo de 1 cps. Para un sonido puro de 2.000 Hz el umbral diferencial es de unos 4 cps y para uno de 8.000 Hz el umbral diferencial es de unos 100 cps.

Podemos deducir, en consecuencia, que la capacidad de resolución tonal del oído humano es mucho más fina de lo que define la escala tonal heptatónica occidental. Nuestra escala divide cada octava en 7 tonos o notas y 12 semitonos. El oído, en cambio, es capaz de diferenciar unos 300 sonidos en una octava situada sobre los 500 Hz.

#### 4.2.7. Sensibilidad absoluta y sensibilidad relativa respecto al tono

En general, la capacidad perceptiva responde mucho mejor a las relaciones entre unos sonidos y otros que a su valor absoluto. La percepción tonal está esencialmente influenciada por este fenómeno. En este sentido, es revelador que los músicos profesionales, a pesar de su altísima especialización auditiva, sólo sean capaces de identificar con precisión la frecuencia de una nota aislada cuando ésta está tocada con su instrumento habitual, pero en el momento en que la nota está construida con cualquier otra fuente sonora, por ejemplo, con un generador digital de sonidos puros, no son capaces de identificar con precisión su altura tonal. Paradójicamente, cualquier individuo no especializado es perfectamente capaz de reconocer una melodía familiar para él a pesar de que ésta sea tocada a distinta altura tonal o con diferentes instrumentos. Reconocemos, por ejemplo, *Noche de paz* independientemente de que ésta sea construida con todas sus notas situadas una octava más arriba o una octava más abajo; e independientemente, también, de que sea tocada con sonido de piano o con un nuevo sonido electrónico recién sintetizado.

Esto se debe a que nuestra sensibilidad auditiva es mucho más alta para las relaciones entre estímulos que para los estímulos aislados. Reconocer *Noche de paz*, que el lector habrá escuchado cantada y tocada de mil modos distintos, supone identificar una determinada estructura relacional entre una secuencia de sonidos ordenados en el tiempo. Y esa relación es independiente de su situación dentro de nuestro mapa auditivo. Es lo mismo que sea tarareada por voces infantiles, cantada en inglés por un tenor, en ruso por un bajo, tocada en un órgano, o por una flauta dulce.

Es algo así como reconocer el esquema visual de una cara formada por unas pocas líneas que dibujan un círculo con dos puntos representando los ojos, una pequeña raya vertical en el centro para hacer la nariz y una línea horizontal en la parte inferior para mostrar la boca. Da igual el color de las líneas, su grosor, tampoco importa el color del fondo ni el lugar en el que esté situado el dibujo, seguiremos reconociendo la cara. Lo importante en este reconocimiento es la relación entre los propios elementos que configuran la forma reconocible.

Este tipo de reconocimiento sonoro de formas relacionales es muy claro en el habla. El efecto de interrogación, por ejemplo, se construye mediante el ascenso tonal de la última sílaba de la frase. Y éste es un efecto reconocible con absoluta independencia de la extensión tonal que tiene la voz que hace la pregunta. Sentimos el efecto interrogación tanto en la voz aguda de un niño como en la voz grave y oscura de cualquier hombre maduro. En suma: toda estructuración tonal responde a una organización global de relaciones de frecuencia y no a las frecuencias en sí mismas como valor absoluto.

#### 4.2.8. *Umbrales temporales*

Los umbrales temporales de la audición varían tanto en función del tono como de la intensidad del sonido y esto hace especialmente complicado recoger y proponer ideas claras y útiles sobre ellos. Podemos, no obstante, trabajar con algunas referencias orientativas.

*Umbral mínimo.* La duración mínima que ha de tener un sonido aislado para que pueda ser oído es de 5 ms (milésimas de segundo). Por debajo de esta cantidad de tiempo el sonido no se percibe. La sensibilidad es algo más alta en la situación inversa, o sea, cuando el sonido es constante e introducimos en él un fragmento de silencio. En este caso podemos ir reduciendo la duración de la interrupción hasta que dejamos de percibirla cuando ésta es inferior a 3 ms.

Es importante aclarar que la sensación de presencia sonora es independiente de las sensaciones de intensidad tono y timbre.

Respecto a la intensidad: se observa que manteniendo constante la presión sonora de un tono audible y haciendo escuchar *sucesos sonoros*<sup>6</sup> cortísimos de éste que se hacen progresivamente más largos, la sensación de intensidad comienza muy débilmente cuando los *sucesos* son de 5 ms, y crece desde los 5 ms del umbral hasta que los sucesos duran entre 100 o 150 ms, entonces la sensación de intensidad se torna ya estacionaria. Resumiendo: es necesario que un sonido dure por lo menos 150 milésimas de segundo para que el oído humano perciba su intensidad real. Y a la inversa, cuando el sonido tiene una intensidad muy débil, concretamente, cuando estamos escuchando un sonido en los umbrales de intensidad, es necesario que el *suceso sonoro* dure 0,5 segundos para que pueda ser audible (véase Schaeffer, 1988, págs. 115 y 119).

6. Véase el apartado 2.3.3.

Respecto al tono: cuando un *suceso sonoro* dura menos de 0,1 segundos se pierde la sensación tonal y el sonido se percibe solamente como un «clic». Dicho de otro modo, los *sucesos sonoros* han de superar los 0,1 segundos de duración para que su tono sea perceptible.

Respecto al timbre: para que sea posible reconocer simultáneamente las características de tono y timbre de un sonido es necesario que el *suceso sonoro* escuchado tenga una duración de más de 1 segundo (véase Guski, 1992, pág. 165).

*Umbral diferencial.* El ser humano es capaz de percibir variaciones de intensidad de 3 dB (umbral diferencial de intensidad) cada vez más próximas en el tiempo unas de otras hasta que el tiempo de una variación a otra se reduce entre 2 y 7 milésimas de segundo. Es decir, podemos llegar a percibir entre 140 y 500 variaciones de intensidad cada segundo. Cuando el tiempo entre un cambio de intensidad y otro es menor de 7 ms la sensación auditiva es ya de sonido continuo, y la variación no es perceptible (véase Guski, 1992, pág. 167).

Posiblemente, para el lector que todavía esté trabajando en un entorno exclusivamente analógico de tratamiento del sonido, este tipo de referencias micrométricas sobre los umbrales temporales sean poco menos que incomprensibles. Ciertamente, con un buen cronómetro en la mano y mucha habilidad, o con los contadores de vueltas o de tiempo de los magnetófonos, es imposible afinar más allá de las décimas de segundo. Y aquí estamos hablando de segundos divididos en mil partes. No obstante, para todos aquellos que hayan tenido alguna vez en sus manos un equipo para la edición digital de audio, la manipulación de *sucesos sonoros* de algunas centésimas o milésimas de segundo resultará ya algo perfectamente natural. Justamente en este terreno y trabajando con este tipo de instrumentos es cuando el conocimiento de los umbrales perceptivos adquiere todo su sentido.

