

## EL TIMBRE: UN ACERCAMIENTO DESDE VARIAS DISCIPLINAS

Anita Arrieta\*

### ABSTRACT

This article considers the timbre perception from the perspectives of physics, psychophysics, and speech perception, in order to attempt a complementary vision that can enrich our position concerning audition.

**Key words:** perception, phonology, timbre.

### RESUMEN

En el artículo se considera la percepción del timbre y se caracteriza desde los enfoques de la física, la psicofísica y de la percepción del habla, para intentar una visión complementaria que pueda enriquecer nuestras posiciones frente al campo de la audición.

**Palabras clave:** percepción, fonología, timbre.

### Introducción

Entre los atributos de la sensación auditiva se encuentran la percepción de la intensidad, de la altura, del timbre y de la localización espacial. En este artículo, nos vamos a referir únicamente a la percepción del timbre. Caracterizaremos este término desde diferentes disciplinas que se han preocupado por abordarlo: la física, la psicofísica, la percepción del habla y la psicología, con el fin de obtener una visión de conjunto de esta categoría.

Nos interesa este concepto porque, en apariencia, la noción de timbre es simple: «modo propio y característico de sonar un instrumento o la voz de una persona». Sin embargo, consideramos que el término abarca un terreno mucho más complejo, debido al papel significativo que éste juega en la vida de los seres humanos.

### 1. El timbre desde la física

En Física, «timbre» se refiere a una cualidad de la percepción de la onda sonora y no a una cualidad física de la onda. La mayoría de las ondas que escuchamos son ondas complejas filtradas, es decir, que la mayor parte de las impresiones sonoras que percibimos se deben al efecto de algunos fenómenos, más o menos simples o complejos, de resonancia sobre una onada compleja. El efecto de resonancia se refiere al hecho de que la energía se concentra en ciertas bandas de frecuencias que comprenden los componentes de la onda compleja. Esta distribución de la energía en el espectro de frecuencias es el hecho físico cuya correlación perceptiva se considera como un timbre. Por ejemplo, el timbre de la vocal [a] se reconoce cuando una onda compleja concentra una energía elevada de 750 hz a 1500 hz.

\* Profesora de la Escuela de Filología, Lingüística y Literatura, Universidad de Costa Rica.

La nota musical «fa», emitida por un piano, posee la misma concentración de energía, y los mismo armónicos, que esta misma nota emitida por un trombón. Lo que distingue a estos dos sonidos diferentes es que los armónicos, que tienen mucha o poca energía, son diferentes, es decir que la energía se distribuye de una manera diferente en el espectro de frecuencias.

Lo anterior explica porqué tales timbres son diferentes. Esta distribución de la energía equivalente al hecho físico del timbre es observable si se analizan las señales con la ayuda del espectrograma, aparato que descompone la señal y muestra su estructura en algunas muestras llamadas espectrogramas.

Dos instrumentos pueden tocar la misma nota (en el mismo tono), pero se reconoce sin dificultad si se trata de un piano o de un trombón. Lo que diferencia el sonido de un piano y el de un trombón, y lo que distingue entre una vocal [a] y la vocal [e], pronunciadas por la misma persona, se denomina timbre.

### 1.1. Los sonidos formantes

La diferencia entre los timbres es comprensible si se examinan las características de los sonidos formantes y de los ruidos.

Un formante es un armónico o conjunto de armónicos que conserva una energía elevada porque su frecuencia coincide con la de un resonador. Si se dice que un sonido tiene un formante de 500 hz, significa que este sonido va a concentrar su energía alrededor de esta frecuencia.

Los sonidos formantes son siempre sonidos de una fuerte energía armónica, con una energía fundamental bien establecida y con armónicos más o menos claros. El timbre «formántico» caracteriza a los sonidos que tienen formantes. Es interesante subrayar, de esta estructura formántica, que la energía del sonido se concentra en bandas de frecuencia bien definidas que contrastan con las zonas de frecuencia donde la energía es baja. Toda la energía de la onda se concentra en bandas más o menos estrechas y las zonas de frecuencia que no están favorecidas por el resonador presentan una energía prácticamente nula. Esto

provoca que los timbres sean altamente cromáticos. Por lo tanto, se pueden percibir claramente los matices. La mayoría de los sonidos formantes presentan más de un formante porque varias veces el resonador es complejo y existen numerosas barras del espectro de frecuencia que van a ser reconocidas en el proceso de resonancia.

