

INTELIGIBILIDAD AUDITIVA MÚLTIPLE EN LA COMUNICACIÓN AUDIOVISUAL

Francisco Javier TORRES SIMÓN
Universidad de Sevilla

Resumen: El oído es el sentido encargado de captar, sintetizar y transmitir las ondas sonoras al cerebro, que interpreta el contenido de la masa acústica recibida. Aunque responsable del 50% de la actividad cerebral, el oído es incapaz de reconocer varios mensajes simultáneos. En esta investigación muestro un sistema de edición que, gracias a procesos de postproducción de audio basados en conceptos de música polifónica, hace posible al órgano responsable de la percepción acústica reconocer con claridad diferentes mensajes.

Palabras clave: Inteligibilidad auditiva múltiple, postproducción audio, sonido, radio, polifonía, comunicación audiovisual.

Abstract: The ear is the sense responsible to catch, synthesize and transmit the sound waves to the brain, than interpret the content of the received acoustic mass. Although responsible of the 50% of the cerebral activity, the ear cannot recognize several simultaneous messages. In this investigation I show an edition system that, thanks to processes of audio postproduction based in polyphonic music concepts, make possible to the sense responsible of the acoustic perception can recognize different messages with clarity.

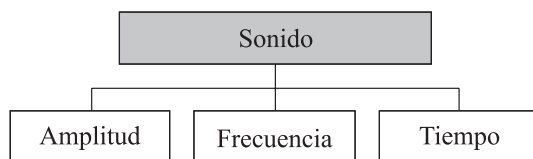
Keywords: Multi audible intelligibility, audio postproduction, sound, radio, polyphony, audiovisual communication.

La percepción auditiva

El proceso de audición consiste en la captación de variaciones de presión del aire sobre el tímpano. A dicho cambio de presión oscilante se le denomina onda sonora o sonido.

Estos cambios de presión dan una combinación de variables generada por formantes, que recoge todo el espectro sonoro. Estos formantes son:

Formantes del Sonido



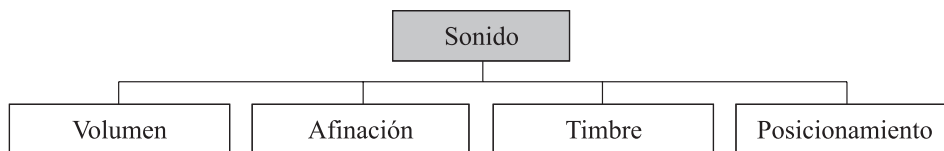
La amplitud es el formante responsable del volumen percibido por el oído. El volumen es medido en decibelios (dB).

La frecuencia es el formante responsable de la afinación de un sonido. Las frecuencias bajas o sonidos graves están formadas por cambios lentos de presión del aire, mientras que las frecuencias altas o sonidos agudos se generan por cambios rápidos. La frecuencia se mide en hercios (Hz).

El tiempo es el formante contenedor de los otros dos.

La combinación de los formantes generan identificadores gracias a los cuales se puede reconocer los diferentes sonidos. Son:

Identificadores Sonoros



El timbre es la resultante de la suma de ondas fundamentales más todos sus armónicos. A las ondas iniciales o fundamentales se les denomina F0, de donde parten todos los parciales o armónicos pertenecientes al mismo. La diferencia en amplitud de dichos armónicos determina el timbre. La onda pura por excelencia es la senoidal y sólo es posible generarla por medios electrónicos ya que en la naturaleza no existen ondas puras. Por ende, todas las ondas generadas en el medio natural estarán formadas por sumas de ondas.

El posicionamiento es la característica acústica de un sonido por la modificación de la afinación, el timbre y el volumen, según el emplazamiento y la orientación del oyente con respecto al emisor. Sumado esto a la capacidad estereofónica del oído, el oyente es capaz de identificar la procedencia de un sonido.

Aunque probada su inexistencia en el universo, dentro de los sistemas analógicos o digitales de edición de audio, el silencio es un elemento más del sonido. Puede ser cuantificado por uno de los formantes sonoros como es el tiempo.

Clasificación de masas sonoras

A diferencia de la vista, el oído es el órgano del que no podemos desconectarnos. Trabaja tanto en estado de conciencia como de inconsciencia. Es el encargado de sintetizar la masa acústica avisándonos de cuanto nos rodea para el posterior procesamiento y selección. El cerebro clasifica los elementos sonoros según los diferenciadores antes descritos.