Cuando tenemos en las manos un instrumento que nos permite afinar hasta la manipulación de las milésimas de segundo, con toda seguridad vamos a trabajar con esta resolución temporal en nuestros montajes. Pero ¿qué sentido puede tener, por ejemplo, manipular e insertar cuidadosamente fragmentos sonoros de unas pocas centésimas (o milésimas) de segundo, cuando la percepción sonora completa no se produce hasta que escuchamos «enormes» *sucesos sonoros* de más de 1 segundo? Pienso que es importante recordar aquí, de nuevo, que la codificación eficaz de los mensajes audiovisuales no debe estar conducida por el dominio de las posibilidades tecnológicas, sino por el conocimiento de las capacidades perceptivas. Y la instrumentación digital de tratamiento del audio abre muchas posibilidades que rebasan la capacidad perceptiva del ser humano. La consecuencia directa de centrar nuestra atención exclusivamente en la tecnología nos lleva con demasiada frecuencia a desaprovechar equipos excelentes, utilizándolos de forma absolutamente absurda e ineficaz; y a no obtener todo el rendimiento posible de otros, en virtud de su supuesta obsolescencia tecnológica.

*Umbral máximo.* No se ha definido un umbral máximo de duración auditiva, no obstante la atención está estrechamente ligada a la variabilidad temporal. Si un sonido permanece estacionario el receptor deja de prestarle atención. Y la atención vuelve a ser de nuevo activa cuando aparecen variaciones en alguna de las

dimensiones del sonido, es decir, en el momento en que el sonido deja de ser es-tacionario. La fuerza de recuperación de la atención es tanto mayor cuanto más rá-pida es la variación del sonido. Podríamos concluir que la variabilidad es un ins-trumento esencial del estímulo sonoro para activar el sistema perceptivo.

Intentaremos desarrollar esto con más detalle.

Antropogenéticamente, el fenómeno perceptivo de la alta sensibilidad al cam-bio puede asociarse a la señal acústica de peligro que constituyen los sonidos ge-nerados por el acercamiento rápido de un predador. Los humanos primitivos, co-mo cualquier especie intermedia dentro de la cadena alimentaria, habían de estar muy atentos a las variaciones del entorno sonoro. Una vez reconocido un espacio por todos los sentidos como lugar sin peligro, el humano podía proceder a des-cansar relajando conscientemente sus músculos y sus sentidos. Excepto el de la audición, que actúa como un sistema de atención pasiva independiente que pone en marcha de forma automática todo el sistema perceptivo al detectar alguna va-riación. Cualquier cambio sonoro en el contexto inmediato es indicación de cam-bio físico en el espacio próximo y, por tanto, posible indicio de la aproximación de un animal peligroso.

La técnica del cazador es, y ha sido siempre, la de evitar la producción de cualquier estímulo perceptible que pueda avisar a la presa de su presencia, hasta el momento en que la proximidad permita sorprenderla con la rapidez suficiente como para que sea inútil su huida. Lógicamente, cambios sonoros rápidos indican movimientos rápidos y, por tanto, un posible predador atacando. En consecuen-cia, la supervivencia de la presa depende de que pueda reaccionar tanto más rá-pidamente cuanto más rápidas sean las variaciones del estímulo sonoro detecta-do. Afortunadamente, este mecanismo perceptivo sigue actuando en el ser humano y permitiéndonos brincar hacia la acera cuando nos sobresalta el claxon de un automóvil.

Pero vayamos al papel que puede desempeñar este fenómeno en el lenguaje audiovisual, que es el objeto de nuestra obra.

En la década de los noventa han aparecido una serie de producciones cine-matográficas que utilizan de manera fundamental la banda sonora para controlar los estímulos emocionales del espectador. Los nuevos sistemas de sonorización envolvente de las salas de cine, que emite con 4, o 6 vías perfectamente diferen-ciadas: 1 tras la pantalla (ubicación de los personajes), 1 para bajas frecuencias, 2 laterales que crean el efecto izquierda/derecha y 1 o 2 que distribuyen el sonido a nbiante de forma envolvente por el resto de la sala, permiten trabajar con mu-cha más eficacia los efectos sonoros sin que éstos enmascaren las voces. Este nue-vo sistema de sonorización ha posibilitado el desarrollo de trabajos muy intere-santes en el uso narrativo del sonido. Concretamente pensamos en películas como *L'amaradas* o *Stargate* que consiguen transmitir al espectador una fortísima sen-sación de tensión mantenida, prácticamente, a lo largo de toda la proyección. Esa tensión depende sobre todo del sonido. En la sala, los espectadores están metidos en una burbuja sonora con formas acústicas muy bien definidas y de gran poten-cia, que se construye desde los altavoces envolventes del cine, y esta burbuja so-nora es constantemente cambiante. Eso obliga al «espectador-presa» a permanecer

en atención activa esperando al virtual «depredador» cinematográfico. Por otra parte, los momentos de la narración que contienen situaciones de peligro están subrayados con una gran estruendo sonoro de espectro dinámico,<sup>7</sup> que comunica una fuerte sensación mantenida de peligro. Por si esto fuera poco, los cambios de espacio de los personajes son tratados también como cambios del punto de audi-ción del espectador. El espectador oye en cada momento desde el lugar en el que están los personajes. Con lo cual el estruendo acústico del peligro cambia global-mente de forma (se escucha desde dentro, luego desde fuera, después se acerca, vuelve a alejarse, etc.) definiendo *sucesos sonoros* distintos con una frecuencia que oscila entre varias decenas de segundos y algunos minutos.

El resultado global es la configuración de un estilo en la manipulación del so-nido que somete al espectador a larguísima períodos de tensión, manteniéndolo en atención constante gracias a la sucesión de cambios sonoros que evitan la re-lajación perceptiva. Cuando el sonido está en cambio permanente, la atención permanece activa con total independencia de la voluntad racional del receptor. Probablemente el lector habrá notado ya en su propia carne, en más de una oca-sión, el efecto agotador de este tipo de recursos expresivos que trabajan con los umbrales temporales de atención.

#### 4.3. La percepción de la complejidad sonora interna

Hasta ahora, hemos revisado en este capítulo los conocimientos disponibles sobre la percepción de la *intensidad*, el *tono* y la *duración* de los sonidos, unos en función de otros, pero siempre como dimensiones aisladas. En este apartado nos dedicaremos ya a estudiar los conocimientos que proporciona la psicoacústi-ca sobre la percepción del sonido, entendiéndolo como un fenómeno global y complejo, en el que interactúan simultáneamente muchas frecuencias diferentes a intensidades diversas. Resumiendo: de acuerdo con lo que definíamos en el apar-tado 3.3.3, vamos a enfrentarnos a la sistematización de la *sensación tímbrica* y a lo que sabemos sobre sus umbrales.