Los sonidos del habla que presentan una estructura formántica más clara son las vocales. Incluso cuando las muestras espectrográficas localizan claramente la frecuencia precisa de los formantes de una vocal determinada, la frecuencia que corresponde a un formante está determinada por el tamaño y la forma del resonador que filtra la señal de la fuente. En el caso de las vocales, la señal se genera en la glotis y el conducto formado por la faringe y la boca es el resonador. El gesto articulador que se puede realizar con la lengua, los labios y el maxilar, va a determinar el tamaño de este tubo y de su forma y, al mismo tiempo, las frecuencias que se reforzarán. Por ejemplo, si el tamaño y la forma del resonador están preparados para una vibración de 500 hz, 1500 hz y 2500 hz (configuración de la vocal schwa), debemos esperar que los armónicos puedan tener una energía elevada y los tres formantes puedan definirse. Pero esto se presenta solamente si existen armónicos en estas frecuencias.

En el sentido estricto, los formantes no son una característica de la señal, sino del resonador: es el resonador el que posee frecuencias formantes, a las cuales éste tiende a vibrar espontáneamente si se le perturba. El resonador condiciona la forma de la onda que incide sobre él.

Cuando se dice que un sonido posee formantes entre 500 hz y 1500 hz, se quiere expresar que se trata de una onda filtrada por un resonador cuyas frecuencias formantes son de 500 hz y 1500 hz.

## 2. El timbre y la psicofísica

La psicofísica ha aportado la formalización matemática para la elaboración de diferentes métodos de medición utilizados por otras disciplinas como la psicología. En el caso de la psicofísica perceptiva, existe un interés por estudiar los efectos

del contexto en la determinación de la respuesta y por considerar las diferencias individuales.

En este campo, se encuentra la teoría de la detección de la señal, tendencia que se originó en las investigaciones aplicadas realizadas en Estados Unidos durante la Segunda Guerra Mundial y después de ella. La aplicación de esta teoría se relaciona con el concepto de timbre.

La importancia de la teoría de la señal radica en que, para estos investigadores, es evidente que el individuo que realiza una observación sensorial toma, al mismo tiempo, una decisión. El sujeto se comporta como un estadístico que se pregunta si la observación que acaba de hacer es más representativa por la presencia de la señal por detectar o por el ruido que la acompaña. Este ruido puede ser un ruido exterior causado por el contexto, o un ruido producido por la actividad bioeléctrica espontánea de las neuronas o de ambos al mismo tiempo. La teoría de la detección de la señal permite que se desarrolle un modelo de discriminación al mismo tiempo que un modelo de decisión. Esta teoría ha suscitado numerosas investigaciones y ha sido aplicada a una gran variedad de dominios de la psicología: situaciones de vigilancia y de control y situaciones de reconocimiento mnémico. Pero, la contribución principal de esta teoría ha sido, sobre todo, cuestionar el concepto de umbral sensorial. La teoría ha permitido también explicar los procesos por los cuales los factores de aprendizaje son capaces de actuar sobre la respuesta sensorial modificando el criterio de decisión del sujeto.

El objetivo de los datos anteriores, como ya se mencionó, es intentar un vínculo entre el tema del timbre y la psicofísica.

Se puede considerar que las teorías de la psicofísica pueden aplicarse para resolver ciertos problemas de la vida cotidiana relacionadas con el reconocimiento y la medición de un sonido y de su timbre. Esta relación es importante para la audiometría tonal de los diapasones, la audiometría radioeléctrica, los tests fonográficos y la audiometría verbotonál.

En general, los exámenes de la audición se fundan en la utilización de métodos psicofísicos «subjetivos», que estudian la sensibilidad de la recepción auditiva de los sonidos artificiales

(audiometría tonal: reloj, diapasón, audiometría electrónica), y del lenguaje articulado no filtrado (audiometría vocal) o filtrada (audiometría verbotonál). Los métodos «objetivos» emplean los indicios observables de la sensación (respuestas de atención, bloqueo del ritmo electroencefalográfico). Estas técnicas permiten definir los umbrales auditivos «normales», detectar los diferentes tipos de sordera o precisar las fallas acústicas de los aparatos de transmisión y de recepción de los sonidos.

### 3. El timbre y la percepción del habla

Las investigaciones en el campo de la percepción del habla son relativamente recientes. El estudio de las propiedades físicas del sonido fueron realizadas por Willis (1892) y Hemboltz (1859) en el siglo XIX. Sin embargo, las investigaciones específicas en el campo de la percepción del habla se desarrollaron en mayor escala durante la Segunda Guerra Mundial. La mayoría de las investigaciones en este campo necesitaron equipos para el análisis y para la síntesis del habla (simulación del habla por computadora).