Por ejemplo: Dos personas en el interior de un tren en movimiento. Hay muchos elementos generadores de sonido a su alrededor. Sin embargo, la conversación es posible gracias a la jerarquización establecida por el cerebro entre toda la masa sonora percibida.

En un primer nivel se encuentra el ruido generado por el vagón en movimiento. Es un sonido constante sin cambios elevados de frecuencia; un *obstinato* donde el traqueteo constante de las ruedas a frecuencias media graves sirve como manto para todos los demás elementos.

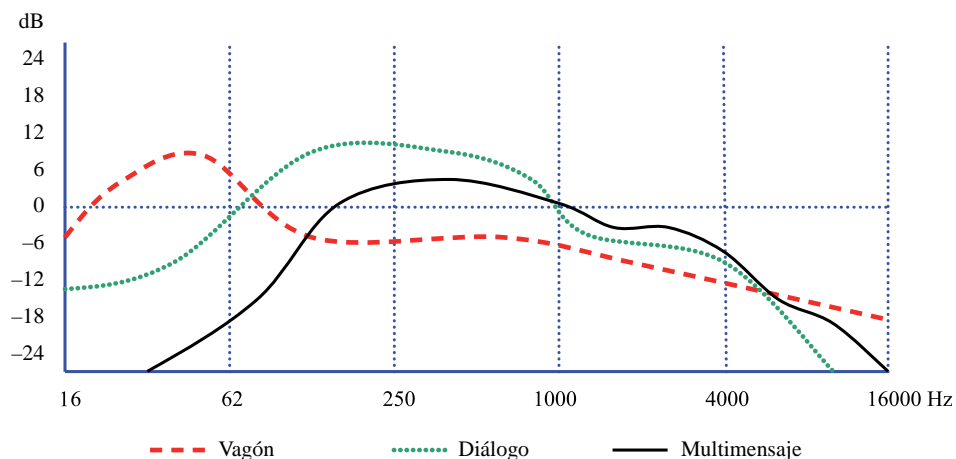
Por otro lado están las dos personas charlando. Entre ellas existe poca distancia y el cerebro ha localizado el timbre de voz. Podríamos llamarlos solistas.

Por último tenemos el resto de conversaciones y ruido flotando en el ambiente. Esta suma de diálogos indistinguibles entre sí se suele denominar Bullicio o Turba. En adelante denominaré el bullicio como “multimensaje”, definiéndolo como combinación de voces paralelas contenedoras de mensajes diferentes en igualdad de identificadores sonoros.

Al no poder distinguir el cerebro ningún mensaje distinto dentro del multimensaje, excepto cuando alguno se destaca con cierta relevancia por medio de cualquier identificador, sintetiza dicho multimensaje como un sonido unitario. Este sonido se graba en nuestro cerebro con identidad propia. El multimensaje se convierte dentro de este formato en un elemento irrelevante.

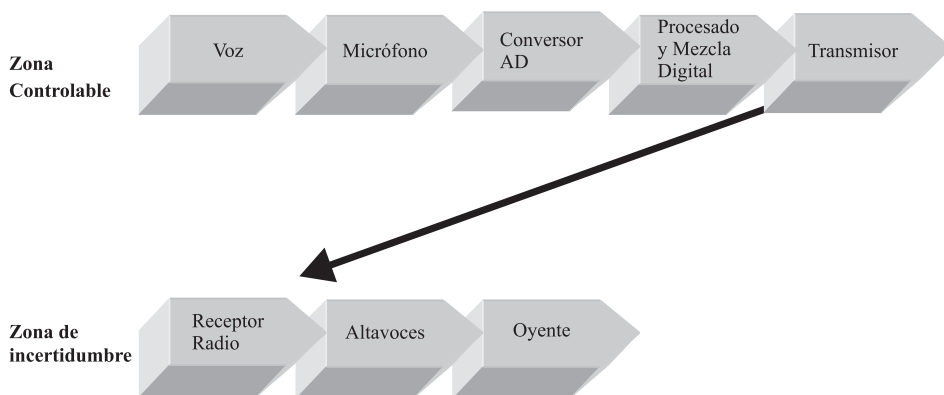
Esta característica interpretativa del oído es debida a que “dos sonidos (o más) de la misma cualidad que suenen juntos se refuerzan mutuamente formando un solo sonido” (Jonas, 2000:195).