Probablemente, la aproximación psicoacústica más completa que se ha desa-rrollado hasta ahora sobre como actúa la percepción humana frente a los sonidos compuestos complejos es la de E. Zwicker. Este investigador propone lo que él mismo denomina *diagramas de sonía* (Zwicker y Feldtkeller, 1981, págs. 182-196). Utilizando estos diagramas es posible traducir los datos sobre intensidad y frecuencia tomados por equipos convencionales de análisis de audio<sup>8</sup> en espec-tros que se ajustan de forma muy aproximada al modo en que el oído humano percibe la complejidad sonora. En el diagrama que Zwicker propone para el

7. Al hablar de *espectro dinámico* nos referimos a que la composición interna del sonido está su-friendo cambios importantes a medida que pasa el tiempo. Por ejemplo, una variación alternativa de 10 dB en la intensidad del 2º y el 4º formantes del espectro global del ruido cada 2 segundos.

8. Nos estamos refiriendo aquí tanto a analizadores de espectro por FFT, como a analizadores de es-pectro en fragmentos de octava.



*cálculo de la sonía* se relacionan simultáneamente la presión física en dB, la sensación subjetiva de intensidad en sones y fones, la sensación de intensidad en función del ancho de banda del sonido, el enmascaramiento de unos sonidos por otros en función de su intensidad y su frecuencia y, finalmente, la relación entre las distintas sensaciones subjetivas de intensidad que determina cada frecuencia fragmentando la zona de audición en tercios de octava.

Pero antes de explicar con detalle este diagrama para el cálculo de la sonía revisaremos dos fenómenos auditivos sobre los que se apoya, que no hemos estudiado todavía: la percepción en *bandas críticas* y el *enmascaramiento*.

4.3.1. *La percepción en bandas críticas*

En el apartado 4.2.3 de este mismo capítulo estudiábamos que la sensación global de intensidad está relacionada con el ancho de banda del sonido. Vimos que la sensación de intensidad subjetiva aumenta a medida que acumulamos frecuencias que suenan simultáneamente al sumarse auditivamente sus energías, y que el cálculo de esta sensación global de intensidad se hacía de acuerdo con unas curvas de ponderación. Este fenómeno tiene también unos umbrales a los que se denomina *bandas críticas*.

Pero estudiamos todo esto desde el principio. Se observa que la sensación de intensidad subjetiva siempre es mayor cuando escuchamos un sonido de cierta anchura de banda, centrado por ejemplo a 100 Hz, que cuando escuchamos un sonido puro de 100 Hz, a pesar de que ambos tengan exactamente la misma presión sonora. Pero esta sensación de que la intensidad aumenta a medida que aumenta el ancho de banda no es homogénea. En realidad, el aumento de la sensación subjetiva de intensidad no se produce hasta que la acumulación de componentes de frecuencia rebasa cierta amplitud de banda.

Pasemos a un ejemplo. Si tomamos un sonido puro de 1.000 Hz con una intensidad de 60 dB y, mientras lo escuchamos, vamos añadiendo progresivamente otros componentes de frecuencia de la misma intensidad, es decir:

|                 |                  |
|-----------------|------------------|
| sonido central: | 1.000 Hz (60 dB) |
| +               | 980 Hz (60 dB)   |
| +               | 990 Hz (60 dB)   |
| +               | 1.010 Hz (60 dB) |
| +               | 1.020 Hz (60 dB) |
| ...             | ...              |
| etc.,           |                  |

Se comprueba experimentalmente que hasta que los componentes de frecuencia acumulados no forman un sonido con un ancho de banda de 160 Hz la sensación subjetiva de intensidad permanece constante. A partir de esta amplitud de banda cada nuevo componente de frecuencia que se añade supone ya un aumento progresivo y homogéneo de la sensación de intensidad. Así pues, en torno

a los 1.000 Hz se observa que la amplitud de banda (o umbral diferencial) necesaria para que una acumulación de componentes sonoros sea percibida como aumento de intensidad es de 160 Hz.

Una forma mucho más cómoda de nombrar este fenómeno es diciendo que existe una anchura de *banda crítica* por debajo de la cual la sensación de intensidad permanece constante. La experimentación a lo largo de toda la superficie audible refleja que diferentes frecuencias centrales tienen *bandas críticas* distintas.

Fletcher detectó ya en 1940 que la percepción de los sonidos compuestos se organizaba en *bandas críticas* (Fletcher, 1961). Actualmente, se asume que la descomposición del espectro de frecuencia en *bandas críticas* es una de las características perceptivas fundamentales del oído. Parece ser que el oído percibe siempre ordenando el sonido en bandas críticas, con independencia del punto de la escala de frecuencias en el que se sitúa el sonido. Zwicker, tras una laboriosa experimentación sobre toda la superficie auditiva, propuso una división de la zona de frecuencias que va de los 20 Hz a los 16 kHz en 24 bandas críticas (Zwicker y Feldtkeller, 1981, pág. 72). Luego, al observarse que la anchura de las *bandas críticas* es relativamente próxima a los anchos de banda de 1/3 de octava, y que la variabilidad entre sujetos para percibir las bandas críticas pueden oscilar en torno al 20 %, o incluso más, la división del espectro propuesta en principio ha sido modificada de manera que se ajusta a los filtros convencionales normalizados de 1/3, 2/3 o 3/3 de octava. Esta división del espectro en bandas 1/3 de octava es la que se implementa en los diagramas propuestos por Zwicker para el cálculo de sonía, que estudiaremos un poco más adelante.

4.3.2. *El enmascaramiento*

El *enmascaramiento* es un fenómeno auditivo mucho más familiar y evidente que el de las *bandas críticas*. Se denomina *enmascaramiento* a la pérdida de sensación de presencia auditiva de un sonido determinado cuando aparece otro similar más fuerte que el primero. Dicho de otro modo: cuando un tono de intensidad fuerte esconde a otros de intensidad igual o menor que están a su alrededor, se dice que los está enmascarándo. Se observa fácilmente que cuanto más cercanos son los tonos de dos sonidos y más alta es la intensidad de uno de ellos, mayor es el enmascaramiento. Una experiencia muy habitual es, por ejemplo, la de estar hablando en un ambiente relativamente silencioso y, de repente, el terrible zumbido que proviene de una obra cercana hace que dejemos de entendernos. Cuando el zumbido es muy intenso parece, incluso, absorber como por arte de magia el sonido de nuestras voces, haciéndolas completamente inaudibles. En la primera situación se está produciendo el enmascaramiento de una parte de las frecuencias de nuestra voz, en la segunda, el espectro sonoro del zumbido es tan amplio y potente que enmascara las voces por completo.