La primera máquina fue desarrollada por Homer Dudley (1936-1939), de Bell Telephone Laboratories. Esta máquina llamada «vocoder» (de voice and coder), fue creada para transmitir de manera «eficiente» las señales del habla a lo largo de los circuitos telefónicos.

La función de la máquina era analizar y grabar el habla en una señal simple que contenía menos información para ser transmitida. A partir de estas investigaciones, Dudley y colaboradores observaron que el habla natural contenía una información redundante (especificada de múltiples formas). Esto quiere decir que varios factores tienen la posibilidad de ayudar en el reconocimiento de los sonidos del habla. Uno de los problemas del vocoder era su «acento eléctrico», pero los principios utilizados en su elaboración facilitaron el desarrollo de la elaboración del espectrógrafo del sonido. Este instrumento analiza las señales audio según la distribución de la frecuencia del sonido comprendido en la señal.

Para el análisis del habla, el espectrógrafo presenta la frecuencia en el eje y, el tiempo en el eje x, y la amplitud, por la oscuridad de los trazos.

Los análisis del habla a partir de la utilización del espectrógrafo existían durante la segunda Guerra Mundial, pero se guardaron como «secreto militar» hasta el final de la guerra. La figura generada por el espectrógrafo se denomina espectrograma del sonido. Esta brinda un despliegue estacionario de la señal acústica que permite evaluar cuáles aspectos de la señal son importantes para la percepción de los segmentos del habla.

Las raíces de la investigación de la percepción del habla han estado, lamentablemente, ligados a los intereses militares y comerciales por desarrollar «un mejor sistema de comunicación».

Actualmente, el trabajo de análisis del habla se realiza por computadoras especializadas que analizan y recrean con una resolución óptima la señal acústica.

En términos de las propiedades acústicas, el habla humana es una señal muy compleja que contiene muchos tipos de información en un momento individual y variable en el tiempo. La señal es rápida, compleja y continua.

Por esta razón, los aportes de las investigaciones en el campo de la percepción del habla son sobre todo demostraciones de los mecanismos utilizados para segmentar y reconocer el habla. Típicamente, los estudios han utilizado las sílabas o las palabras individuales como material experimental para segmentar.

### 3.1. Medir la voz

La voz es perfectamente analizable desde el punto de vista físico: la señal acústica del habla. El análisis de esta señal toma en cuenta diversos parámetros constitutivos: la frecuencia fundamental, la composición armónica (espectro acústico), la amplitud, así como el parámetro tiempo. Estos parámetros físicos corresponden a los fenómenos percibidos auditivamente: la frecuencia fundamental determina la impresión de altura (voz grave o aguda), la composición armónica corresponde al timbre (más o menos rico), la amplitud a la intensidad. El parámetro tiempo interviene en diversas

etapas y permite apreciar la velocidad de elocución de los sujetos.

La evolución del timbre en la duración es a menudo más estable. Algunas características están ligadas a la morfología del sujeto (tamaño de la laringe y de la cavidad bucal), pero ciertos «efectos» se pueden lograr, así como un cantante hace variar la calidad del sonido emitido. La tensión de los músculos de la laringe puede variar según el estado emotivo de un sujeto y puede también modularse en función de las necesidades de la comunicación.

### 3.2. Propiedades acústicas de la señal del habla y la noción de timbre

La fuente de energía acústica para la producción de sonidos procede de la modulación de la corriente de aire que circula desde los pulmones hasta los labios. La producción de un sonido implica el movimiento de moléculas de aire. Durante la producción del habla, algunas partículas del tracto vocal limitan la corriente de aire.

Según la manera como se produce un sonido, la lingüística ha establecido una serie de rasgos distintivos para cada sonido, de modo que puedan ser estudiados. Es aquí donde aparece la noción de timbre. Es principalmente en la descripción acústica de las vocales que se ha estudiado el timbre. Todos los sonidos del habla están compuestos de sonidos complejos (complex waves sounds). Esto quiere decir que contienen frecuencias diferentes de una manera simultánea. Las vocales resultan de una modificación de la fuente del sonido en la manera como éste está determinado por las características resonantes de la cavidad oral del tracto vocal durante la producción de este sonido.

Por lo tanto, un sonido específico está compuesto de la frecuencia de resonancia asociada al tamaño y a la forma del espacio por el cual circula el aire.