Gráfica EQ del reparto aproximado de frecuencias del vagón.



Organigrama de emisión radiofónica

Haciendo uso de la radio como ejemplo más claro de la emisión sonora en la comunicación audiovisual, y centrándonos en los puntos más importantes dentro del interés de nuestro estudio, se puede hacer una gráfica del procesado de la voz desde su emisión hasta la recepción:



En este gráfico se muestran dos zonas de flujo de los elementos acústicos. La primera zona es aquella donde se puede interactuar para mejorar las propiedades

acústicas de las voces. La segunda zona es aquella imposible de controlar al ser casi infinitas las posibilidades existentes. Entre estas posibilidades se puede encontrar el hecho de escuchar el multimensaje con un equipo Hi-Fi o por auriculares. Por otra parte, el entorno de escucha es incontrolable ya que el oyente se puede encontrar en un lugar tranquilo, o por el contrario, encontrarse en un bullicio extremo donde aun mejorando las cualidades acústicas de las voces, sea completamente imposible entender ninguno de los mensajes. De todas formas, este modelo de postproducción está concebido primitivamente para controlar la inteligibilidad en condiciones razonables, aunque de manera adicional, se pueden aplicar paralelamente conceptos expuestos en diferentes trabajos de investigación para el aumento de inteligibilidad en espacios de acústica compleja.

La polifonía y el multimensaje

La música polifónica es un claro ejemplo de control sobre el multimensaje. Esta música vivió su época dorada entre el siglo XV y XVI con compositores tales como Josquin des Prés o Giovanni Pierluigi da Palestrina. La música polifónica organizaba los sonidos según modos y no en tonos, como la música tonal. Los dispone verticalmente en número de cuatro (tetracordos) de lo que resultan una serie de combinaciones generadoras de escalas. El arte polifónico se fundamenta en el contrapunto. El contrapunto enseña a organizar los sonidos simultáneos con sus predecesores y posteriores bajo normas estrictas. Con esta forma singular de composición, el control sobre la audición simultánea de las voces es ejemplar, pues se consigue separar con claridad cada una de las voces. Es de aquí de donde parte mi concepto de inteligibilidad auditiva múltiple.

Es cierto que el proceso de emisión de la palabra no es igual que el del canto. El ancho de frecuencia usado en la narración es ínfimo con respecto al canto. De igual manera ocurre con la dinámica. La dinámica utilizada en una comunicación lingüística será mucho menor que la utilizada en el canto.

Otra de las características que se advierte en la música polifónica procede de la distinción de voces masculinas y femeninas por sus rangos de frecuencia. De manera natural, la afinación de la voz femenina se encuentra una octava por encima de la masculina. Esto significa que la voz de la mujer oscila exactamente el doble que la del hombre para emitir sensitivamente el mismo sonido. Esto no ocurre en los procesos del habla. Las ondas fundamentales de la voz masculina oscilan aproximadamente a 100Hz mientras que las de la femenina trabajan entre 160 y 180Hz. “El rango de F0 de un locutor puede adquirir cualquier valor entre 75Hz y 300Hz” (García, 2000: 168). La frecuencia de la voz femenina sigue

siendo mayor que la del hombre pero su diferencia no es lo suficientemente notable (por el rango de frecuencia en la que trabaja) como para ser perceptible.

Por todo ello es precisamente por lo que en la música polifónica resulta más fácil distinguir las voces que integran un multimensaje.

El oído pierde capacidades de distinción en frecuencias extremas por lo que, cuanto más se aproxime a rangos de frecuencias sobre 40Hz para sonidos graves y 12kHz para los agudos, la diferenciación entre sonidos de frecuencias similares será menor que en el rango medio.

Otro rasgo diferenciador entre habla y música reposa en el concepto de organización. Los sonidos se rigen por reglas preestablecidas y concebidas como tales, tanto los sonidos lineales generados por una voz (melodías), como los sonidos generados simultáneamente por diferentes voces (armonía).