La explicación del enmascaramiento, tal como ocurre con la de las *bandas críticas*, no es física sino fisiológica. Las investigaciones sobre el funcionamiento de



la membrana basilar del oído humano muestran que un sonido simple no excita sólo a un nervio concreto correspondiente al tono con el que se experimenta, sino que, además, excita todo un conjunto de nervios correspondientes a una gama de frecuencias que están alrededor del tono simple en cuestión, actuando éste, en realidad, como la frecuencia central de una banda más o menos ancha (véase Lancaster, 1973). Es decir, la membrana basilar reacciona igual frente a un tono puro que frente a un sonido compuesto por ese mismo tono y las frecuencias más próximas a él. Este funcionamiento del oído parece explicar tanto el fenómeno tonal como el enmascaramiento como el fenómeno de las bandas críticas.

Lógicamente, si se tiene en cuenta el fenómeno que acabamos de explicar, la representación gráfica de un sonido puro desde el punto de vista perceptivo no debe estar definido por una raya, sino por una superficie que tenga en cuenta la influencia de cada tono sobre la percepción de las frecuencias vecinas. Se habla, entonces, de *curvas de densidad específica* del sonido. Estas curvas fueron establecidas por Fletcher realizando numerosos experimentos sobre enmascaramiento (Fletcher, 1961), y han sido normalizadas y sistematizadas por Zwicker para cada una de las bandas de frecuencia (Zwicker, 1960).

### 4.3.3. Los diagramas de Zwicker

Hecha ya la revisión que proponíamos sobre la percepción en *bandas críticas* y sobre el *enmascaramiento*, pasaremos a desarrollar la explicación detallada de los diagramas para el *cálculo de la sonía*.

Más arriba, decíamos que Zwicker, partiendo de los conocimientos existentes sobre la organización perceptiva del sonido, ha desarrollado una experimentación sistemática sobre percepción psicoacústica en una parte muy amplia del espectro audible. Este trabajo le ha permitido integrar sobre un mismo diagrama experimental la sensación subjetiva de intensidad en dB, fones y sones, la sensación de intensidad en función del ancho de banda, el enmascaramiento y, finalmente, la relación entre las distintas sensaciones subjetivas de intensidad en función de la *densidad específica*<sup>9</sup> que tiene cada banda de frecuencia, ajustándose mucho al mecanismo de la audición compleja humana.

Pero veamos todo esto gráficamente en la figura 26. En esta figura se muestran dos de los diez diagramas preparados por Zwicker para el cálculo de los espectros de sonía. El diagrama superior es para calcular la sonía en *campos sonoros libres*, y el inferior para el cálculo en *campos sonoros difusos*.

Un *campo sonoro libre* es aquel que se genera al emitir sonido en un espacio en el que no existe ningún tipo de posibilidad de reflexión. En él el sonido llega al punto de medida desde una única dirección. Sería el caso de un lugar total-

9. El concepto de *densidad específica* hace referencia a la distinta sensación de intensidad subjetiva que producen los sonidos de la misma presión acústica, en función del lugar del espectro audible en el que estén situados. Dicho de otro modo, a igual presión sonora, se dice que un sonido tiene mayor *densidad específica* que otro cuando transmite una sensación de intensidad mayor.

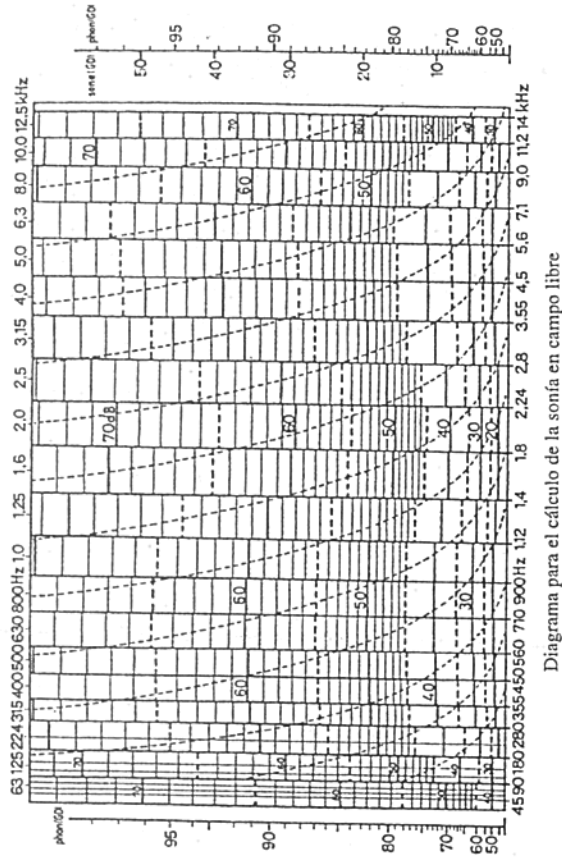


Diagrama para el cálculo de la sonía en campo libre

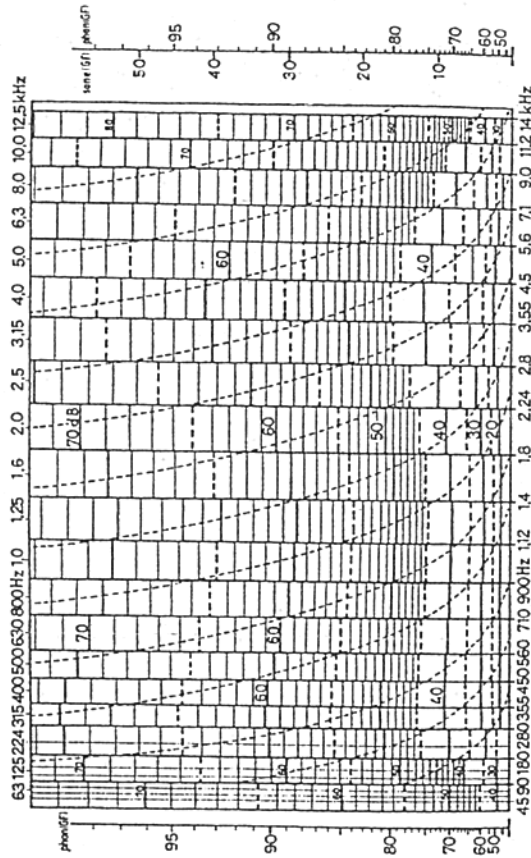


Diagrama para el cálculo de la sonía en campo difuso

FIGURA 26. Diagramas de Zwicker para el cálculo de la sonía. En esta ilustración sólo se muestran dos de los diez diagramas que son necesarios para representar todo el espectro audible: cinco para la percepción en campo libre y otros cinco para campo difuso.

mente abierto y con el suelo absorbente, por ejemplo, una pradera con la hierba crecida, una cima recién nevada, o una cámara anecoica. En este tipo de campo sonoro, la fuente actúa sin que al sonido emitido se le oponga ninguna vibración que modifique el espectro sonoro original.