En el caso de las vocales, el componente de frecuencia de la fuente de estos sonidos en la glotis incluye la frecuencia fundamental (FO) y los múltiplos de la frecuencia fundamental.

Los múltiplos de FO se denominan armónicos. Por ejemplo, si la FO es de 100 hz,

los armónicos pueden ser de 200 hz, 300 hz, 500 hz y así sucesivamente hasta llegar a 8000 hz. Algunas de estas frecuencias tendrán mayor o menor amplitud. Por ejemplo, la forma de la cavidad bucal para producir la vocal [u] va a determinar la frecuencia que será prominente cuando la corriente de aire circule por intermediario de la forma particular de una cavidad oral.

La forma de la cavidad oral para la vocal [u] permite que las bandas de frecuencia baja puedan resonar. Y las bandas de frecuencia de resonancia para el habla cambian según los movimientos de los articuladores durante la producción del habla. Estas bandas se denominan formantes. La forma de la cavidad oral cambia según la vocal pronunciada.

En un espectrograma, se pueden apreciar los formantes de las vocales. Cada vocal se caracteriza al menos por tres bandas horizontales. Estas bandas son los formantes de las vocales. Por lo tanto, las vocales se diferencian por la posición relativa de los primeros dos formantes: F1 frecuencia baja y F2 frecuencia elevada es el patrón que caracteriza a una vocal como [i]; mientras que, para la identificación de la vocal [u], se distinguen dos formantes muy bajos de frecuencia.

El tamaño y la forma del tracto vocal influyen en los componentes de la frecuencia. Por esta razón, los hombres, las mujeres y los niños asignan valores diferentes para la formación de vocales.

Si los oyentes reconocen las diferentes vocales pronunciadas por un hombre, una mujer o un niño, es porque existe un sistema de reconocimiento que opera, y no porque exista una referencia a valores absolutos.

#### 4. El timbre y los estudios de psicología

El timbre ha sido estudiado, en psicología, principalmente por la psicología social. Es sobre todo en las investigaciones que se interesan por la voz humana y por su influencia en las relaciones sociales donde el timbre ha sido abordado. Sin embargo, el timbre se presenta, en estos estudios, como una característica de la voz, y no ha sido considerado de manera aislada.

En este contexto, los estudios de psicología social plantean la gran importancia que tiene, para las relaciones sociales más elementales, la habilidad para modificar de una manera apropiada el conjunto de conductas verbales y no verbales para tratar de influir sobre las actitudes y comportamientos de los otros.

Estas investigaciones se centran en el tono y la altura de la voz, así como en el número y la localización de las pausas y la naturaleza de las líneas de entonación como elementos que favorecen el contacto social. Estos estudios siguen las metodologías que se detallan a continuación:

a) *Manipulación de la voz y de los tipos del habla («cue manipulation approach»)*

Estas propuestas presentan una muestra de habla (generalmente persuasiva) a locutores nativos para obtener «ratings» de las fuentes de credibilidad. Esta estrategia ha sido muy popular en las investigaciones de psicología social. Se manipulan series enteras de claves vocales (acentos, calidad de la voz...). Algunos estudios en esta dirección son los de Lambert *et al.*, 1960; Giles y Powesland, 1975.

Las variables dependientes utilizadas en estas investigaciones están ligadas, en su mayoría, con las escalas que pertenecen a dimensiones mayores: dominio, dinamismo, bienestar y honestidad. Los resultados de tales estudios tienden a mostrar que los acentos estándares, el dinamismo, la velocidad de la voz y una falta relativa de pausas y de repeticiones, ligada a una calidad de la voz «normal», provoca resultados elevados de competencia, dominio y dinamismo.

b) *Codificación de una fuente específica «via role playing»*

Este tipo de investigación ha sido muy popular en los estudios de expresión de las emociones (Kramer 1963; Scherer 1970). Se han realizado investigaciones para analizar cómo los sujetos interpretaban las diferentes maneras que los actores utilizan para expresar sentimientos, ideas. Lamentablemente, los resultados

han estado influenciados por el desempeño de los actores. Algunas veces ellos empleaban formas muy tradicionales para expresar ciertas emociones y estados psicológicos. No obstante, estas investigaciones han permitido establecer relaciones entre ciertas características y el efecto obtenido en los espectadores. Por ejemplo, en una de los experimentos, por la voz que un actor debía utilizar, los espectadores interpretaron que se trataba de un hombre «que no había tenido suerte en la vida, que no tenía trabajo...» y otros estereotipos sobre el éxito social.