De todo ello se deduce que el multimensaje no puede ser entendible de manera natural, a no ser que sus elementos se encuentren organizados por pautas o cualquier organización similar que trabaje la diferenciación de planos acústicos.

El éxito de este sistema, que llamaremos procesador de inteligibilidad múltiple, radica en aplicar, por medio de estos procesos de postproducción, los conceptos de control de voces existentes en la polifonía, aportando a cada voz una identidad tímbrica, dinámica y espectral para que el oído pueda fácilmente reconocer a cada una de ellas por separado.

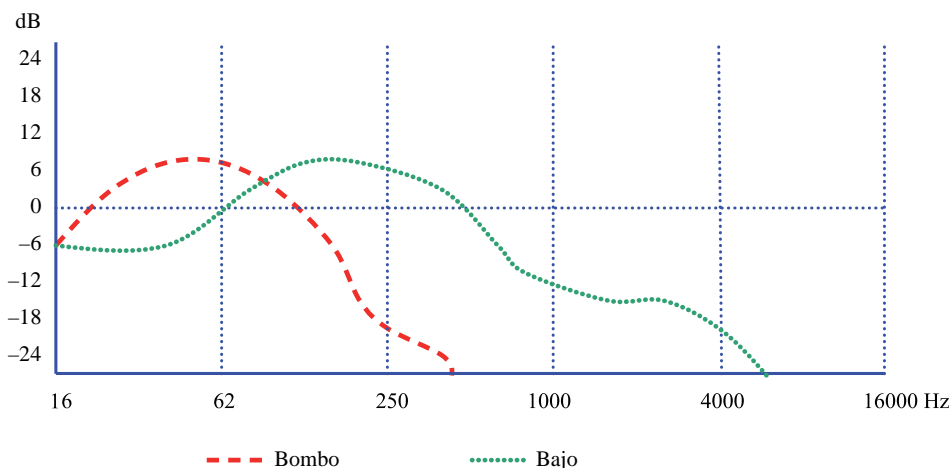
Los preliminares del sistema de inteligibilidad múltiple

Uno de los factores a tener en cuenta para garantizar el éxito del sistema de inteligibilidad múltiple es preparar las voces de manera que el procesado de señal sea lo más correcto posible para su posterior uso.

El micrófono es el primer elemento encargado de la transformación de la voz en una señal editable (sea analógica o digital). La diversidad de características entre los diferentes modelos existentes en el mercado, antes de considerarlo como una desventaja, se puede interpretar como una solución. La diversidad en respuestas hará que se puedan usar diferentes micrófonos en una misma sesión, para acentuar o atenuar ciertas características de las diferentes voces.

En la postproducción musical es habitual el uso de distintos micrófonos, incluso para un mismo instrumento. Cabe destacar el uso del modelo 414 de AKG para grabaciones de bajos o bombo de batería gracias a su gran diafragma. La tipología de micrófonos es extensa tanto por su patrón polar, tamaño de diafragma o su tecnología (condensador, dinámico, válvula o cinta).

Gráfica EQ del reparto aproximado de frecuencias de un bombo y un bajo.

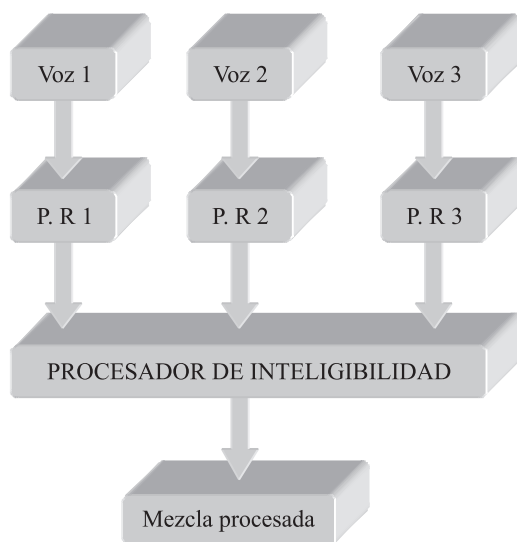


Este puede considerarse el primer factor donde crear un reparto de frecuencias para las distintas voces, al intento de conseguir, al igual que en la polifonía, una clara distinción entre la voz masculina y la femenina.