Contrariamente, *un campo sonoro difuso* es el que se produce cuando el sonido proviene de todas partes, sin que haya ninguna dirección de propagación del sonido que predomine con claridad sobre las otras. En un *campo sonoro difuso* el sonido llega al punto de medida desde todas las direcciones a la vez. Esto ocurre en las cámaras reverberantes, es decir en las habitaciones cerradas con el suelo y las paredes lisas y duras. En los *campos difusos* el espectro original que produce la fuente sonora queda alterado al mezclarse con sus propias reflexiones.

Zwicker ha desarrollado un total de cinco diagramas para cada tipo de campo con el fin de que el resultado fuese gráficamente manejable. Cada uno de los cinco diagramas para *campo libre* y para *campo difuso* recoge una parte distinta de la gama audible de intensidades. Tomando como unidad de intensidad, por ejemplo, el dB, y como banda de referencia la de 900 Hz a 1.120 Hz, cada uno de los cinco diagramas para *campo libre* y para *campo difuso* recoge respectivamente la siguiente franja de intensidades:

|                           | <i>en campo libre</i> | <i>en campo difuso</i> |
|---------------------------|-----------------------|------------------------|
| 1 <sup>er</sup> diagrama: | de 5 a 37 dB          | de 5 a 34 dB           |
| 2 <sup>o</sup> diagrama:  | de 10 a 51 dB         | de 10 a 53 dB          |
| 3 <sup>er</sup> diagrama: | de 20 a 70 dB         | de 20 a 73 dB          |
| 4 <sup>o</sup> diagrama:  | de 40 a 92 dB         | de 40 a 89 dB          |
| 5 <sup>o</sup> diagrama:  | de 60 a 110 dB        | de 60 a 109 dB         |

Como el lector puede comprobar, los diagramas que se muestran en la figura 26 son los centrales (20 a 70 dB en *campo libre*, y 20 a 73 dB en *campo difuso*). Las diferencias entre uno y otro se deben a la influencia del campo acústico con o sin reflexiones, y pueden observarse si se comparan detenidamente las distribuciones relativas de las líneas.

Desarrollaremos a partir de ahora toda la explicación de los diagramas en singular puesto que, a pesar de que la distribución de los datos en cada uno de ellos es distinta, todos actúan exactamente del mismo modo y configuran, en realidad, un único gráfico representado en varias partes.

En el eje de abscisas (horizontal) del cuadro se representarán las frecuencias del espectro audible organizadas en tercios de octava (aproximadamente en bandas críticas). En las frecuencias más bajas los tercios de octava están representados por líneas verticales delgadas y las líneas gruesas muestran saltos de una octava. A partir de los 280 Hz (numeración horizontal inferior) las líneas gruesas verticales indican ya divisiones de 1/3 de octava. En el cuadro aparecen dos numeraciones horizontales, una en el lado horizontal inferior y otra en el superior. Ambas

expresan la misma escala de frecuencias, no obstante, las cifras del lado horizontal superior lo que indican son la frecuencia central de cada una de las bandas críticas. Ambas escalas son complementarias, y permiten situar en el diagrama cualquier sonido compuesto por frecuencias comprendidas entre 20 y 14.000 Hz (le una forma muy aproximada desde el punto de vista de la sensación auditiva).

En el interior del cuadro, en el eje de ordenadas (vertical) el lector habrá observado una serie de escalas distintas que fragmentan cada una de las bandas (le frecuencia en unas 35 divisiones de anchura irregular. Y habrá notado, también, que la distribución de estas divisiones irregulares es diferente para cada banda (le frecuencia. Esas escalas internas expresan la distinta sensación relativa de intensidad que producen los sonidos que tienen la misma presión sonora pero que están situados en un lugar distinto dentro del espectro de frecuencias audibles.

Desarrollaremos un ejemplo para explicar esto.

Imaginemos tres sonidos A, B y C: el sonido A tiene su tono central a 8.000 Hz y una anchura de banda de 1.900 Hz; el B está centrado tonalmente a 10.000 Hz, tiene un ancho de banda de 2.200 Hz; y, finalmente, el sonido C está situado tonalmente a 12.500 Hz y su amplitud de banda es de 2.800 Hz. Ahora, medimos estos sonidos con un sonómetro y los ajustamos de modo que los tres tengan la misma presión sonora: exactamente 65 dB. Experimentalmente, se comprobaba que estos tres sonidos no producen la misma sensación subjetiva de intensidad, sino que se perciben de modo que el sonido A resulta el más fuerte, el B se percibe como más débil que el A y, por fin, el C se oye como comparativamente bastante menos intenso que los dos anteriores. Y este fenómeno es completamente independiente de que la medida física en dB de las tres intensidades sea la misma.

Este tipo de fenómeno perceptivo en el que la frecuencia influye sobre la sensibilidad del oído a la intensidad, queda perfectamente recogido en las escalas internas verticales del diagrama de Zwicker que estamos revisando.

Véamoslo ahora gráficamente.

Las divisiones de las escalas internas de cada banda crítica están expresadas en decibelios (dB), cada división marcada por una línea delgada representa 1 dB, (le una línea gruesa a otra hay un salto de 10 dB, y entre una línea gruesa y una d s- continua el salto es de 5 dB.

Sinemos ahora los sonidos A, B y C, de nuestro ejemplo en el diagrama (le Zwicker y observemos el resultado. La forma de hacerlo será simplemente rellenando las franjas verticales que representan los distintos tercios de octava del m- do más aproximado posible a los anchos de banda mencionados, hasta la altura que indica su intensidad en dB. Es decir, tal como se representa en la figura 27.

Al dibujar los tres sonidos, podemos observar cómo la altura de cada una (le las representaciones es distinta, a pesar de que la presión sonora de los tres en (B es exactamente la misma. Éste es el modo en que el diagrama recoge el fenómeno de las distintas *densidades específicas* de cada sonido. O, dicho de otra manera, la distinta sensación subjetiva de intensidad que producen los sonidos que tienen la misma presión sonora pero que están situados en un lugar distinto dentro del espectro de frecuencias audibles. Así, utilizando los diagramas de sonía, podemos disponer de una «imagen perceptiva» de la presencia que va a tener cualquier