- c) *El enfoque de la interacción*  
Este enfoque dirige la discusión entre dos personas o entre ciertos grupos en el laboratorio para tratar de detectar la influencia de uno o de varios participantes sobre los otros. Estas interacciones se graban para poder analizar «objetivamente» los comportamientos de los sujetos. La información visual y auditiva se les presenta posteriormente a los participantes que tenían una influencia significativa en la situación. A pesar del interés de tales investigaciones, los experimentos requieren de mucho tiempo y dinero, por lo que los resultados no son generalizables a gran escala.
- d) *Otros tipos de análisis estudiados*  
Otros estudios analizan la medida de las pausas y la importancia de las variaciones melódicas gracias a la informática y a sus posibilidades de cuantificar. Los trabajos de Addington (1969), del equipo de Brown y sus colaboradores (1973, 1974) y las observaciones de Chevie - Muller y otros (1978), han aclarado el papel de los diversos parámetros acústicos de la percepción de la personalidad. Por ejemplo, plantean que la calidad de la voz (la riqueza del timbre), refleja la competencia y el dinamismo. Esta calidad se presenta como uno de los elementos que atestiguan, para el oyente, el deseo de contacto y de comunicación de su interlocutor.

Otros estudios psicológicos se han preocupado por las consecuencias, para la apreciación musical, que la afasia severa provoca en los pacientes. Estas investigaciones han dirigido su atención principalmente a la reproducción melódica y al reconocimiento de timbres diferentes. Han propuesto que, debido a la preservación frecuente de la reproducción y de la apreciación musical en los pacientes que sufren de una afasia profunda, la habilidad musical se controla en el hemisferio derecho. Por ejemplo, Milner (1962) utilizó los tests de aptitud musical para evaluar pre y postoperatoriamente a un grupo de 16 pacientes. Encontró que solamente los pacientes que tenían una lesión en el hemisferio derecho mostraban, después de la cirugía, un descenso en la realización de la tarea en cuestión. Estas situaciones se evidenciaban en las tareas de juicios («igual o diferentes»), de timbres y de memoria tonal.

No obstante, estudios más recientes (Miceli, 1982; Auerbach, Allard, Naeser, Alexander y Albert, 1982) han mostrado como evidencia clínica que el déficit denominado amusia se presenta también si la lesión ocurre en el hemisferio izquierdo.

## 5. Conclusiones

Después de haber presentado la noción de timbre desde diversos enfoques, concluimos que el estudio de este concepto, generalmente reducido a sus componentes acústicos y físicos, también está relacionado con significados de tipo social y cultural que, lamentablemente, todavía se encuentran en su etapa inicial, así como también lo está la investigación de este término y sus relaciones con la organización cerebral.

No obstante, esbozar algunos de los lineamientos que el estudio del timbre ha tomado según las disciplinas abordadas, evidencia la dificultad de reducir el término a una sola concepción o a una única interpretación. Más bien, el recorrido planteado remite a la complejidad y a la interacción que los diversos componentes

del sonido mantienen unos con otros y, por ende, a la dificultad de aislar el concepto de timbre.

### **Bibliografía**

- Alès, Catherine. 1990. *L'esprit des voix. Etudes sur la fonction vocale*. La Pensée Sauvage Editions.
- Berko Gleason, J.; Bernstein Ratner, N. 1993. *Psycholinguistics*. U.S.A.: Harcourt Brace Jovanovich College Publishers.
- Bregman, A. 1990. *Auditory Scenes Analysis. The Perceptual Organization of Sound*. London: MIT Press.
- Carterette E.; Friedman, M. 1978. *Handbook of Perception*. London: Academic Press.
- D'Introno, Francesco. 1995. *Fonética y Fonología*. Madrid: Ediciones Cátedra, S.A.
- Fellowes, J. 1997. *Perceiving sex and identity of a talker without natural vocal timbre*. *Perception and Psychophysics*, 59, 839-849.
- Giles, H.; St. Clair, R. 1979. *Language and Social Psychology*. Oxford: Basil Blackwell.
- Pittam, J. 1994. *Voice in Social Interaction. An Interdisciplinary approach*. California: Sage Publications, Inc.
- Steven's Handbook of Experimental Psychology. Vol. 1 y 2. «Perception and Motivation». 1998. New York: John Wiley and Sons.
- Tiberghien, G. 1984. *Psychophysique*. Paris: Presses Universitaires de France.