Hay que disponer por tanto de un EQ así como de un compresor clase A por canal para corregir las anomalías existentes en la voz. De igual importancia es el uso de un de-esser ya que “es un compresor de acción rápida que suele estar incorporado dentro de un compresor-limitador encargado de bajar las altas frecuencias que están presentes en los sonidos sibilantes” (Gibson, 2002a: 77).

Modelo de procesamiento de inteligibilidad múltiple del multimensaje

En el siguiente diagrama se expone un modelo de producción usando un procesador de inteligibilidad. Este modelo consta de un supuesto de tres voces.



Tal como se ve en el diagrama, las voces pasan previamente por unas puertas de ruido (Noise Gate) encargadas de activar el procesador de inteligibilidad. Esto se debe a que si una de las voces se encuentra en un punto del discurso sonando en solitario, no tendría la necesidad de ser tratada. Por el contrario, si una voz escuchada en solitario recibiera un procesado de ecualización o compresión de manera acusada, podría correr el riesgo de producir un sonido antinatural.

De esta manera, para que el procesador de inteligibilidad entre en funcionamiento es necesario que al menos dos P.R. den actividad acústica.

Formantes del Procesador de Inteligibilidad

El procesador de inteligibilidad está compuesto por diferentes procesadores acústicos comunes que, usados de manera específica, arrojan los resultados esperados. Dichos procesos son:

1. Compresión

El compresor es un procesador de dinámicas. Se usa para corregir las fluctuaciones entre volúmenes altos y bajos. Se puede configurar de múltiples formas, de manera que picos de dinámica producidos por sonidos con saltos repentinos de volumen, se ajusten homogéneamente al resto con valores preestablecidos.

Gracias al compresor, y al balance de volumen de una mesa de mezclas, es posible escuchar un pasaje de oboe en *mezzoforte* por encima de un *tutti* orquestal *fortísimo*. Estos son los contrastes de las bandas sonoras de Hollywood sólo posibles gracias a los sistemas de postproducción. Por el contrario, quedarían como ocurrencias imposibles del compositor si contemplara su ejecución por una orquesta clásica profesional al no ser viables acústicamente.

Estas propiedades características del compresor son aplicables en el sistema de inteligibilidad múltiple para corregir los planos sonoros, consiguiendo igualar las dinámicas de las voces participantes entre ellas.

2. EQ

A diferencia de la radio, en la postproducción musical es habitual la manipulación exagerada de la ecualización para la localización de instrumentos. Esto ocurre al tener que repartir el espectro sonoro entre instrumentos que comparten frecuencias. Uno de los errores frecuentes es la aglomeración de masa acústica en la zona de 40 Hz a 400 Hz entre el bombo de la batería y el bajo. Este efecto produce poca definición en el bombo y ninguna inteligibilidad por parte del bajo en las notas producidas. Para corregir el error, se asigna a cada instrumento una pequeña parcela dentro de ese rango, de tal manera que pueda oírse con mayor identidad cada uno de estos instrumentos. Entiéndase este proceso como la creación de una pequeña ventana para cada sonido dentro de una misma habitación donde sólo se les permite asomarse por la ventana asignada.

Aunque es extremadamente efectivo este proceso, si se escucha uno de estos instrumentos por separado el resultado será artificial, por lo que el uso de este tipo de EQ debe regirse por la precaución.

3. Excitador armónico

El excitador armónico es un procesador acústico muy usado en clubs y radios como un previo al masterizador. Tiene la capacidad de realzar los rangos de armónicos preestablecidos, dando brillo y nitidez así como consistencia a la totalidad del sonido. Es poco frecuente usar el excitador armónico como efecto por canal. Sin embargo, la capacidad de este procesador para matizar la tímbrica de un sonido es inigualable, ya que no modifica la personalidad acústica del sonido, sino que reestructura la combinación de los armónicos internos, haciendo que estos interactúen con el resto de sonidos de manera completamente distinta.

Un uso moderado del excitador así como un estudio detallado de sus posibilidades puede hacer que el posicionamiento espectral de cada voz sea fijada sin necesidad de retocar exageradamente la EQ.

4. Panorama

En la radio, por tradición, no se suele espaciar las voces. Posiblemente se deba a que en sus comienzos era monoaural. En la actualidad el sistema estereofónico está generalizado en los sistemas de comunicación audiovisual. Por ello, el uso moderado del panorama en coloquios o debates, puede ayudar al oyente a localizarlas. Este procesado simple de señal ayudaría al cerebro a realizar una escucha selectiva con menor esfuerzo.