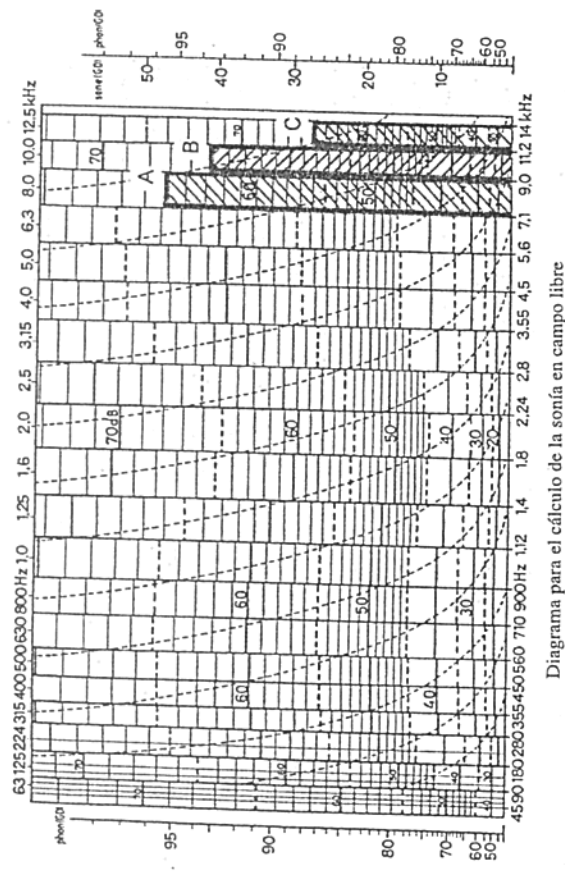


Diagrama para el cálculo de la sonoría en campo libre

FIGURA 27. Representación sobre el diagrama de Zwicker de la distinta sensación de intensidad que producirán sendos sonidos A (8.000 Hz), B (10.000 Hz) y C (12.500 Hz), a pesar de que la presión sonora de los tres sea exactamente la misma: 65 dB.

quier sonido en nuestro oído, y prever qué sonidos van a predominar sobre otros aplicando criterios estrictamente perceptivos.

Pasemos ahora a revisar otra de las posibilidades del diagrama. El lector notará que en el gráfico aparecen una serie de líneas discontinuas de forma curva que, recorriendo el gráfico de izquierda a derecha, descienden de arriba a abajo cruzando varias bandas críticas. Estas líneas expresan cómo actúa el fenómeno del enmascaramiento en función de la intensidad y la frecuencia del sonido enmascarador y de los sonidos enmascarados. Estas líneas de enmascaramiento nos permiten prever si un sonido va a enmascarar a otro, o no. Veamos un ejemplo.

Nos interesa saber si tres sonidos distintos (A, B y C) que tonalmente están bastante próximos entre sí, serán oídos claramente, o si se van a interferir entre ellos enmascarándose unos a otros. Los sonidos son los siguientes: A es un sonido compuesto, cuyo tono está centrado en 500 Hz, que tiene un ancho de banda de 110 Hz y una intensidad de 68 dB; B y C son dos tonos puros de 60 dB de intensidad que tienen unas frecuencias de 630 Hz y de 1.000 Hz, respectivamente. Ubiquemos ahora estos sonidos en el diagrama. El lector puede observar los tres sonidos representados en la figura 28.

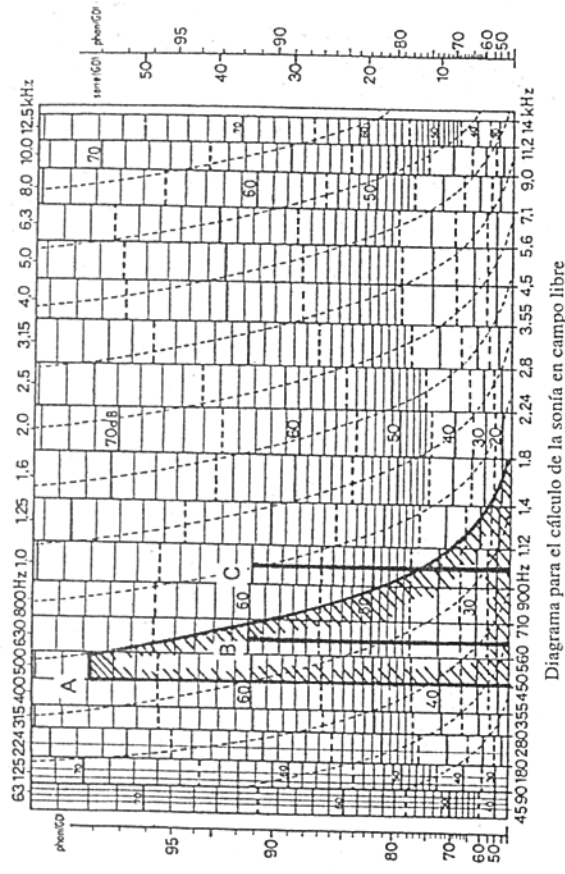


Diagrama para el cálculo de la sonoría en campo libre

FIGURA 28. Representación sobre el diagrama de Zwicker de un enmascaramiento acústico. El sonido A (500 Hz; 68 dB) enmascarará completamente al sonido B (630 Hz; 60 dB), pero no al C (1.000 Hz; 60 dB).

El sonido enmascarador será siempre el de frecuencia más baja. Para estudiar el enmascaramiento que producirá el sonido más grave se ha de considerar que este sonido (en este caso el sonido A) recubre toda la superficie contenida entre el espacio definido por la banda de frecuencias en cuestión y la línea discontinua (de enmascaramiento) que parte de esta misma banda de frecuencias. Toda la superficie acústica que queda recubierta con este trazo es la zona audible que enmascara el sonido A. Los otros dos sonidos (B y C), al ser puros, han sido representados con un único trazo grueso. En la figura 28 se ve con toda claridad cómo el sonido B queda totalmente incluido en el área de enmascaramiento mientras el C está en su mayor parte fuera de ella. Esto significa que el sonido B quedará completamente enmascarado por el A y, por tanto, que desde el punto de vista perceptivo podemos prescindir de él puesto que será inaudible. No obstante, el sonido C sí podrá ser escuchado sin ningún problema.

El sonido B si que sería audible si aumentásemos su intensidad en 7 u 8 dB, puesto que entonces escaparía ya al área de influencia del enmascaramiento.

Respecto a la estricta explicación del diagrama, sólo queda ya comentar la presencia a ambos lados del gráfico de sendas escalas de fones y una escala de sones a la derecha, cuyo sentido es el de permitir la comparación entre los valores de intensidad dados en dB o dados en fones o en sones. Y posibilitar la representación

sobre el diagrama de sonidos cuyos valores de intensidad vengan dados en fones o sores.<sup>10</sup>

Pasemos ahora a revisar la utilidad del diagrama de Zwicker como instrumento funcional aplicado a la comunicación audiovisual.

La aplicación de los diagramas de sonía tiene desde hace ya algunos años un papel importante en el universo de la ingeniería de las telecomunicaciones. El modelo de Zwicker sobre la percepción del sonido se está utilizando para abaratar las transmisiones de señal acústica. A partir del conocimiento de fenómenos perceptivos como el de la audición en *bandas críticas* y el del *enmascaramiento*, la ingeniería de tratamiento de la señal sonora procede a eliminar todo aquello que el oído humano no es capaz de percibir, ahorrando así una parte importante de la información a tratar, almacenar o transmitir y consiguiendo, en consecuencia, un abaratamiento importante del precio.

Pero toda la información múltiple y simultánea que proporciona el diagrama de Zwicker no sólo es útil a la economía de las telecomunicaciones, sino que pensamos puede y debe ser utilizado como un instrumento básico para la economía de la expresión acústica en su sentido más amplio y general. Desde el punto de vista expresivo, disponer de un modelo teórico que permita prever qué sonidos resultarán inaudibles al quedar enmascarados y qué sonidos no, y cuál es la presencia real que tiene cada sonido ante el oído humano ha de resolvernos muchos de los problemas que aparecen en la producción de los diseños sonoros complejos.

Pensemos, por ejemplo, en la sonorización cinematográfica de una batalla. En este tipo de escenas auditivas es frecuente acumular y superponer numerosos *objetos sonoros* caros y de difícil producción. Y es frecuente, también, que una buena parte de los *objetos sonoros* que componen el escenario auditivo de la batalla queden completamente disueltos (enmascarados) en una mezcla múltiple y confusa que, en realidad, los hace completamente inaudibles. La aplicación racional del diagrama de Zwicker en este tipo de producciones sonoras como un instrumento de apoyo técnico-narrativo puede suponer un ahorro importante de tiempo y esfuerzo, y un fuerte incremento del rendimiento expresivo en el resultado final de la mezcla. La selección y la organización sonora de todo el material que ha de componer un decorado acústico suele depender exclusivamente de la intuición y el oído del realizador. No obstante, este trabajo puede y debe estar apoyado en instrumentos como el que estamos estudiando, porque al manejarlos se están teniendo en cuenta los mecanismos esenciales que organizan la percepción sonora del ser humano.

Cuando se trabaja, por ejemplo, en sonorizar una persecución, resultaría muy útil saber exactamente a qué diferencia tonal ha de estar la melodía que subraya la acción dramática, del tono medio de cada uno de los efectos sonoros (motores de coche, frenazos, ruidos urbanos, gritos), para que el ruido no enmascare la música. De este modo, sería posible conseguir con rapidez y eficacia que ambos sis-

temas sonoros (música y efectos) fuesen audibles llegando a una mezcla equilibrada y natural.

Los efectos sonoros son a menudo sonidos compuestos altamente enmascaradores. Normalmente, para evitar el enmascaramiento, por ejemplo, de la melodía, o de otros efectos con menor potencia sonora (ruidos de pasos, crepitar del fuego, murmullo de hojas, etc.) se tiende a potenciar los más débiles. Si uno de los sonidos no se oye porque queda enmascarado, simplemente se aumenta su intensidad. El resultado es un efecto atronador y antinatural en el que sonidos que normalmente son de muy poca intensidad estallan en la sala de cine o en el altavoz del televisor como auténticos cañonazos. El trabajo sobre las composiciones sonoras utilizando los diagramas de sonía supondría disponer de un auténtico mapa sobre las posibilidades humanas de audición. Un mapa que nos permitiera saber por dónde hemos de movernos acústicamente para desarrollar el máximo rendimiento narrativo con el mínimo esfuerzo. Diseñando adecuadamente la ubicación de los sonidos simultáneos sobre la superficie auditiva, podemos evitar el enmascaramiento sin necesidad de aumentar la intensidad, sin necesidad de que el sonido global resultante sea atronador, consiguiendo así una mezcla mucho más próxima a las intensidades naturales de nuestro entorno cotidiano.

Con frecuencia escuchamos en las obras audiovisuales, sonidos supuestamente delicados y volupuosos como el «fru-fru» de las ropas, pasos en ambientes silenciosos, murmullos de agua... Este tipo de sonidos, que deberían ser tenues y sutiles, suelen estar tratados con unas intensidades tan fuertes que resultan completamente antinaturales. Para evitar esto, resultaría muy eficaz determinar con precisión cuál ha de ser la composición espectral y la intensidad mínima imprescindible para superar los 40 dB de frecuencias bajas que produce el ambiente de la sala donde se ve la televisión o se proyecta la película, evitando que el almidón de las ropas siga sonando como 50 kilos de cartón piedra aplastados por una api-sonadora.

10. El lector puede encontrar en el apartado 3.4.2 una explicación detallada sobre estas unidades.

## CAPÍTULO 5

### LAS FORMAS DEL SONIDO

*La lógica de desarrollo de este apartado será ir revisando, de menor a mayor complejidad, todo aquello que sabemos sobre la organización auditiva del sonido hasta llegar a establecer una primera taxonomía de las formas sonoras fundamentales. Así, en este capítulo, se intentan localizar las piezas fundamentales del rompecabezas acústico con el que construimos el sentido sonoro.*



## 5.1. La percepción de formas sonoras

Hasta aquí hemos estado estudiando la *sustancia sonora* dentro del lenguaje audiovisual en el sentido más hjelmsleviano del término (véase Hjelmslev, 1980). Hemos ido analizando y estudiando cómo es la «pintura sonora», cómo está compuesta, qué tipo de «color auditivo» transmiten al oyente los distintos «colores sonoros» básicos. Por qué razón percibimos de manera distinta una mezcla u otra. De qué modo podemos manipular la intensidad del material sonoro puro. Cómo influyen las intensidades en el «color» final que da la mezcla... Y hemos ido estudiando sistemáticamente este análisis en el marco de la narración audiovisual. Ahora estudiaremos el siguiente nivel expresivo, el nivel de las formas. Es decir, el nivel en el que la sustancia sonora se organiza en modos y distribuciones concretas, componiendo formas, para que a estas formas se les pueda asignar sentido).

Es cierto que la percepción de formas es en gran parte cultural y, por tanto, aprendida. Pero también lo es que responde y se organiza sobre los cientos de sonidos con las leyes físicas que rigen el sonido, y con las leyes fisiológicas y psicológicas que rigen la percepción auditiva. Sin duda, la percepción de formas sonoras rebasa la organización física de la *sustancia sonora* y los mecanismos primarios de funcionamiento del oído humano, pero, a la vez, depende rigidamente de ambos.

El reconocimiento de las formas sonoras se organiza en grandes sistemas.

Unos sistemas los descubrimos por la simple observación de fenómenos distintos que suelen estar próximos en el espacio y ser sincrónicos en el tiempo. Y otros son el resultado de un largo y complejo aprendizaje. El primer tipo de sistemas sería el que organiza la identificación de formas sonoras como la del ruido de un motor o la del crepitar del fuego. El dominio del habla y la competencia musical serían, en cambio, dos ejemplos paradigmáticos del segundo tipo de sistemas sonoros, los que son altamente dependientes de la educación.