La incapacidad de controlar el entorno del oyente

Considerando que no existe en el organigrama de emisión radiofónica ningún tipo de control sobre los sistemas de emisión del sonido generado; y considerando las condiciones en las que el oyente se encuentra en el momento de la recepción del multimensaje, habría que generar un sistema que funcionara de igual manera en cualquier situación, aún en las extremas, como la contaminación acústica o aislamiento.

Conclusiones

La capacidad de entendimiento del ser humano por medio de mensajes acústicos es limitada ya que la fisiología del oído impide la diferenciación de varias voces entre sí, agravándose dicha limitación cuando aumenta la distancia entre el sujeto receptor y los emisores.

Sin embargo, en un entorno controlado, donde las leyes acústicas pueden ser tecnológicamente manipuladas, se pueden aplicar conceptos de la música polifónica para obtener una masa acústica diferenciada en sus distintos mensajes, que el oído alcanza a identificar.

Bibliografía

a) Libros:

- ADLER, Samuel. (2002): *The Study of Orchestration*. New York: W.W. Norton & Company.
- AMYES, Tim. (1990): *Técnicas de postproducción de audio en video y Film*. Madrid: Instituto Oficial de Radio y Televisión RTVE.
- ANDERTON, Craig. (1980): *Electronic Projects for musicians*. New York: Amsco Publications.
- BELTRÁN MONER, Rafael. (1991): *Ambientación musical*. Madrid: Centro de Formación RTVE.
- BORWICK, John. (1996): *Micrófonos. Tecnología y aplicaciones*. Gipuzkoa: Escuela de Cine y Video.
- COPLAND, Aaron. (1994): *Cómo escuchar la música*. Madrid: Fondo de Cultura Económica de España.
- GIBSON, Bill, (2002a): *Sound Advice on Compressors, Limiters, Expanders & Gates*. Vallejo (California): ProAudio Press.
- GIBSON, Bill, (2002b): *Sound Advice on Mixing*. Vallejo (California): ProAudio Press.
- GIBSON, David, (1997): *The Art of Mixing*. Vallejo (California): MixBooks.
- GÓMEZ JUAN, Eduard y CUENCA DAVID, Ignasi. (2006): *Tecnología básica del Sonido II*. Madrid: Thomson Editores Spain, Paraninfo.
- JONAS, Hans. (2000): *El principio Vida. Hacia una biología filosófica*. Madrid, Trotta.
- JORDÁ PUIG, Sergi. (1997): *Audio Digital y MIDI*. Madrid: Ediciones Anaya Multimedia.
- KOSTKA, Stefa y PAYNE, Dorothy. (2000): *Tonal Harmony*. Boston: Mc-Graw-Hill Companies.
- MOLINA, Radamés y RANZ, Daniel. (2000): *La idea del Cosmos. Cosmos y Música en la Antigüedad*. Barcelona: Ediciones Paidós Iberia.
- MOORE, Douglas. (1988): *Guía de los Estilos Musicales*. Madrid: Altea, taurus, Alfaguara.
- NISBETT, Alec. (1990): *Usos de los Micrófonos*. Madrid: Instituto Oficial de radio y Televisión. RTVE.
- PISTON, Walter. (1984): *Orquestación*. Madrid: Real Musical.
- SCHÖNBERG, Arnold. (1974): *Armonía*. Madrid: Real Musical.
- SHEA, Mike. *Studio Recording Procedures*. (2005): *How to record any instrument*. Columbus: McGraw-Hill Companies.
- ZAMACOIS, Joaquín. (1997): *Tratado de armonía*. Cooper City (Florida): SpanPress Universitaria.
- RUMSEY, Francis y McCORMICK, Tim. (2004): *Sonido y grabación. Introducción a las técnicas sonoras*. IORTV.

b) Artículos de revistas científicas:

TAPIAS MERINO, Daniel / GARCÍA, Carlos. (2000): “La frecuencia fundamental de la voz y sus efectos en el reconocimiento de habla continua”. *Procesamiento del lenguaje natural*, nº 26, 2000 pp. 163-168.