Ciertamente, todos estos sistemas determinan que un conjunto de formas sonoras pasen a ser formas-patrón. Son las formas sonoras que al ser reiteradamente identificadas y asociadas a otros fenómenos concretos adquieren valor de modelo sonoro de referencia y valor de signo. Así, en la memoria auditiva se van acumulando formas-patrón y constituyendo un amplio muestrario que orientará nuestra manera de escuchar, organizar, seleccionar e interpretar las mezclas de sonidos que provienen de nuestro entorno. Finalmente, el uso reiterado de estos patrones hará que el ser humano tienda con mucha fuerza a la percepción categorial, es decir, a escuchar sólo aquello que busca, o aquello que está acostumbrado a reconocer habitualmente, a menudo, incluso, ignorando todo lo demás.

No obstante, como fenómeno anterior, y a la vez paralelo, a la percepción categorial de las formas sonoras, han de actuar necesariamente la propia física del sonido y los mecanismos perceptivos primarios. Pensamos, en consecuencia, que

En el proceso de identificación de las formas sonoras existe una etapa previa a la percepción de categorías sonoras y a la construcción de sentido que esto comporta. No nos estamos refiriendo ahora al reconocimiento de formas sonoras como al sinónimo de identificación, sino a su percepción misma. Intentaremos aclarar esto.

Para poder reconocer una forma sonora, es necesario haberle asignado antes un sentido. Haber experimentado que esa forma es relevante. Ese valor de relevancia puede haber sido otorgado por varias vías: la repetición reiterada de la misma forma sonora, el impacto emocional de la situación a la que está asociada, el valor que supone identificar esa forma para la supervivencia del oyente, etc. Como consecuencia de esa experiencia, la forma en cuestión se almacenará en nuestra memoria como un patrón sonoro. Pero antes de esta etapa ha sido necesaria la percepción estricta de la forma. Ha sido necesario poder percibirla antes de poder reconocerla.

Estamos, pues, diferenciando entre la *percepción* de formas sonoras y el *reconocimiento* de formas sonoras. La percepción es previa al reconocimiento. Percibir formas sonoras supone experimentar sensaciones subjetivas asociadas a estímulos acústicos complejos; antes de asignar a estas formas algún sentido o algún valor expresivo. Vimos ya algo de esto cuando estudiábamos la sensación tímbrica (véase apartado 3.3.3). Contrariamente, reconocer formas sonoras es identificar sensaciones auditivas ya conocidas a las cuales están asociadas determinadas relaciones de contigüidad o de sentido.

Esta diferencia entre la *percepción* y *reconocimiento* de las formas sonoras no es baladí. El narrador audiovisual tiene en sus manos una serie de instrumentos de producción que le permiten tratar los objetos sonoros, *acusmatizándolos* y alterando su forma. De modo que estos *objetos sonoros* no sólo dejan de estar asociados a su fuente original, sino que, además, pasan a ser identificados como algo que no tiene nada que ver con su origen real. Piense el lector, por ejemplo, en los efectos sonoros sintetizados que se utilizan normalmente en el cine de ciencia-ficción. Los zumbidos vibrantes de las armas futuristas y el sonido vertiginoso de las aves que huyen hacia el hiperespacio han sido originados, normalmente, en el traido de un sintetizador y provienen físicamente de altavoces colgados en las paredes de la sala de proyección. ¿Qué es lo que se está reconociendo entonces? No existe un referente objetivo, ni las armas ni las naves existen; en todo caso, reconocemos algo que hemos oído ya en alguna otra narración audiovisual del mismo género. No obstante, lo que nos interesa es la primera vez, el primer sonido inventado para contar acústicamente los ruidos de las armas del futuro y el salto al hiperespacio. Para el narrador que tiene capacidad de diseñar nuevos sonidos, lo importante no es la identificación de una forma sonora familiar, sino el saber cómo ha de ser la forma misma de aquello que crea, para que pueda ser interpretado por el receptor tal como él quiere.

Para crear un sonido podemos seguir dos estrategias distintas. La primera es, simplemente, imitar, copiar otras formas sonoras que sabemos reconocibles. La segunda es articular una composición de distintos rasgos parciales de estas mismas formas que, antes, cuando estaban completas, eran identificables, pero que

ahora, al ensamblar una forma nueva a partir de fragmentos de otras, dejará de serlo. Para hacer esto necesitamos diferenciar entre *percepción* y *reconocimiento*, porque cada uno de los rasgos acústicos parciales que ahora serán percibidos como formas específicas, ya no configuran un patrón sonoro almacenado en nuestra memoria. Ahora componen un suceso sonoro nuevo al que hay que asignar sentido, y ese sentido será el que nos proponga el narrador audiovisual con una nueva asociación sonido-imagen o sonido-concepto. Recurriendo a la teoría que comenzamos a desarrollar en el apartado 2.3, el dominio de las formas sonoras, es decir la capacidad para percibir y manipular formas sonoras con independencia de los patrones almacenados previamente en nuestra memoria auditiva, es lo que permite la creación audiovisual de *entes acústicos* nuevos.

Cuando disponemos de capacidad para crear formas sonoras nuevas, lo que nos interesa saber es qué características sonoras son las que el oyente va a considerar como relevantes y por qué. Del sonido nos va a interesar su carácter diferenciador, su forma misma; y del receptor, la percepción estricta que va a tener de ella, porque el sentido también lo podremos crear nosotros. Así, partiendo de toda la reflexión anterior, definiremos *forma sonora* como:

Toda configuración acústica que, aun siendo analizable en dimensiones más simples, tiende a ser percibida como un bloque sonoro unitario y coherente.

De acuerdo con este planteamiento, en este apartado y los siguientes, desarrollaremos una primera revisión sobre la organización de las formas sonoras, antes de adentrarnos en el fenómeno estricto de la producción de sentido a partir del sonido. Es decir, lo que nos va a ocupar ahora será la observación de las formas más elementales de agrupamiento y organización de las distintas dimensiones de los sonidos complejos; y el tipo de percepciones auditivas que estas organizaciones sonoras desencadenan en el receptor. Probablemente las formas sonoras simples más familiares y mejor clasificadas en la cultura occidental sean el *ruido* y el *silencio*. En consecuencia con esta reflexión, consideramos que nuestro estudio de las formas sonoras debe comenzar revisando estos dos conceptos.

### 5.1.1. El ruido

El concepto de *ruido* en el ámbito de la comunicación audiovisual es tan extremadamente familiar como poco concreto. Se utiliza la palabra «ruido» para nombrar algo tan vago como «cualquier sonido no deseado»; para hacer referencia a los *efectos sonoros*, es decir, a los sonidos del ámbito audiovisual que no son de origen verbal ni musical, por ejemplo, el timbre de un teléfono o el crujido de una puerta; a menudo la palabra «ruido» se asimila, también, al concepto de *sonido*, es decir, a cualquier vibración perceptible por el oído; otras veces se entiende por «ruido» todo aquello que dificulta un proceso de comunicación.

Evidentemente, la palabra «ruido» aporta, entonces, muy poca información para el narrador audiovisual. No obstante, tras el concepto de *ruido* subyacen formas